



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN FILOSOFÍA DE LA CIENCIA // INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

FILOSOFÍA DE LAS MATEMÁTICAS Y LÓGICA DE LA CIENCIA

UN MODELO NO MONOTÓNICO Y PARACONSISTENTE

DE EXPLICACIÓN CIENTÍFICA

TESIS

QUE PARA OPTAR POR EL GRADO DE DOCTOR EN FILOSOFÍA DE LA CIENCIA

PRESENTA:

DAVID GAYTÁN CABRERA

DIRECTOR DE TESIS: DR. J. RAYMUNDO MORADO ESTRADA, IIF-UNAM.

**COMITÉ TUTOR: DRA. ÍTALA MARÍA LOFFREDO D'OTTAVIANO, POSGRADO EN FILOSOFÍA DE LA CIENCIA;
DR. AXEL ARTURO BARCELÓ ASPEITIA, IIF-UNAM; DR. MAX FERNÁNDEZ DE CASTRO, POSGRADO EN
FILOSOFÍA DE LA CIENCIA; DRA. MARÍA INÉS PAZOS, POSGRADO EN FILOSOFÍA DE LA CIENCIA.**

MÉXICO, D.F. abril 2014.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mis padres, María Luisa Cabrera y René Gaytán

La tentativa de determinar el universo mediante principios a priori ha fracasado; la lógica, en lugar de ser, como antes, una barrera para las posibilidades, se ha convertido en la gran liberadora de la imaginación, presentando innumerables alternativas cerradas al sentido común irreflexivo y dejando a la experiencia la tarea de decidir, cuando la decisión es posible, entre los varios mundos que la lógica ofrece a nuestra elección.

Bertrand Russell, Los Problemas de la Filosofía, 1912.

AGRADECIMIENTOS

El núcleo de la presente investigación doctoral lo concebí en el seno de los estimulantes seminarios impartidos por Raymundo Morado en los años 1995 y 1996, en el Instituto de Investigaciones Filosóficas, en el marco del *Posgrado en Filosofía de la Ciencia* de la UNAM. La solución principal para la elaboración del modelo, la forma en que combiné la Lógica del Razonamiento por *Default*, de Raymond Reiter, y las jerarquías de cálculos C_n , $1 \leq n \leq \omega$, de Newton da Costa, la propuse en esos años, y presenté sucesivas versiones del sistema en diversos coloquios y congresos a partir de entonces. Las dificultades que tuvo que enfrentar el sistema entonces propuesto fueron de índole filosófica tanto como de adecuación lógica.

Un motor claro de mi propuesta fue el intento de solución para el problema hempeliano de las salvedades, presentado en mi tesis de maestría. Pero esto me llevó naturalmente a proponer una salida a los contraejemplos paradigmáticos contra la modelación clásica de la explicación científica, con dos motivaciones remotas principales: 1) mostrar la potencialidad de las lógicas no clásicas para la modelación formal útil de ciertos aspectos de la Ciencia, y 2), estrechar vínculos entre dos áreas de investigación aparentemente distantes, a saber, la Filosofía de la Ciencia y la Inteligencia Artificial. Pero la tarea que me propuse, implicaba integrar soluciones a muy diversos problemas filosóficos y lógicos, y tenía consecuencias sobre el nivel de complejidad de una propuesta formal que pudiera representar tan diversos aspectos de la explicación supuestos en tales problemas. Las soluciones normales a una modelación formal de la explicación, a través sobre todo de estudios relacionados con la abducción, no atendían la complejidad del problema tal como se ponía éste sobre la mesa de discusión en el debate al interior de la Filosofía de la Ciencia. Así, la elaboración del marco filosófico que serviría de contexto para que la solución formal tuviera sentido, aunque en gran parte se llevó a cabo en esos años y hasta 1998, fue madurando a partir de entonces desde una posición no deductivista y epistémica a una solución que, integrando las cualidades anteriores, se convirtió en una posición claramente pragmática y minimalista acerca de la modelación de la explicación. También, los detalles para construir la propuesta formal que integrara bien todos estos elementos tuvieron que madurar a la luz de las orientaciones siempre extraordinariamente lúcidas de parte de mi director de tesis, Raymundo Morado, quien puso siempre amablemente a disposición su gran sensibilidad filosófica y sus contraejemplos directamente encaminados a reflexiones fructíferas sobre racionalidad y Lógica.

Debo agradecer la principal orientación teórica de la tesis que ahora presento, a las disquisiciones de Bas van Fraassen y de Michael Scriven, sobre la explicación. Pero, de un modo más cercano, también a Eduardo Flichman, por mi discusión con él sobre el problema de las salvedades que Hempel propuso; a Larry Laudan, por el modo abierto en que examinó mi propuesta de modelación lógica para solucionar un problema relacionado con su propia noción de evidencia; y con grande estimación a Raúl Orayen, a quien considero el pilar por excelencia de la tradición lógica mexicana. Mi discusión con Orayen sobre uno de sus artículos (que considero muy notable) donde propone un problema contra la noción deductiva hempeliana de explicación, fue clave en mi orientación hacia una solución intensionalista y pragmática al problema de la modelación de la explicación, aunque quizá él no hubiera estado de acuerdo con algunos detalles de mi solución formal final.

Para la versión final de mi propuesta fue sumamente importante la intervención de Ítala D'Ottaviano, de la *Universidade Estadual de Campinas, UNICAMP, SP, Brasil*, propuesta que en realidad se convirtió, de *facto* y con toda la intención de mi parte, de la suya y de mi director oficial de la tesis, en una co-dirección de mi trabajo doctoral. En ese aspecto, debo mencionar todo lo que aprendí con ella en mi estancia doctoral en el *Centro de Lógica, Epistemología e História de la Ciencia, CLE, de la UNICAMP, SP, Brasil*. Además, agradezco los cursos de posgrado tomados con Walter Carnielli y con Marcelo Coniglio en esa época, a quienes directamente agradezco mucho también, pues ayudaron indirectamente a la forma que cobró la presentación de mi trabajo final. A Carnielli agradezco mucho sus preguntas provocadoras que ponían en cuestión la utilidad de la modelación no monotónica. A Coniglio, su paciencia para hacernos comprender las demostraciones y su manera ingeniosa de pensar las estrategias para ellas. Por supuesto, en esa estancia mi directora de investigación fue Ítala D'Ottaviano, a quien debo agradecer en particular una revisión detallada de toda la tesis y una discusión siempre alentadora y profundamente aguda sobre los pormenores formales, lo que por un lado me permitió una comprensión precisa de los rumbos por los que habría que construir las soluciones formales del sistema que se propone para la modelación de la explicación y, por otro, a la postre, contribuyó determinantemente a mi formación lógica.

Así, la presente tesis debe mucho a la fortuna de haberme encontrado con personas muy inteligentes que decidieron discutir mis ideas. En este marco, agradezco mucho a Julián Molina, su discreto pero agudo sentido lógico; a Alicia Pazos, su nítida abstracción conceptual y su rapidez de pensamiento; a Abel Hernández, su extraordinaria creatividad para la interdisciplina; a Agustín Rayo, su entusiasta y lúcida charla; y a Inés Pazos, *Universidad Autónoma de la Ciudad de México, UACM*, sus profundas clarificaciones argumentativas y sus críticas. También agradezco mucho la lectura

cuidadosa e inteligente de mi trabajo, que hizo Axel Barceló, como miembro del comité, y su charla siempre amistosa, llena de profundas conexiones lógico-filosóficas. Fue imprescindible la lectura detallada, amable, con profunda sensibilidad formal y clarificadora, de Max Fernández de Castro, de la *Universidad Autónoma Metropolitana, Iztapalapa, UAM-I*, a quien agradezco mucho. En esta misma línea de felices encuentros en mi vida académica estoy agradecido en especial por la fortuna de contar con la ayuda y la amistad enriquecedora de uno de los principales expertos en No Monotonidad en Latinoamérica, que tiene además un gran conocimiento de la Filosofía, Raymundo Morado. Y también debo mucho a la suerte de contar con la guía, la amistad y el pensamiento agudo y entusiasta de una de las principales pensadoras contemporáneas en Lógica Paraconsistente, Ítala D'Ottaviano.

En relación con mi estancia en Brasil, estoy muy agradecido también por las discusiones que mantuve en *el Seminario de Lógica*, y en el seno del *Grupo de Lógica Teórica y Aplicada, GTA*; ambos del CLE. En específico, al apoyo siempre paciente recibido por parte de Anderson de Araujo, como adjunto en una clase de Carnielli, y a la charla de mis ideas con algunos que eran entonces estudiantes del posgrado en el CLE, como Ramón Souza, Juliana Bueno-Soler, Roberto Dumaresq, Dante Cardoso, Leandro Suguitani y Samir Gorsky.

Finalmente, agradezco mucho a Verónica Gaytán Cabrera, por todo el trabajo y esfuerzo con el que me apoyó en toda la serie de trámites sin los cuales la titulación no hubiera podido llevarse a cabo.

INDICE

Agradecimientos	3
Introducción General	10
I.1. Modelar	11
I.2 Modelar la Explicación	13
Capítulo 1. Algunas discusiones acerca de los modelos clásicos de explicación científica.	28
Introducción	28
1.1. La Propuesta Hempel-Oppenheim	29
1.1.1.El modelo D-N	30
1.1.2. El Modelo I-S y las Diferencias entre ambos Modelos	36
1.2. La Tensión Hempel-Scriven	44
1.3. La Tensión Hempel-Coffa	56
1.4. Otros Modelos de Explicación Científica	62
1.4.1.Salmon	64
1.4.2.Railton	70
1.4.3.Van Fraassen	78
Capítulo 2. Los problemas clásicos para la modelación argumental de la explicación: inferencia y pertinencia explicativa.	86
Introducción	86
2.1.Desarticulando dos problemas clásicos	87
2.1.1.El Supuesto Argumental	87
2.1.1.1.El problema	87
2.1.1.2.El análisis del problema o el Supuesto de Alto Grado de Certeza	94

2.1.2.La ambigüedad epistémica	101
2.1.2.1.La reacción estándar	101
2.1.2.2.Sugerencia de una nueva perspectiva	108
2.1.3. Supuesto Argumental y No Monotonidad	114
2.1.3.1.La deducción en las propuestas de explicación	114
2.1.3.2.Un modelo argumental	119
2.1.3.3.Un modelo No Monotónico	125
2.1.4.Las relaciones de pertinencia	128
2.1.4.1.Las asimetrías	129
2.1.4.2.El supuesto causalista de la explicación	137
2.1.5.Causalidad, contexto y explicación	153
2.1.5.1.Dos problemas de la causalidad en la explicación	153
2.1.5.2.La noción de relación causal	166
Capítulo 3. El contexto y la modelación inferencial de la explicación científica: paraconsistencia y no monotonicidad.	184
Introducción	184
3.1.Explicación en contextos inconsistentes	184
3.1.1.La explicación de Frege	184
3.1.2.El <i>explanans</i> inconsistente: la paradoja de Russell	187
3.1.3.Implicaciones: T. de Conjuntos y Matemáticas	189
3.1.4.Racionalidad e Inconsistencia en la modelación de la explicación	195
3.1.5.Un sistema Paraconsistente: Los Cálculos Cn	200
Introducción	200
3.1.5.1. El sistema de da Costa	203
3.1.6.Paraconsistencia y Racionalidad	206

3.1.6.1. La vía modal	207
3.1.6.2. La vía sintáctica	209
3.1.6.3. Algunas Consecuencias	214
3.2. El razonamiento del sentido común	216
3.2.1. Algunas intuiciones sobre el razonamiento del sentido común	217
3.2.2. Una introducción informal a la lógica del razonamiento por <i>default</i>	231
3.3. Explicación e ignorancia	247
3.3.1. No monotonicidad en la Ciencia	255
3.3.2. No Monotonicidad y relación de pertinencia explicativa	270
3.4. Algunas reflexiones sobre la lógica <i>default</i> y la no monotonicidad	272
3.4.1. Inteligencia artificial y explicación	273
3.4.2. No monotonicidad	276
3.4.3. El análisis interno	287
3.4.3.1. La regla <i>default</i>	287
3.4.3.2. Prerrequisito y salvedad	291
3.4.3.3. Prerrequisito y conjetura	297
3.4.3.4. Inconsistencia entre salvedades	300
3.4.3.5. Certeza	303
3.4.4. El análisis externo	304
3.4.4.1. El contexto de un <i>default</i>	304
3.4.4.2. Inconsistencia en la extensión	308
3.4.5. No monotonicidad y transitividad	310
Capítulo 4. Un modelo argumental de explicación científica	320
4.1. Un modelo paraconsistente y no monotónico de explicación	320
4.1.1. El contexto de esta propuesta formal	320

4.1.2. Algunas asunciones filosóficas acerca de la explicación, que se hacen en el modelo	323
4.1.3. Un Marco Teórico Formal para el Análisis de las Explicaciones	325
4.1.3.1. Sistemas Epistémicos	326
4.1.3.2. Construyendo una <i>Teoría Default</i>	330
4.1.3.3. Una Teoría <i>default</i> paraconsistente	335
4.1.3.4. El modelo de explicación	342
4.2. El modelo y la explicación	353
Introducción	353
4.2.1. Pertinencia explicativa	354
4.2.1.1. Asimetrías	354
4.2.1.2. Explicaciones y relaciones de explicación	359
4.2.1.3. Causalidad	362
4.2.2. Ambigüedad epistémica y no monotonidad	366
4.2.3. Representación de la ignorancia	368
4.2.3.1. <i>Explanans</i> postulado	368
4.2.3.2. <i>Explanandum</i> falible	380
4.2.3.3. Explicación dinámica	397
4.2.4. Paraconsistencia	408
4.2.4.1. Inconsistencia en el sistema epistémico	411
4.2.4.2. Inconsistencia y explicación	421
Conclusiones	435
Referencias Bibliográficas	444

INTRODUCCIÓN GENERAL

La explicación es un elemento dinámico de la actividad científica. La relevancia de una explicación particular se encuentra en relación directa con información proveniente de un contexto teórico. En este sentido, un análisis formal sobre la estructura general de la explicación parece ser una empresa muy compleja. Algunos de los elementos que contribuyen a esta complejidad y que son especialmente interesantes son los siguientes:

- 1) Algunas veces el *explanans* depende de teorías inconsistentes, o de teorías inconsistentes entre sí, y a pesar de ello nadie toma las teorías o el *explanans* como lógicamente triviales.
- 2) Parece poder suponerse, aunque sólo sea para fines modelativos, una relación inferencial entre el *explanans* y el *explanandum*; esta relación generalmente no es deductiva.

El problema 1 es más bien novedoso pero el problema 2 junto con otros más que examinaremos después con brevedad han hecho muy difícil la tarea de desarrollar un análisis formal de la explicación. Es posible que estos obstáculos sean librados mediante el uso de lógicas no clásicas, particularmente lógicas paraconsistentes y lógicas no monotónicas. Esta perspectiva conforma el núcleo de la propuesta para la elucidación de la explicación científica en la presente tesis.

No obstante, antes de internarnos en una reflexión que nos lleve directamente sobre los detalles de la explicación que podrían justificar la introducción de herramientas de lógica no clásica para su análisis, es útil aclarar algunos puntos respecto del papel de la modelación en términos generales. A continuación presentaré algunas características muy generales de la modelación, que tomaré luego como punto de partida para dejar por sentado algunos aspectos en relación con la perspectiva desde la cual se erige mi investigación.

I.1 Modelar

En cierto sentido, caracterizar lo que normalmente entendemos por "modelación" es algo que puede equipararse con la caracterización de un tipo de explicación o, más específicamente, con la caracterización de una subclase de las explicaciones que pueden verse como respuestas a preguntas-cómo. Algunos autores han enfatizado que las explicaciones pueden verse como respuestas a preguntas. Ello nos permite identificar, al menos de forma práctica, algunos tipos de explicación: la "explicación-qué", la "explicación-cómo" y la "explicación-porqué" según sea que encajen en un determinado tipo de cuestión como se sugiere por los nombres. Así, modelar a un objeto x puede verse como un estilo para explicar cómo es x . Consecuentemente, un modelo de la explicación de esta clase sería cierta especie de "metaexplicación-cómo" de la explicación.

Como he estado tratando de usar el término "modelación", el sistema que conforma al modelo es, en cierto sentido, es una cosa de tipo epistémico, en particular, una representación. Debe ser entendido como una construcción semántica capaz de ocupar lugares en la articulación del conocimiento, del mismo modo en que una teoría en su sentido semántico puede hacerlo. La construcción de un modelo no necesariamente tiene que limitarse a las relaciones internas del objeto que se pretende modelar; en

ocasiones es necesario vincular ciertas relaciones internas con algunas relaciones que ocurren en el entorno del objeto a modelar. Por lo menos en la discusión en Filosofía de la Ciencia, a lo largo de la segunda mitad del siglo XX, en la modelación de la explicación muchas veces se asumen también algunas relaciones externas como la situación contextual o las teorías aunque frecuentemente en estos casos las relaciones externas no quedan modeladas con precisión. De esta forma la modelación debería entenderse como la construcción de un sistema de representación que simule las relaciones dentro de un sistema (susceptible de incluir el entorno) que es considerado pertinente para la elucidación del objeto de estudio.

Hasta ahora he dicho solamente que no todo tipo de teorización consta de la construcción de un modelo simulador. No pretendo defender alguna especie de calidad superior de elucidación entre la teorización y la modelación de algo, y considero que la modelación es sólo otra alternativa de investigación o, mejor dicho todavía, un estilo de teorizar. De cualquier modo es pertinente mencionar algunos otros detalles que relacionan más directamente la teorización con la modelación.

Teorizar para elucidar x , donde x puede ser un individuo o una clase de individuos, puede consistir en una de las tres clases siguientes de actividad: 1) proporcionar un conjunto de características que sea necesario y no suficiente para saber que tenemos un x , 2) proporcionar un conjunto de características que sea suficiente y no necesario para saber que tenemos un x , 3) proporcionar un conjunto de características que sea suficiente y necesario para saber que tenemos un x . En el primer caso, tenemos elementos para determinar cuándo algo es susceptible de ser un x , es decir, tenemos características inherentes a x y por tanto en ocasiones sabemos definitivamente que algo no es un x ; sin embargo no podemos estar seguros si los individuos y que cumplen con esas características son realmente miembros de la clase de los x y tenemos casos de individuos z tales que no pueden ser identificados (aunque deberían serlo) como no pertenecientes a la clase de los x . A este caso lo llamaremos "acotación necesaria de x ". En el segundo caso,

que llamaremos "acotación suficiente de x", tenemos elementos para saber definitivamente que estamos ante un x; sin embargo no podemos estar seguros de que, si algo no cumple con estas características, no es realmente un miembro de la clase de los x; por tanto, quizá haya algunos z que no puedan ser identificados (aunque deberían serlo) como pertenecientes a la clase de los x. En el último caso, tenemos elementos para negar que algo es un x y elementos para identificar un x; lo llamaremos "biacotación de x".

1.2 Modelar la Explicación

En el caso de la modelación de la explicación científica podría parecer que existe un problema intrínseco en el intento por realizarla utilizando herramientas lógicas. La intuición que subyace a esta afirmación es correcta en un cierto sentido pero incorrecta si se toma en un sentido absoluto. Una de las motivaciones principales del presente trabajo de investigación es, precisamente, mostrar el doble aspecto de esta intuición mediante un ejemplo concreto de modelación.

El pensamiento de Hempel suele ponerse como ejemplo paradigmático de la propuesta de modelación de la explicación que se encuentra fundamentada en herramientas lógicas. Como veremos en detalle, las críticas que, en relación a la inferencia, se esgrimen contra esta clase de modelación pueden clasificarse analíticamente en tres tipos principales:

- 1) Las que atacan la necesidad de la inferencia lógica.
- 2) Las que atacan la suficiencia de la inferencia lógica.
- 3) Las que rechazan la asunción argumental.

No expondré ahora detalladamente cada una de estas críticas pues esto supondría un análisis muy elaborado para una introducción pero sí mencionaré algunos puntos importantes para mostrar la línea de trabajo que desarrollaré después en el capítulo I.

Tomaré como eje de la exposición al modelo nomológico-deductivo (D-N) de Hempel aunque críticas similares pueden hacerse respecto de su modelo inductivo-estadístico (I-S). Para ello sólo es importante tener claro, entonces, los siguientes puntos:

- i) cuando A y e se encuentran en una relación explicativa se dice que A es el *explanans* y e el *explanandum*,
- ii) los modelos D-N e I-S de explicación suponen que las explicaciones pueden modelarse mediante argumentos de una cierta clase, a este supuesto me refiero cuando digo "supuesto argumental" (SA),
- iii) muy simplificada, un conjunto de premisas A explica a e si y sólo si ocurre que de A se infiere e.

La línea de argumentación en contra de la necesidad de la inferencia lógica para un modelo de explicación científica está basada en contraejemplos que tengan dos características: A) que sean considerados intuitivamente como explicaciones y B) que el *explanans* no confiera apoyo deductivo al *explanandum*. De hecho, cualquier *explanans* que confiera sólo probabilidad, alta o baja, al *explanandum* es un contraejemplo perfecto contra el modelo D-N de explicación. Luego mostraré cómo esta crítica toma un carácter más relevante que está vinculado con el tipo de crítica del inciso 3.

Respecto de la crítica contra la suficiencia de la inferencia lógica los contraejemplos deben tener dos propiedades: \neg B) que el *explanans* confiera apoyo deductivo al *explanandum* y \neg A) que no sean considerados intuitivamente como explicaciones. Como se ve, estos contraejemplos son argumentos deductivos que, sin embargo, no pueden ser considerados explicaciones intuitivas. Los casos presentados

generalmente son casos que ejemplifican el llamado "problema de la asimetría", que explicaremos en el capítulo I.

Hempel admite que puede haber casos de inferencia probable; lo manifiesta explícitamente y ello le sirve de antecedente para la construcción de su modelo I-S de explicación. Por lo tanto los casos de inferencia por alta probabilidad encontrados como contraejemplos a la necesidad de la inferencia lógica no serían en realidad ejemplos relevantes contra la propuesta de Hempel, pues Hempel diría que pertenecen a otra clase de explicación, a saber, la explicación I-S, que se encuentra fundamentada sobre un tipo de inferencia en la que el *explanans* le confiere alta probabilidad al *explanandum*. Trivialmente, dichas consideraciones explican que los críticos de los modelos hempelianos centren su atención en los casos de inferencia mediante baja probabilidad más bien que en los de alta probabilidad. Los casos de inferencia por baja probabilidad son verdaderos desafíos para la propuesta modelativa de Hempel.

El tipo de crítica contra la asunción argumental está relacionado con lo que aludí en el párrafo anterior. Entiendo "Supuesto argumental" (SA) como el supuesto de que la explicación es susceptible de ser modelada mediante argumentos.

Demostrar, mediante contraejemplos, la no necesidad y la no suficiencia de la inferencia lógica, constituye una refutación contundente para las pretensiones del modelo D-N de explicación. La crítica es concluyente si consideramos a la inferencia lógica como característica modeladora, es decir, la crítica afecta al modelo D-N en tanto que fracasa su pretensión de tener, con este tipo de inferencia, conformada fundamentalmente una biacotación de la explicación.

Sin embargo la clase de inferencia que está siendo anulada en esta crítica es la inferencia deductiva. La inferencia probable podría entonces seguir apoyando SA. La inferencia deductiva puede ser considerada como la inferencia límite más fuerte, en el

sentido de que las cosas que son derivadas deductivamente son derivadas suponiéndolas con un grado de certeza absoluto. Cuando derivo q a partir de la suposición de la verdad de p implica q y de p , estoy concluyendo la verdad de q sin lugar a dudas. Ahora bien, si la crítica es extensible a los casos de inferencia por alta probabilidad, por ejemplo mediante casos concretos de explicaciones en las que el *explanans* confiere sólo una baja probabilidad al *explanandum*, parece evidente que la relación inferencial no tiene nada que ver con la relación explicativa puesto que no podríamos construir inferencias basadas en una baja probabilidad. De ser así, los argumentos serían en gran medida irrelevantes para la modelación de la explicación.

Las críticas que he esbozado se han ido conformando históricamente quizá desde el surgimiento de la tensión Hempel-Scriven que examinaremos en el capítulo I. Hay otras críticas en este mismo respecto que no son menos importantes pero que desarrollarlas aquí requeriría, aún en el caso de que lo hiciera muy abstractamente, de una desmedida extensión para la introducción de esta tesis. Sólo mencionaré que dos de ellas, las cuales van al corazón de nuestra posibilidad de modelación lógica de la explicación, se deben a Raúl Orayen por una parte y al mismo Hempel, por otra.

En la tesis de maestría exploré una aplicación del formalismo de la Lógica del Razonamiento por *default*, de Raymond Reiter, para solucionar el problema hempeliano de las salvedades. Dicho problema, como yo lo interpreto en esa investigación, es básicamente la demostración que hace Hempelⁱ de que no se pueden representar formalmente ninguna de las 3 funciones básicas de la ciencia, a saber, el ascenso inductivo, la inferencia teórica y las explicaciones. La razón fundamental es que existen a menudo salvedades a las inferencias, las cuales no se pueden determinar anticipadamente y no se pueden formalizar para una inferencia eficiente de acuerdo a los objetivos pragmáticos de una representación de este tipo para la ciencia. En ese trabajo, hice notar el enorme parecido del planteamiento hempeliano, con el planteamiento que hace Reiter, al presentar uno de los problemas básicos para la representación del razonamiento del

sentido común. A partir de esto, propuse que si eliminamos la perspectiva deductivista desde la cual está Hempel planteando el problema, y partimos por lo tanto más bien de un enfoque epistémico de representación y no, como lo propone él, de uno ontológico, podemos resolver este problema de representación. Así, propuse una representación inferencial de la propuesta epistémica de una inferencia en ciencia y no de la relación ontológica que guarden efectivamente los fenómenos en cuestión en cada una de las 3 funciones de la ciencia aludidas por Hempel. Plantearé ahora en este trabajo de investigación que esta estrategia que pasa, básicamente, de un planteamiento deductivista de representación a uno que es no deductivista, podría desarticular los problemas que suponemos que hay para una modelación argumental de la explicación científica. Pero es un cambio que también supone una debilitación de nuestros criterios de racionalidad para el tratamiento de estos aspectos de la Ciencia.

La representación que hice en ese trabajo de maestría estuvo basada en el esquema propuesto por Reiter para una regla *default* para la representación del razonamiento sentido común. La idea básica fue representar las salvedades hempelianas como las justificaciones de una regla *default*. Estas justificaciones pueden bloquear la inferencia cuando, dicho de manera simple, se genera una inconsistencia en las extensiones construidas mediante las reglas de este tipo. De este modo, pueden representarse las salvedades pero no sólo eso, sino también el tipo de interacción que tienen con la expansión de una teoría a partir de supuestos básicos, condiciones iniciales, y reglas.

Una estrategia del tipo, que usa herramientas que sea han aplicado con éxito en Inteligencia Artificial, y que he descrito en el párrafo anterior, puede dar cuenta de los problemas sobre la inferencia que han sido planteados históricamente en la Filosofía de la Ciencia, en contra de los modelos clásicos de explicación de Hempel.

Sin embargo, como lo dejé claro también en mi tesis de maestría, los problemas contra los modelos clásicos, por lo menos los problemas que están vinculados con la intención de construir un modelo formal de explicación científica, van más allá de la debilitación de la noción de inferencia, la Filosofía de la Ciencia se ha encargado de plantear cuestiones aún más difíciles. Si bien es cierto que esta debilitación ya coloca el problema en una situación diferente y abre la posibilidad para nuevas formas de modelación lógica de algunos aspectos de la Ciencia, hay algunos problemas que seguirían pendientes y que son quizá, desde un punto filosófico, más difícilmente abordables desde un planteamiento lógico. En el análisis de los contraejemplos dados en la Filosofía de la Ciencia alrededor de la noción de inferencia, uno puede descubrir, paralelamente, otra serie de problemas que pueden agruparse en dos conjuntos. El primero de ellos es un conjunto de problemas relacionados con la noción de contexto. Al parecer una modelación más efectiva de la explicación científica requeriría la inclusión del contexto y esto significa tener una representación del contexto en un marco inferencial no deductivo. Pero una vez que incluimos el contexto, surge la clase de problemas mencionados apenas mencionados al principio de esta introducción. Uno de los problemas que surgen es que a veces tenemos contextos en los que una explicación parecería generar contradicciones. Así, esto no solo implica la necesidad de representar relaciones inclusive no deductivas en diferentes contextos, sino que también la necesidad de que nuestro formalismo pueda manejar contextos inconsistentes. El segundo, un conjunto de problemas que aglutina algunas cuestiones alrededor de la noción de pertinencia explicativa, o de lo que pueda llamarse la relación de relevancia supuesta en las conexiones referidas en una explicación.

Así, en el abordaje del problema de la posibilidad de una modelación formal de la explicación científica, podemos distinguir tres clases de problemas íntimamente relacionados:

- 1) Problemas conectados con la relación de inferencia supuesta en el modelo.

- 2) Problemas implicados por la inclusión del contexto en un modelo de explicación científica.
- 3) Problemas para la determinación y modelación de lo que puede llamarse la relación de pertinencia explicativa en la explicación.

Como anticipé en mi trabajo de maestría, el abordaje de estas tres clases de problemas, hace muy difícil la modelación de la explicación, y lanza a estas tentativas, más allá de una propuesta formal que de cuenta de la relación de inferencia, más bien a la propuesta de una perspectiva diferente para abordar el problema de la modelación de la explicación. Nos invita a abrazar una perspectiva no deductivista y epistémica del problema de esta modelación. Una perspectiva no deductivista ya posibilita la desarticulación de las críticas paradigmáticas emitidas contra los modelos clásicos. Pero, en la discusión filosófica sobre el problema, se fueron dejando claras algunas otras dificultades más profundas que implican un cambio de perspectiva más radical respecto de la modelación de la explicación, uno que nos lleva a una representación pragmática de cómo funciona la explicación en relación con diversos contextos, incluyendo contextos inconsistentes. Para ello intentaré construir un modelo mínimo de explicación que se centre en cómo funciona la interacción de sus elementos internos y la interacción entre la explicación y elementos externos, como el contexto y las lógicas que asumimos como subyacentes a estos contextos. Será una biacotación sí, pero del caso mínimo de explicación científica. Pero será también una acotación necesaria también sólo respecto de la intención de determinar el tipo de argumento que uno puede asumir que podría capturar lo que llamamos explicación, respecto de sus funciones internas y externas.

Así, en el presente trabajo intentaré un abordaje de las tres clases de problemas que nos vaya introduciendo a una perspectiva diferente para abordar el problema general de la modelación. La perspectiva que finalmente propondré será no deductivista, epistémica y pragmática. No deductivista en el sentido de que no nos importarán valores altos de certeza en la inferencia para esta modelación. Epistémica en el sentido de que la

explicación será entendida como una propuesta teórica de elucidación que es falible, y que por tanto no tiene porqué atender las expectativas difíciles de alcanzar que se ven supuestas en parte de las discusiones filosóficas sobre el punto. Y es pragmática en el sentido de que el modelo que propondré intentará rescatar el sentido de la explicación que normalmente aceptaríamos desde el punto de vista del sentido común, para lo cual se propondrán modelos de razonamiento del sentido común para elucidar las explicaciones de tipo científico.

Haré esto desarrollando la discusión a través de lo que llamaré la tensión Hempel-Scriven y la tensión Hempel-Coffa. La articulación de estas dos tensiones puede servirnos como dos postes indicadores que nos lleven a identificar dos tendencias en la racionalidad que se supone a la hora de intentar elucidar la explicación científica.

Esta forma de abordar el problema de la modelación de la explicación se encuentra relacionado con algunos problemas generales sobre la racionalidad científica.

Un cambio de perspectiva en esta dirección supone una ruptura con la asunción de la lógica clásica como lógica subyacente de la racionalidad por excelencia. Es necesario, sin embargo, tomar en cuenta que dicha ruptura con la asunción de la lógica clásica como lógica subyacente no fue producida por primera vez por la lógica paraconsistente. Más bien pareciera que siempre estuvo implícita la divergencia en algunos pensadores importantes al menos desde la modernidad y no sólo en el análisis de la explicación. La ruptura se reinauguró con críticas de la historia de la ciencia a las posiciones más estándares en Filosofía de la Ciencia. Críticas que, precisamente, fueron concretadas en los años cincuentas y sesentas. Podríamos hablar de las posiciones de la línea sociologista dura de la Filosofía de la Ciencia, pero es suficiente sólo con mencionar una posición más moderada: la de Thomas S. Kuhn; quien, a fin de cuentas, marcó con su sello la discusión posterior en esta disciplina hasta nuestros días.

En realidad, sólo hay dos tesis suficientemente enfáticas involucradas con nuestra noción de racionalidad, basada en la Lógica Clásica, que son enunciadas por Kuhn y después repetidas por diversos filósofos posteriores a la aparición de "La Estructura de las Revoluciones Científicas" (1962):

(i) La imposibilidad de encontrar algoritmos que describan los procesos de la actividad científica.

(ii) La negación del sentido acumulativo del progreso científico.

[En la versión del F.C.E., México, 1991, sobre todo en Caps. IX y XIII, (para acumulatividad ver pp. 149-175)]

(i) señala la inadecuación de modelos de racionalidad basados sólo en propiedades que son armónicas con la Lógica Clásica, va contra la noción de inferencia clásica entre otras cosas; (ii) puede verse como en contra de la monotonicidad de la noción de consecuencia clásica como una noción que refleje el progreso científico. Ambas afirmaciones golpean el sentido "duro" de racionalidad científica que había sido todavía defendido por Popper y heredado a él desde el positivismo lógico. Golpean la tendencia de colocar la Lógica Clásica como única lógica subyacente de la racionalidad científica. No obstante, (i) y (ii) no atacan ni a la relevancia de un análisis formal de los aspectos de la actividad científica ni, mucho menos, van en contra del supuesto de un progreso racional en la ciencia.

Empero, debido al carácter agudo con que fue hecha la crítica aludida, se identificó el error cometido por algunas de las posiciones estándares, con el uso general de modelos lógicos en las disquisiciones sobre filosofía de la ciencia y sobre la filosofía en general. En este sentido, el rechazo a la lógica per se no se hizo esperar, pero esto es un error y una muestra de ignorancia sobre las posibilidades de la lógica.

Paralelamente, el impacto que añaden las lógicas no clásicas, en especial las paraconsistentes, incluye la posibilidad de obtener sistemas formales que representen alguna parte de esta distinta racionalidad filtrada en toda esta crítica. Esto es, a la par que se suman a la crítica de ciertos supuestos de racionalidad al estilo sólo de la Lógica Clásica, proporcionan elementos para la solución del problema.

Así, sin hacer caso omiso de que la crítica a la que me he referido atacó nuestra noción de racionalidad y de que esto es de una importancia mayúscula, tenemos la opción de enfrentar el reto de elucidar algunos aspectos de la racionalidad científica a través de resultados concretos en Lógica No Clásica. En especial, uno puede descubrir que podría haber interacciones más interesantes entre los problemas y las teorías propuestas en Filosofía de la Ciencia, y campos disciplinares más recientes, como la Inteligencia Artificial. Creo que éste es un punto de partida crucial para cualquier filosofía que pretenda tomar conciencia de los acontecimientos lógicos ocurridos en este siglo.

De este modo, el uso de herramientas de Lógica No Clásica en la modelación de la explicación tiene en mi trabajo un papel central. Mi tesis además de intentar ofrecer una propuesta alternativa a la elucidación de la explicación pretende, aunque sea mediante una especie de curriculum oculto, constituir también un argumento a favor de la posibilidad de la modelación lógica de la explicación, mediante algunas lógicas no clásicas.

La introducción de herramientas lógicas no clásicas se ve motivada en mi trabajo, a través de los puntos M1-M3 que anoto en seguida.

(M1) La reflexión sobre una breve reconstrucción histórica de algunos de los problemas generales que han enfrentado algunos filósofos en la búsqueda por modelar a la explicación.

(M2) Una observación que se relaciona con una inesperada consecuencia de la modelación de contextos en una explicación basada en relaciones inferenciales

deductivas. Dicha observación recae sobre el hecho de que en ocasiones el *explanans* tomado en conexión con un conocimiento de fondo en el que surge la explicación, conforma un conjunto de proposiciones inconsistentes. En consecuencia, de esperarse que este conjunto apoyara una inferencia deductiva, tal conjunto serviría de explicación para cualquier *explanandum*. Aunque esta observación pudiera interpretarse tan sólo como una crítica débil al deductivismo debido a su carácter aparentemente minoritario en medio de los casos consistentes, revela algunas cosas interesantes:

- a) la importancia de distinguir el *explanans* respecto de los elementos contextuales,
- b) una visión más realista acerca del tipo de cosa a la que nos enfrentamos al elucidar la explicación y, consecuentemente,
- c) la necesidad de un modelo más complejo de la explicación.

Además, el carácter minoritario que he mencionado debería sopesarse con la idea de que se trata en no pocos casos, de una suerte de "inconsistencia" de transición que tiene un interesante valor en la legitimación de las teorías.

(M3) describe el enigma de encontrar la forma de construir una noción de inferencia que representara mejor la relación de explicatividad. Esto último surgió como consecuencia de los resultados acerca de las formas en que la explicación se escapa de las nociones de inferencia mencionadas y, paralelamente, de la idea de que SA es un supuesto valioso para la modelación de la explicación.

Describiré ahora brevemente el desarrollo de la tesis tomando en cuenta las motivaciones M1-M3 anteriores.

M1 pone sobre la mesa de discusión el valor de SA en contraste con las dificultades encontradas para la modelación lógica de la explicación. Esto será desarrollado en el capítulo I y en el II. En el capítulo I realizo una revisión histórica de los problemas, luego examino conceptualmente los problemas generales para la modelación, y esto me sirve después de punto de partida para el análisis de un texto relativamente reciente de Hempel que constituye una crítica medular sobre la utilización de la noción deductiva de inferencia en la representación teórica. Intento entonces subrayar la importancia del cambio de perspectiva respecto de la modelación de la explicación y, finalmente, enfatizo los problemas de consistencia en el *explanans* y la idea de que la explicación pueda considerarse como un nuevo tipo de inferencia no deductiva, a la luz del supuesto de que los *explanantes* de nuestras explicaciones son falibles. Ya este análisis nos llevará a sugerir que una posición no deductivista será imprescindible para capturar algunos detalles importantes de la explicación. En el capítulo II abordaré con mayor detalle los contraejemplos paradigmáticos y trataré de hacer ver que será imprescindible el abordaje de los problemas relacionados con la relación de pertinencia explicativa. Este análisis nos conducirá posteriormente a la necesidad de un cambio de perspectiva en la modelación que, como dije, va más allá de la no deducción. Un cambio de perspectiva como el anunciado en mi tesis de maestría: necesitaremos una perspectiva epistémica y pragmática para la efectiva disolución de algunos de los problemas que se implican a través de los casos paradigmáticos presentados a lo largo de la discusión en Filosofía de la Ciencia. La propuesta que defenderé para la disolución de estos problemas irá por dos vías: una formal y otra filosófica. La primera vía es la construcción de un modelo que permita representar las funciones internas y externas de la explicación, y en el que subyazca una racionalidad más débil para la interpretación de las explicaciones. Describiré aquí, más detalladamente, la aproximación filosófica al problema. Primero haré algunas distinciones para clarificar el debate sobre la modelación argumental de la explicación. Distinguiré entre relaciones de explicación, explicaciones y relaciones de pertinencia explicativa. Esto me permitirá distinguir dos clases de *relata* asociados a la relación de pertinencia explicativa, lo cual nos encaminará a una solución pragmática, que considero

teóricamente importante, para las asimetrías en la explicación científica. El modo en que completo esta aproximación se encuentra motivado por el callejón sin salida al que nos arroja la discusión filosófica sobre la posibilidad de un modelo argumental de explicación: la necesidad de la inclusión del contexto en contraposición con la conveniencia de evitar que la noción de explicación, a través de una relación de pertinencia explicativa demasiado laxa, implique un relativismo. ¿Cómo entonces, al respecto de este tema, mantener un contextualismo sin abrazar un relativismo? La manera en que mi propuesta filosófica contesta a esta cuestión se constituye de un intento de recuperar la intuición de que la relación causal es importante para entender la pertinencia explicativa y, a la vez, de un intento de capturar la intuición de que el lado no causal de una asimetría también puede interpretarse como conteniendo una relación explicativa. Así, la caracterización que haré de la relación de pertinencia explicativa se constituye de dos casos: la relación causal y la relación inferencial, en una apuesta por la idea de que si reinterpretemos los *relata* involucrados en los casos de explicaciones no causales, podremos capturarlos, como casos de explicación de la verdad de una proposición. En detalle, primero supondremos que la explicación tiene forma de un argumento, y en esto nos encontramos siguiendo la intuición van Fraasseniana de que explicar es *dar buenas razones* para creer en el *explanandum*. En segundo lugar, consideramos que necesitamos responder a la pregunta de cómo distinguimos los argumentos que sí son explicaciones, acá nos encontraremos intentando recuperar la intuición de que la explicación debería “*dar cuenta del explanandum*”. Así, usamos la relación de pertinencia explicativa para tal distinción. Si la explicación está dirigida a dar cuenta de sucesos, entonces la relación de pertinencia explicativa que debe verificarse es la de causalidad. Y si se pretende dar cuenta de proposiciones, entonces la relación de pertinencia explicativa importante será la relación inferencial. Tercero, realizamos una distinción más. En el caso de la explicación causal, tenemos dos casos, el caso en que la relación causal es más parecida a una relación de causalidad eficiente, y el caso en el que se trata de una relación de fundamentación de un suceso en otro, en ambos casos podemos decir que se trata de una relación causal, al nivel de los hechos. Tanto los dos casos causales, como el caso inferencial, pueden verse como

formas diversas de recuperar de manera integral, la intuición de “dar razones para creer en el *explanandum*” tanto como la intuición de “dar cuenta del *explanandum*”; además, los dos casos (relación causal y relación inferencial) serán casos de relaciones de pertinencia explicativa que estarán, en alguna medida, contextualizadas. Para ello será importante, por supuesto, anticipar en la discusión la conveniencia de la distinción entre explicación y relación de pertinencia explicativa. Propondré entonces entender la explicación como una clase particular de argumento que, sólo cuando está dirigido intensionalmente a explicar sucesos, exige conexiones causales entre los sucesos que hacen verdaderas sus premisas (*explanans*) y el suceso que describe su conclusión (*explanandum*). Por supuesto, esto no es la solución completa, se requerirá adoptar un modo en que podamos limitar la manera en que el contexto determina de antemano la relación de pertinencia, sea causal o inferencial, evitando que uno pueda construir un contexto *ad-hoc* para cualquier relación. La manera de limitar esto, en el caso causal, será manteniendo una definición intensional constante de causalidad, mientras que se permite el cambio extensional de esta relación de contexto a contexto. Esto se defenderá en el capítulo II, pero se intentará representar formalmente en el capítulo IV.

El cambio de perspectiva nos lleva de modo natural a la inclusión del contexto en un modelo de explicación científica. Sin embargo, como he anticipado desde el inicio de esta introducción, esto representará un problema adicional. M2 puede ser descrito simplifícadamente mediante la siguiente afirmación: Algunas veces el *explanans* depende de una teoría inconsistente y a pesar de ello no se concluye que dicho *explanans* explica trivialmente cualquier cosa. ¿Cómo bloqueamos la inferencia deductiva clásica para estos casos? Este problema lo analizaré en el capítulo III. En este capítulo utilizo como ejemplo el caso de la teoría de los conjuntos de Cantor en relación a la explicación de la noción de número de Frege, después presento una breve revisión histórica de un tipo de lógicas no clásicas que pueden contribuir a solucionar esta cuestión: las lógicas paraconsistentes.

M3 puede ser descrito como sigue: Algunas veces no hay una relación inferencial deductiva entre el *explanans* y el *explanandum* a pesar de que puede representarse una relación inferencial. ¿Cómo construir este nuevo tipo de relación inferencial?. Esto será también un tema abordado en el capítulo III. Análogamente al capítulo II, comienzo con la exposición de un ejemplo de explicación; dicho ejemplo pone de manifiesto una relación inferencial que no obedece al Principio de Monotonidad de la lógica clásica; este principio puede explicarse llanamente como el principio de que si aumento información a las premisas de un argumento deductivo la conclusión se sigue infiriendo del conjunto de premisas aumentado. El ejemplo me sirve de pretexto para presentar una revisión histórica sobre un tipo de lógicas no clásicas que pueden ayudarnos a encontrar la relación inferencial idónea para la explicación: las lógicas no monotónicas. Después presento un sistema "clásico" de lógica no monotónica que utilizaré de manera fundamental para la construcción del modelo presentado en el capítulo IV, la lógica del razonamiento por *default* de Raymond Reiter, usada normalmente con fines de representación en Inteligencia Artificial. Finalmente examino algunos puntos en relación con las propiedades generales de este tipo de lógicas, algunas características importantes del sistema de Reiter y los vínculos pertinentes al problema de la modelación de la explicación.

En el capítulo IV desarrollaré el modelo de explicación científica que propongo. Realizo en este capítulo la construcción del modelo en forma de una estructura inmersa en un marco formal general; esto lo hago en un sentido inverso a la argumentación general de la tesis, primero presento la parte no monotónica del modelo y después añado los instrumentos paraconsistentes con sus respectivos cambios en la parte no monotónica del modelo. Por último muestro cómo el modelo salva algunos casos paradigmáticos y resuelve ventajosamente algunos problemas expuestos inicialmente en la tesis. También en esta parte menciono algunas limitaciones del modelo que pueden anticiparse.

CAPÍTULO 1

ALGUNAS DISCUSIONES ACERCA DE LOS MODELOS CLÁSICOS DE EXPLICACIÓN CIENTÍFICA

Introducción

Un análisis de las propuestas de modelos de explicación, hechas a partir del debate sobre los modelos clásicos de Hempel y Oppenheim, es indispensable para la identificación de los problemas que tendría que superar un modelo de explicación en el contexto de la Filosofía de la Ciencia. En este capítulo comenzare exponiendo los modelos clásicos y después dare una caracterización breve de algunos de los principales modelos que les sucedieron en el debate. Podremos identificar con este ejercicio dos tendencias diferentes en la propuesta de un modelo de explicación, una que apunta a una racionalidad, que bien podríamos calificar de dura, y otra que apunta a una idea más bien flexible de racionalidad.

Cuando Hempel y Oppenheim [**Hempel & Oppenheim 1948**]¹ propusieron su Modelo Nomológico-Deductivo prácticamente sistematizaron los vectores de la discusión que serían abordados en la discusión de las décadas posteriores. Después Hempel lo haría más complejo mediante su propuesta del Modelo Inductivo-Estadístico que, aparentemente, trajo consigo una visión contraria a la de la propuesta por él y Oppenheim. Llamaremos a estos dos modelos, los “modelos clásicos”. La discusión

¹ Hempel, Carl, G. & Oppenheim, Paul, "Studies in the Logic of Explanation", *Philosophy of Science*, vol. 15, pp. 567-579, 1948. Reimpreso en Joseph Pitt, (ed.) *Theories of Explanation*, Oxford University Press, E.U.A., 1988, p. 10. Haré las referencias en relación con la reimpresión de Pitt.

posterior a ellos puede verse como atravesada por dos ejes: una tensión en la propuesta del modelo Nomológico-Deductivo que se produce por una tendencia a debilitar el modelo; y una tensión provocada por el intento de ajustar el modelo Inductivo-Estadístico a ciertos cánones más rígidos de racionalidad.

En este capítulo ofreceré un panorama general de algunos de los problemas en la discusión de la explicación científica. Haré esto mediante una breve descripción de sólo algunos de los principales modelos de explicación científica y de las tesis y problemas que les subyacen, presentando la propuesta Hempel-Oppenheim, en contraste con la de algunos de sus críticos, específicamente de Scriven, Coffa, Railton, Salmon y van Fraassen.

1.1. La Propuesta Hempel-Oppenheim

Hempel propuso en realidad tres modelos de explicación científica: el nomológico-deductivo (D-N), el deductivo-estadístico (D-S) y el inductivo-estadístico (I-S). Los tres modelos se encuentran basados por lo menos en tres supuestos comunes. El primero es lo que propongo que sea llamado el "supuesto argumental" (SA). Esto es, el supuesto de que un argumento de cierto tipo puede modelar adecuadamente la explicación; dado SA, una explicación se compone básicamente de dos elementos: el *explanans* y el *explanandum*. Como he dicho en la introducción, el *explanandum* es el enunciado que describe al hecho que se pretende explicar, mientras que el *explanans* es el conjunto de enunciados que se aducen para explicarlo [Hempel & Oppenheim 1948, p. 10]. El segundo supuesto común de los modelos de Hempel es la necesidad de que el *explanans* contenga alguna clase de regularidad general, ya sea que ésta sea una ley deductiva o una regularidad estadística. El tercer supuesto es que el tipo de explicación que se intenta caracterizar con estos modelos es la que responde a preguntas del tipo "¿Por qué?".

Cada modelo puede representarse mediante un tipo de esquema de argumento distinto. La diferencia entre los tres esquemas está basada en algunas condiciones que se imponen a los componentes de los mismos. Las condiciones se refieren tanto al tipo de fórmulas a las que se permite formar parte del esquema, como a las relaciones entre dichas fórmulas. Los tres modelos definen tres distintos tipos de argumento en relación no sólo a su esquema sino a condiciones extra-lógicas y al tipo de derivación supuesto entre el *explanans* y el *explanandum*. En términos generales, puede decirse que el modelo D-N está basado en un tipo de derivación deductiva con un *explanans* que contiene enunciados de ley universales. El modelo D-S es un modelo que también tiene una noción de derivación deductiva y enunciados universales que contienen aseveraciones probabilísticas². Por último, el modelo I-S posee una noción de derivación probabilística con enunciados generales probabilísticos. Sólo presentaré los modelos D-N e I-S en este trabajo, pues son los que se encuentran en relación más directa con las críticas hechas a la propuesta Hempel-Oppenheim.

1.1.1. El Modelo D-N

El modelo D-N, de [Hempel & Oppenheim 1948], se constituye de un esquema bajo ciertas condiciones de adecuación.

Esquema:

$$C_1, C_2, \dots, C_k$$

$$\underline{L_1, L_2, \dots, L_r}$$

$$E$$

Donde C_1, C_2, \dots, C_k representa un conjunto de enunciados singulares que describen ciertas condiciones iniciales; L_1, L_2, \dots, L_r , un conjunto de leyes generales y E el

²La deducción en este caso está permitida en virtud de las leyes de la teoría matemática de la probabilidad estadística.

explanandum. Las premisas del esquema constituyen el *explanans*. La línea del esquema es una relación lógica deductiva clásica.

Este esquema define una relación de inferencia entre el *explanans* y el *explanandum*. Además, el esquema está sujeto a algunas condiciones que Hempel y Oppenheim clasifican bajo el rubro de “condiciones lógicas de adecuación”. Tales condiciones son las siguientes:

- 1) El *explanans* debe contener por lo menos una ley.
- 2) El *explanans* debe contener por lo menos una condición inicial.
- 3) El *explanandum* debe derivarse deductivamente del *explanans*.
- 4) Toda ley en el *explanans*, debe requerirse para la derivación del *explanandum*
- 5) El *explanans* debe tener contenido empírico. Es decir, debe ser susceptible de prueba mediante observación o experimentación por lo menos en principio [Hempel & Oppenheim 1948, p. 11].³

Hempel agrega una condición más, que llama "Condición Empírica de Adecuación.":

- 6) Los enunciados del *explanans* deben ser verdaderos.

Dado el papel central de los enunciados de ley en el modelo D-N de explicación, Hempel y Oppenheim realizaron un análisis muy detallado de la forma lógica de un enunciado de ley. No expondré este aspecto de su modelo por dos razones: Primero, construiré un modelo que, como algunos otros modelos, no supondrá la necesidad de incluir leyes en el *explanans* y por ello no será de interés crucial en la discusión desarrollada en esta investigación. Segundo, aún en el caso de que Hempel y Oppenheim lograran una definición adecuada para la forma de los enunciados de ley, la mayoría de las

³En [Hempel & Oppenheim 1948] se presentan estas condiciones de forma más compacta. En el texto de ellos las condiciones 3-5 que he descrito aparecen juntas como la condición R2.

críticas a sus modelos no dependen directamente de su forma de entender dichos enunciados⁴. Sin embargo es necesario extraer y mencionar, del análisis que ellos realizan para su caracterización de enunciado de ley, algunas distinciones que serán importantes para entender su presentación formal de las condiciones de una explicación.

En relación a la noción de ley, de un análisis del artículo de 1948 se desprende lo siguiente. Formalmente hablando, pueden distinguirse dos clases de enunciados generales: los universales y los existenciales⁵. En la clase de los enunciados universales se encuentra la clase de los enunciados de ley [Hempel & Oppenheim 1948, p. 27]. Hay dos subclases de enunciados de ley: la subclase F de los enunciados puramente universales y la subclase D de los enunciados esencialmente universales que pueden ser derivados de los puramente universales⁶. Un enunciado puramente universal es un enunciado universal que no contiene constante individual. Por ejemplo, el enunciado $(x)(Px \vee \neg Px)$ es puramente universal y, nótese, es equivalente al enunciado singular $Pa \vee \neg Pa$. Un enunciado esencialmente universal es un enunciado universal que no es equivalente a un enunciado singular⁷. Por ejemplo, los enunciados $(x)(Qxa \supset Vx)$ y $(x)(Px \supset Vx)$ son enunciados esencialmente universales⁸. A las leyes que pertenecen a F Hempel y Oppenheim les llaman leyes fundamentales y a las que son miembros de la clase D, leyes derivadas. Así, tenemos dos clases principales: la clase de enunciados generales G y la clase de enunciados de ley L tales que si $U \subseteq G$ es la clase de los universales, $L \subseteq U$. La clase L se divide, a su vez, en dos subclases F y D.

⁴Excepto la crítica de [Orayen 1976].

⁵Los universales son los enunciados que comienzan con un cuantificador universal y los existenciales con un cuantificador existencial.

⁶La letras "F" y "D" corresponden, respectivamente, a "Fundamental" y "Derivada".

⁷Un enunciado singular es un enunciado que contiene sólo constantes individuales.

⁸Hempel está suponiendo que el dominio sobre el que corren los cuantificadores es finito y que no tenemos fórmulas infinitas.

Hempel y Oppenheim presentaron una caracterización formal de la explicación. Esta caracterización formal tiene algunos detalles que conviene resaltar⁹. Consideremos L el lenguaje de la lógica clásica de primer orden con funciones [Hempel & Oppenheim 1948, p. 26]¹⁰. Supongamos que T es un enunciado general que no es equivalente a un enunciado singular. Supongamos también que C es un enunciado singular y E es el enunciado singular que describe al *explanandum*. Supongamos que $\{T, C\}$ cumple con la condición 5 antedicha. Entonces tenemos el siguiente par de definiciones:

(ExP)

El par ordenado $\langle T, C \rangle$ es un ***explanans potencial*** para un enunciado singular E si y sólo si:

- (1) $\{T, C\} \mid\!\!\!-\ E$
- (2) No ocurre que $C \mid\!\!\!-\ E$.

(Ex)

El par ordenado $\langle T, C \rangle$ es un ***explanans*** para el enunciado E si y sólo si¹¹:

- (1) $\langle T, C \rangle$ es un *explanans potencial* para E .
- (2) T y C son verdaderos.

Observemos algunos detalles. La separación en dos definiciones implica una división en el cumplimiento de las condiciones 1-6. Esta es una forma importante de exponer, separadamente, las condiciones lógicas de adecuación de las condiciones empíricas de adecuación.

Dadas las características de T y C , se cumplen las condiciones 2 y 5. Por ExP(1) se cumple la condición 3 y por ExP(1) y ExP(2) se cumple la condición 4.-Finalmente, por Ex(2)

⁹La siguiente caracterización formal difiere de la de Hempel y Oppenheim en algunos detalles. Intento con esta formulación evitar los detalles que refieren a la definición formal de T .

¹⁰ L es un lenguaje cuyos predicados primitivos son puramente cualitativos. Esto significa que para su definición no es necesario referir a entidades particulares. Esto no es esencial para la exposición del modelo D-N.

¹¹[Hempel & Oppenheim 1948, p. 28.]

se cumple la condición 6. No obstante, la condición 1 no se cumple. T , tanto en ExP como en Ex no es, necesariamente, un enunciado de ley. Mientras que Hempel y Oppenheim aducen al principio del artículo que una condición necesaria para una explicación es que el *explanans* contenga leyes, la definición formal de un *explanans* no incluye esta condición. Esta interesante distinción muestra que o bien Hempel y Oppenheim contemplan casos de explicación no basados en leyes, o por lo menos es claro que una caracterización formal de la explicación como la de los modelos clásicos, pero que prescindiera de las leyes, es posible. Por supuesto, ya las estipulaciones hechas para la caracterización de T como las características del lenguaje L , implican que T sea un enunciado general de cierto tipo particular, específicamente, del tipo de enunciados que no son equivalentes a enunciados singulares. Pero esto deja abierta la puerta a *explanantes* sin enunciados de ley.

Hempel y Oppenheim expresan su preocupación por el hecho de que un *explanans* tenga también enunciados generales que no son leyes y, por ello, definen anticipadamente los enunciados de tipo T : enunciados generalizados (no necesariamente universales) no equivalentes a un enunciado singular. A estos enunciados los llaman "teorías". Una teoría definida así puede ser un enunciado de ley, pero no es necesario que lo sea. Así, Hempel y Oppenheim estarían contemplando los casos de explicaciones que no contienen leyes aunque no admiten explicaciones sin enunciados generales de un tipo preciso: los que no son equivalentes a enunciados singulares¹². De cualquier forma, su modelo se dedica a un caso especial de estas explicaciones, a aquellas que contienen leyes.

Sin duda hay formas de integrar la condición 1 a la caracterización formal que he presentado. Por ejemplo mediante la consideración de *explanantes* compuestos por dos conjuntos. Uno, digamos, que tenga como miembros a enunciados del tipo de T , y otro, χ , que contenga a enunciados singulares C . De esta forma, podríamos estipular que debe

¹²A estos enunciados les llaman enunciados "esencialmente generales", que no necesariamente tienen que ser universales como las leyes. Una cita al respecto es la que sigue: "The explanation of a phenomenon may involve generalized sentences which are not of universal form.", [Hempel & Oppenheim 1948, p. 28].

contener al menos un enunciado T' de ley y con ello se estaría cumpliendo la condición 1. Me parece que esto es lo que Hempel y Oppenheim debieron de haber tenido en mente al construir su caracterización formal de la explicación.

Finalmente, el modelo D-N supone dos tesis más. La primera es que con la **subsunción por leyes** (o subsunción legal) el modelo captura las conexiones causales que normalmente se suponen entre los hechos mencionados en el *explanans* y el hecho descrito por el *explanandum*. En general, se refiere al tipo de explicación que ha sido considerado hasta aquí, como una explicación causal.¹³ Por supuesto, la relación causal es asumida en el sentido empirista tradicional: como una relación de la cual sólo podemos verificar propiedades de sucesión, conjunción constante y contigüidad.

La segunda tesis es a veces llamada el "quinto requerimiento de Hempel", por ejemplo en [Scriven 1962]¹⁴, y es la idea de que la explicación y la predicción tienen la misma estructura interna. Esto es en el sentido de que ambos tipos de actividad científica, la explicación y la predicción, tienen las mismas características lógicas [Hempel & Oppenheim 1948, p. 12]. Lo único que las distingue es la dimensión pragmática de su uso: la primera da apoyo al *explanandum* y la segunda nos produce una expectativa racional de la ocurrencia de ciertos hechos en el futuro [*idem*]¹⁵. Esto después se conoció como la "tesis de la simetría entre la explicación y la predicción" y es una de las ventajas del modelo argumental de Hempel, pues permite que la predicción y la explicación puedan tener formalmente un mismo modelo.

¹³"The type of explanation which has been considered here so far is often referred to as a causal explanation." [Hempel & Oppenheim 1948, p. 13].

¹⁴Scriven, Michael, *Explanations, Predictions, and Laws* University of Minnesota press, Minneapolis, 1962. Reimpreso en Joseph Pitt, *Theories of Explanation*, Oxford University Press, E.U.A., 1988. La paginación, en adelante, se refiere a la reimpresión de [Pitt 1988]. De acuerdo a la organización de los cuatro requerimientos R1-R4 en [Hempel & Oppenheim 1948], R1 es la condición 3; R2 corresponde a las condiciones 1, 2 y 4; R3 es la condición 5; y R4 es la condición 6.

¹⁵*Idem*.

1.1.2. El Modelo I-S y las Diferencias entre ambos Modelos

En el artículo de 1948, Hempel y Oppenheim mencionan muy brevemente otras clases de explicación que requerirán otros modelos. Posteriormente, en [Hempel 1965]¹⁶, Hempel presenta uno de esos modelos: el modelo I-S de explicación [Hempel 1993, p. 44]. Si “ $P(G/F) = r$ ” designa que la probabilidad de que algo tenga la propiedad G, dado que tenga la propiedad F, es igual a r; y “Fb” designa que el objeto b tiene la propiedad F; entonces el esquema que define parcialmente el modelo I-S es el siguiente [Ibid., p. 44]:

(I-S)

$$\begin{array}{l} P(G/F) = r \\ Fb \\ \text{=====}[r] \\ Gb \end{array}$$

Donde la primera premisa expresa una probabilidad condicional general y la segunda es una instancia en que se cumplen las condiciones. La conclusión es una instancia de la consecuencia del enunciado de probabilidad. La doble línea entre las premisas y la conclusión representa una noción de derivación probabilística: la conclusión es derivada con un grado alto de probabilidad r . Para Hempel es imprescindible que r sea un grado alto de probabilidad. Al parecer, la motivación de Hempel para estructurar el modelo I-S de esta manera es que las conclusiones que vienen de enunciados probabilísticos corresponden a una certeza deductiva. Veamos esto.

¹⁶ Hempel, Carl, G., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press y Collier Macmillan, Nueva York y Londres, 1965. Parcialmente reimpresso en [Hempel 1993]: *Aspects of Scientific Explanation*, en David-Hillel Ruben, (ed.), *Explanation*, Oxford University Press, E.U.A. 1993. Una traducción completa al castellano es [Hempel 1979]: *La Explicación Científica*, Edit. Paidós-Argentina, 1ª edición, Buenos Aires, 1979. La paginación a la que referiré en adelante, es la de [Hempel 1993].

Una forma alternativa para concluir Gb es usar una noción de derivación deductiva y concluir un enunciado modalizado. Como sigue:

(DP)

$$P(G/F) = r$$

Fb

$$P(Gb) = r$$

Sin embargo, Hempel piensa que una derivación de este tipo cometería el error de concluir con seguridad un enunciado modal cuya verdad es relativa a un cuerpo de evidencia particular. En realidad, se dice en [Hempel 1993, p. 43], la conclusión no es "autocontenida" pues no depende únicamente de las premisas del esquema. Por esta razón, Hempel prefiere relativizar la noción de derivación a la probabilidad del enunciado que funciona como regularidad.

El cambio de noción de derivabilidad, le trajo a Hempel una nueva forma de entender la explicación científica que alejará mucho a este modelo de las virtudes que se le podían suponer al modelo D-N. Hempel parece darse plena cuenta de esto cuando dijo que los argumentos I-S no otorgan un grado de certeza completo a la conclusión pero que:

"En lugar de negarles *status* explicativo sobre la base de que la no realización del *explanandum* es compatible con el *explanans*, es decir a pesar de su invalidez, debemos reconocer que constituyen explicaciones de un carácter lógico diferente que reflejan, podríamos decir, un sentido diferente de la palabra "porque"
[Hempel 1979, p. 387].

La idea de que los argumentos I-S deben considerarse como explicaciones puede estar apoyada además por el supuesto de que en esta clase de argumentos subyace la

representación de relaciones causales [Hempel & Oppenheim 1948, p. 32]¹⁷ entre los hechos descritos por el *explanans* y el descrito por el *explanandum*. Sin embargo, una cosa es darle a los argumentos I-S la categoría de explicaciones y otra muy distinta es suponer que sus características son muy semejantes a los D-N.

Hempel parece no darse cuenta completamente, en su artículo de 1965, cuán distintas son las concepciones supuestas por los modelos D-N e I-S. Las diferencias entre los modelos pueden verse a través de los grupos de problemas, muy distintos entre sí, para cada uno. Estos grupos de problemas pueden explicarse respectivamente, como he dicho antes, mediante la tensión Hempel-Scriven y la tensión Hempel-Coffa que abordaré más adelante.

Gran parte de los problemas principales del modelo I-S fueron detectados por Hempel desde el artículo recién mencionado. La manera más fácil de exponer las diferencias más inmediatas entre el modelo D-N y el modelo I-S, es acudiendo primero a un ejemplo problemático que presenta el mismo Hempel. Consideremos el siguiente esquema [Hempel 1993, p. 44]:

(Rec)

Px: a x se le aplica penicilina.

Sx: x tiene una fuerte gripa.

Rx: x se recupera de la gripa.

j : Juan

$$P(Rx / Sx \wedge Px) = .8$$

$$Sj \wedge Pj$$

$$===== [.8]$$

$$Rj$$

¹⁷ Aunque esta representación sea considerada sólo como una conjetura causal, como una hipótesis de trabajo, véase por ejemplo [Hempel 1979, p. 346].

No obstante, Si Juan es un señor de 92 años, las probabilidades de que Juan no se recupere pueden ser altas y podemos tener el siguiente esquema:

($\neg Rec$)

Nx : x tiene 92 años.

$$P(\neg Rx / Sx \wedge Px \wedge Nx) = .8$$

$$Sj \wedge Pj \wedge Nj$$

$$===== [.8]$$

$$\neg Rj$$

De manera que si nuestro conocimiento tiene la información descrita por $Sj \wedge Pj \wedge Nj$, entonces los esquemas Rec y $\neg Rec$ nos conducirían a inconsistencias: tendríamos Rj y $\neg Rj$ en el mismo sistema. Hempel llama a estos resultados "inconsistencias inductivas". Hempel se ve entonces en el problema de manejar dos argumentos con conclusiones contradictorias, pero además debe notarse que ambos argumentos pueden ser contruidos a partir de la misma base de datos sobre el mundo. No necesitamos rechazar las premisas del primer argumento, Rec , para poder construir el segundo, $\neg Rec$, sino que el cuerpo de premisas del primero forma parte del segundo. Así, un mismo cuerpo de evidencia nos conduce a dos argumentos en conflicto. En otra sección de su libro de 1965 Hempel escribe que las reglas elementales de inducción violan el requisito mínimo "que toda regla de procedimiento científico debe llenar" [Hempel 1979, p. 79]: pueden conducir a inconsistencias.

La pregunta de cómo evitar considerar como legítimos ambos grupos de premisas contenidas en dos esquemas I-S que conducen a inconsistencias, puede verse como un problema en relación a un conjunto de datos determinado K . Este es el modo como Hempel lo aborda. Pero este conjunto K , considerado como un estado epistémico, puede no poseer todos los datos acerca del sujeto en cuestión. Por ejemplo, podría no tener datos acerca de, digamos, la edad de Juan. Esto es lo que constituye lo que se ha llamado

el "**problema de la ambigüedad epistémica**" [Hempel 1993, p. 46]. Una perspectiva para abordar este problema, la que sigue Hempel, es la búsqueda de la clase de referencia adecuada: no bastaría con saber que Juan tiene una fuerte gripa y que se le ha aplicado penicilina sino que también deberíamos haber tomado en cuenta su edad. Encontrar la clase de referencia adecuada es equivalente a saber lo *suficiente* acerca de Juan. Pero esto casi nunca es factible. Así, el problema de la ambigüedad epistémica nos conduce a la relativización de la explicación respecto de un conjunto de conocimientos. Este conjunto de conocimientos está determinado por un tiempo t . Escribamos K_t para designarlo. Nuestra clase de referencia estará relativizada a K_t . Hempel atribuye [Ibid., p. 47] las siguientes propiedades a K_t :

- a) Los miembros del conjunto K_t no pueden ser considerados como verdaderos, sino sólo aceptados en forma tentativa.
- b) La membresía del conjunto K_t cambia en el transcurso del tiempo. Por nueva "evidencia desfavorable".
- c) El conjunto K_t es consistente.
- d) El conjunto K_t está cerrado bajo implicación lógica.

La propiedad descrita por el inciso c proporciona una forma de solucionar el problema de la ambigüedad epistémica. Si se supone que se ha reconocido científicamente la ocurrencia de una de las dos conclusiones de los argumentos en conflicto, ese *explanandum* estaría en K_t . Dado que K_t es consistente, el *explanandum* contradictorio no aparecería en K_t . Por tanto, no habría ambigüedad epistémica. Esto podría ayudarnos a decidir entre ambos argumentos. Sin embargo uno desearía, como también nota Hempel, decidir entre ellos sin importar si el *explanandum* ocurrió de hecho o no [Hempel 1993, p. 47-8]. Para ofrecer una solución más óptima, Hempel postula el **Requerimiento de Máxima Especificidad**:

(RME).

"Considérese una explicación propuesta de la forma estadística básica

(3o)

$$\begin{array}{l} "p(G / F) = r \\ Fb \\ ===== [r] \\ Gb" \end{array}$$

Sea s la conjunción de las premisas y, si K es el conjunto de todos los enunciados aceptados en un tiempo dado, sea k un enunciado que es lógicamente equivalente a K (en el sentido de que k es implicado por K y a su vez implica todo enunciado en K)¹⁸. Entonces la explicación propuesta (3o), para ser racionalmente aceptable en la situación de conocimiento representada por K , debe cumplir la siguiente condición (el requerimiento de máxima especificidad): si $s \wedge k$ implica que b pertenece a una clase F_1 , y que F_1 es una subclase de F , entonces $s \wedge k$ también debe implicar un enunciado que especifique la probabilidad de G en F_1 , digamos

$$p(G / F_1) = r_1$$

Y en esta fórmula r_1 tiene que ser igual a r "[Hempel 1993, p. 51]¹⁹

¹⁸Por supuesto, aquí el enunciado k no puede ser considerado exactamente la conjunción de todo enunciado en K , pues k podría ser infinito.

¹⁹"Consider a proposed explanation of the basic statistical form

(3o)

$$\begin{array}{l} "p(G, F) = r \\ Fb \\ ===== [r] \\ Gb" \end{array}$$

Lets s be the conjunction of the premisses, and, if K is the set of all statements accepted at the given time, let k be a sentence that is logically equivalent to K (in the sense that k is implied by K and in turn implies every sentence in K). Then, to be rationally acceptable in the knowledge situation represented by K , the proposed explanation (3o) must meet the following condition (the requirement of maximal specificity): If $s \wedge k$ implies that b belongs to a class F_1 , and that F_1 is a subclass of F , then $s \wedge k$ must also imply a statement specifying the statistical probability of G in F_1 , say

$$p(G, F_1) = r_1$$

Si hemos de considerar un esquema de argumento como un esquema I-S, RME nos obliga a que toda subclase de la clase F (tomada en cuenta en el argumento) debe tener la misma probabilidad que la clase F. Puede decirse entonces que la clase de referencia²⁰ del sujeto en cuestión es máxima desde un punto de vista intensional en el sentido de que incluye todas las características pertinentes del sujeto en cuestión. En este caso una característica pertinente es aquella que es capaz de modificar la probabilidad del *explanandum*. Supongamos que, en los ejemplos *Rec* y \neg *Rec*, todas las características relevantes al caso son *Sx*, *Px* y *Nx*. Consecuentemente, el argumento que ha tomado en cuenta la clase referencia adecuada, esto es, el argumento que cumple con RME, es \neg *Rec*. Por lo tanto, este argumento debe ser elegido preferentemente a *Rec*. Por otro lado, podemos suponer que Juan tiene una cicatriz de 5 pulgadas en el brazo izquierdo, esto es, Juan pertenece a la clase F_1 de personas que cumplen con *Sx*, *Px*, *Nx* y tienen semejante cicatriz. Si \neg *Rec* cumple con RME entonces la probabilidad de $\neg R_j$ respecto de F_1 debe ser la misma que la que resulta de tomar en cuenta sólo a *Sx*, *Px* y *Nx*. La cicatriz es, pues, una característica que no añade más probabilidad al *explanandum*.

La solución de Hempel al problema de la ambigüedad epistémica supone la idea de la **relativización epistémica** de las explicaciones I-S. Esta relativización implica un cambio de perspectiva radical respecto de la explicación D-N. Entre las consecuencias de esta perspectiva está el hecho de que las premisas de un argumento I-S claramente pueden no ser completas. Idealmente uno siempre tendría identificadas todas las características pertinentes para un dado *explanandum*. Pero esto no siempre ocurre así. En general, uno no tiene bien claras cuáles son todas las características pertinentes y, sobre todo, la idea de cuáles son pertinentes puede cambiar si nuestra información aumenta. Consecuentemente, el problema de la ambigüedad epistémica refleja otro problema complejo: el hecho de que **nueva información** puede cambiar nuestras conclusiones. Para

Here, r_1 must equal r "

²⁰"Problema de la clase de referencia" es el nombre con que se conoció el problema en teoría de probabilidades.

explicar esto, pensemos que ante la posibilidad de obtener datos para K_t mediante argumentos inductivos, y el hecho de que este tipo de argumentos puede perder fuerza o alguna clase de legitimidad ante nueva evidencia (como el caso de *Rec*), es claro que la información considerada en K_t puede cambiar debido a la aceptación de nueva evidencia, por lo menos mediante el mecanismo de considerar una sola de las conclusiones de los argumentos en conflicto.

Así, la tentativa de una modelación deductiva tuvo que ser complementada, por Hempel, mediante un modelo no deductivo de inferencia para ciertos casos de la explicación científica. No obstante, el nuevo modelo trajo consigo su conjunto de problemas. Los problemas del modelo I-S se originan de un rasgo conocido de las inferencias probabilísticas: la ambigüedad epistémica. Es prudente poner sobre relieve el tipo de solución que Hempel intenta, se basa en RME. Preguntarnos si podríamos construir una solución diferente es una cuestión interesante. Notemos también que el problema de la ambigüedad epistémica es el problema de que un esquema de inferencia arroje conclusiones en conflicto, contradictorias entre sí. La solución de Hempel incluye, como hemos visto, una relativización epistémica, y el intento de erradicar contradicciones. La tentativa de modelación hempeliana ha tocado fondo, al parecer, en un problema que lo arroja de pronto en una dimensión contextual de la representación de explicaciones.

Además del problema señalado por Hempel, el de la ambigüedad epistémica, hubo otros problemas para los modelos clásicos. Expondré enseguida brevemente algunos de éstos. Los expondré a propósito de dos artículos críticos. Michael Scriven es autor de uno de ellos; Alberto Coffa, del otro.

1.2. La Tensión Hempel-Scriven

Michael Scriven [Scriven 1962]²¹ construyó una fuerte crítica al modelo D-N de explicación. Su idea de explicación se funda en una noción de entendimiento y un supuesto, que parece subyacer a todo el texto de 1962: la idea de que la elucidación de la explicación debe estar basada en el uso lingüístico que se hace del término "explicación" [Ibid., p. 52]²². Hempel y Oppenheim, en cambio, no pretenden que la explicación que ellos caracterizaron con su modelo D-N represente la forma en que los científicos realmente hacen sus explicaciones [Hempel 1979, p. 405]. Aunque tampoco puede esperarse que ellos pretendan que sus modelos deban capturar explicaciones de ámbitos no científicos.

El artículo de Scriven es un importante ejemplo de una perspectiva distinta de los modelos clásicos para abordar el problema de la elucidación de la explicación. Una perspectiva que involucra, como veremos, la dimensión pragmática de esta actividad científica, es decir, involucra como elementos importantes de análisis dos nuevos elementos: el contexto y los sujetos que exigen o reciben una explicación. En lo que sigue haré primero una síntesis de las críticas de Scriven y después expondré brevemente las principales tesis sobre la explicación, que se proponen en su artículo.

Veamos las críticas. Scriven dice que el modelo D-N:

- (1) No distingue entre la estructura de la explicación y la de la predicción.
- (2) Enfatiza la descripción del hecho por encima del hecho mismo.
- (3) Excluye algunos tipos de explicación.

²¹ Scriven, Michael, *Explanations, Predictions, and Laws*, University of Minnesota press, Minneapolis, 1962.

²² Aunque aclara que esto no significa que uno pueda identificar explicaciones tan sólo mediante la localización de ciertos términos en el uso del lenguaje.

(4) Incluye casos que no deberían ser considerados como explicaciones.

Otra forma de describir el punto 1 es diciendo que Scriven argumentó en contra de la idea de establecer una simetría entre la estructura de una explicación y la de una predicción. No me detendré mucho en la descripción de esta crítica pues la aceptación o rechazo de la tesis de simetría no me parece crucial contra el modelo D-N ni contra ningún otro modelo de explicación científica. Valga el argumento de que la principal motivación para la construcción de un modelo de explicación es su elucidación, y esto podría resultar en el rechazo de la tesis de simetría entre explicación y predicción. Por lo menos intuitivamente, la idea de que tengan la misma estructura no parece que sea un factor indispensable de la explicación.

Básicamente, Scriven muestra que A) no toda predicción es correlativa con una explicación y que B) no toda explicación es correlativa con una predicción. Así, la tesis de la simetría sería completamente errónea.

Uno de los contraejemplos para mostrar A es que puede apelarse al hecho de que las vacas yazcan en el campo abierto durante el día, para predecir que lloverá en unas pocas horas. Sin embargo, nadie consideraría que este mismo hecho pueda explicar que lloverá en esas condiciones. Así, quedaría claro según Scriven que hay algunas predicciones que no pueden correlacionarse con explicaciones de lo predicho.

La parte B de su crítica toma como caso la explicación de leyes. La explicación de una ley general no tiene predicción que le pueda ser análoga [Scriven 1962, p. 54]. El razonamiento es como sigue. Una ley debe cumplirse en el pasado, en el presente y en el futuro. Lo que podría significar una predicción de una ley sería algo como la confirmación de todas las retrodicciones y predicciones posibles de la ley. En todo caso, si fuera posible entender esto como una predicción, tendría que extenderse el concepto de predicción de forma que incluyera al de retrodicción. Pero esto, según Scriven, no es plausible. La razón

es que el único rasgo característico de una predicción es su referencia a hechos futuros de modo que si el concepto se extiende también para hechos pasados entonces el concepto de predicción se pierde.

Una estrategia más para atacar la tesis de la simetría es mostrar que podemos explicar hechos que no han ocurrido mientras que no podemos tener predicciones de hechos que no ocurrirán [Scriven 1962, p. 61]. Pero esta crítica a la propuesta Hempel-Oppenheim no es justa dado que la clase de explicación que ellos consideran en el modelo D-N es, precisamente, una que asume que el *explanandum* es verdadero. La explicación en este caso es pensada como surgiendo de la constatación del *explanandum*. De modo que esta estrategia no es efectiva contra el modelo D-N, pues en este modelo tanto las explicaciones como las predicciones lo son de hechos que ocurren. A pesar de ello, Scriven revela aquí una intuición importante: la idea de que uno podría considerar la actividad de explicar, de manera independiente de la verdad de los componentes de la explicación.

La crítica 2, como vimos, se propone negar la prioridad de las descripciones de los hechos sobre los hechos. Hempel y Oppenheim pretenden que una explicación puede describirse como una relación entre las descripciones de los hechos involucrados (en el *explanans* y en el *explanandum*), cuando lo importante, según Scriven, son los hechos que uno pretende explicar. Bajo el supuesto de Scriven de que el objeto de una explicación es un hecho y no su descripción, él atribuye a Hempel y Oppenheim una inconsistencia. Según Scriven, la inconsistencia se produce porque ellos no se dan cuenta de las consecuencias de la distinción que hacen. Cuando uno explica, explica hechos. A sabiendas de esta distinción, ellos usan descripciones de los hechos y estipulan la deducibilidad como requisito. Este requisito es planteado por el modelo D-N como la única relación entre los elementos de una explicación. Sin embargo esto los lleva a una contradicción pues, dado que los hechos son parte de la explicación, la relación que puede establecerse entre la explicación y los hechos es con toda seguridad una relación no

deductiva [Scriven 1962, p. 61]²³. De modo que el modelo D-N tendría que suponer deducibilidad y no deducibilidad al mismo tiempo.

Para que la inconsistencia señalada por Scriven tuviera lugar en el nivel del discurso (proposicional u oracional) de Hempel y Oppenheim, Scriven tendría que asumir que ellos suponen que el hecho a explicar está incluido (de algún modo distinto al de una descripción) dentro del modelo D-N. Si Hempel y Oppenheim negaran tal inclusión la contradicción, al menos en el nivel del discurso, no tendría lugar. No obstante, lo que puede desprenderse de todo esto es que Scriven encuentra señalando en este punto que el modelo D-N es una sobresimplificación de la explicación, basado en la idea de que reduce la relación entre el *explanans* y el *explanandum* a una mera deducción, es decir, basado en la idea de que la reduce a una relación entre descripciones.

Hay otra línea de argumentación en [Scriven 1962] que puede tomarse como en contra del énfasis descrito por 2, el énfasis de las descripciones por encima de los hechos correspondientes. Esta línea consiste en mostrar que la no verificación de una conexión real entre los fenómenos referidos por el *explanans* y el descrito por el *explanandum*, conduce a casos de inadecuación del modelo D-N. La idea es suponer que esto ocurre justo por no tomar suficientemente en consideración los hechos y más bien dedicarse a una reconstrucción basada en descripciones. Esta línea de argumentación subyace en el argumento relacionado con el ejemplo del puente y la dinamita, el cual veremos en la descripción de la crítica 4.

El ataque contra la adecuación del requerimiento de deducibilidad lo constituyen los puntos 3 y 4, respectivamente contra la suficiencia y contra la necesidad de tal requerimiento. Scriven dirige esta crítica a través de contraejemplos. Algunos de los ejemplos de Scriven ofrecen un apoyo fuerte a la crítica 3 y algunos claramente apoyan la 4. Los ejemplos que apoyan 3 se encuentran ligados con la tesis de que hay explicaciones

²³"For it is a consequence of this distinction that a nondeductive step is involved between the statements in an explanation and the phenomenon explained."

que no reflejan relaciones causales entre los hechos que involucran. Un par de ejemplos a favor de 3 está en [Scriven 1962, p. 54]: explicar el simbolismo de una danza tribal y explicar una teoría o una prueba. Scriven dice que estos casos son tomados como casos de explicación de significado, por Hempel y Oppenheim. Scriven argumenta entonces que, a pesar de sus diferencias importantes con casos que establecen conexiones causales, son casos que guardan estrechas semejanzas con los casos causales.

Aparentemente, Hempel y Oppenheim podrían escapar a la crítica del párrafo anterior. En última instancia los casos mencionados pueden entenderse como respuestas a preguntas-qué o preguntas-cómo, y no a preguntas-porqué, preguntas a las cuales está dirigido el estudio de Hempel y Oppenheim. Sin embargo, creo que la crítica debe entenderse como un ataque contra los supuestos de Hempel y Oppenheim acerca de la explicación. Atacan el supuesto de que la explicación es sólo una respuesta a preguntas-porqué, de modo que, si se le toma de este modo, esta salida sería una petición de principio. Es importante percatarse de que los casos mencionados por Scriven obedecen al supuesto de que la elucidación de la explicación debería contemplar casos que son considerados como explicaciones en otros ámbitos. En este sentido, su propuesta puede ser tomada como una muestra de cómo el modelo D-N no captura algunos tipos de cosas que normalmente consideramos como explicaciones. Aunque ello no constituye una obligación para modelarlas, sí enriquece el área de investigación sobre la cual emprender un trabajo de modelación. No obstante lo anterior, para que la crítica de Scriven pudiera surtir efecto en contra de la necesidad del criterio de deducibilidad, lo cual era el punto, él tendría todavía que explicar cómo el hecho de no tomar en consideración relaciones no causales afecta la posibilidad de que la relación de deducibilidad pudiera modelar estas relaciones. La crítica de Scriven aparece así, por lo menos, incompleta.

Debe considerarse que el apoyo a la crítica 4 está dado en parte mediante las propuestas teóricas acerca de la naturaleza de la explicación, que Scriven desarrolla a lo largo de todo el artículo, pero principalmente mediante la idea de que la subsunción por

leyes del modelo D-N no captura la conexión causal que debería suponerse en una explicación D-N: No basta con la apelación a una ley general, sino que hay que verificar que existe efectivamente la conexión causal entre los eventos involucrados [Scriven 1962, p. 68]²⁴. Una reconstrucción del ejemplo principal de Scriven [*Idem*] para mostrar lo que no debería incluir el modelo D-N es la siguiente: supongamos que a un puente G le es lanzada cierta cantidad de dinamita D a una de sus arcadas. Pero supongamos también que, unos segundos antes de la explosión de la dinamita, ocurre una fractura en esa misma arcada, suficiente para derribar la parte del puente que depende de la misma. En este caso, es posible elaborar una explicación D-N para el hecho de que el puente ha sido derribado en esa parte, tal que use la ley de que a los puentes (del tipo al que pertenece G) que son dinamitados, en una de sus arcadas, con una cierta cantidad de dinamita D, les ocurre que la parte del puente que depende de esa arcada se derriba. Sin embargo, en realidad la conexión causal que provocó la caída de esa parte del puente no fue la dinamitación sino la fractura de la arcada. Así, la conexión ficticia podría ser incluida en el esquema D-N cuando no debería ser considerada como una explicación del hecho en cuestión²⁵, y por tanto, el modelo no sería suficiente para la caracterización de una explicación de este tipo.

La importancia del ejemplo del puente de Scriven radica en que aún a pesar de que podemos construir una explicación tipo D-N del fenómeno, cuyas premisas son completamente verdaderas (es verdad, por ejemplo que se dinamitó la arcada)²⁶, tal explicación no es la explicación correcta del hecho. Esto es, el modelo D-N incluye demasiado. La crítica de Scriven en este punto parece realmente contundente, pues muestra casos que podrían intuitivamente no ser admitidos como explicaciones, que sin embargo quedarían modelados por el esquema D-N junto con su requisito de adecuación empírica. Finalmente el ejemplo también nos invita a tomar en cuenta las relaciones entre los hechos mismos y no sólo entre sus descripciones, lo cual es una de las consecuencias

²⁴Bajo el supuesto de que queremos construir explicaciones completas.

²⁵Esto es un caso típico de sobredeterminación causal.

²⁶ Así, cumpliría no sólo con deducibilidad, sino que también con el requisito empírico de adecuación.

que Scriven quiere proponer, a favor de la crítica que hemos ordenado más arriba con el numeral 2.

Las propuestas más importantes de Scriven para la elucidación de la explicación son las que siguen:

(A) La distinción entre explicación y fundamentos de la explicación.

(B) La incorporación de las nociones de entendimiento y de contexto.

A es una propuesta interesante. Al parecer, Scriven entiende por "fundamentos de la explicación" los datos que apoyan la creencia de que el *explanans* de una explicación es verdadero²⁷. Por ejemplo, en el caso del puente, las pruebas de laboratorio acerca de la estructura química y la resistencia de la arcada en cuestión, etc. Para Scriven no debería pretenderse que la explicación contenga sus propios fundamentos. Para justificar esta afirmación Scriven propone primero un argumento analógico. Para Scriven la justificación de una proposición, entendida como los datos que apoyan la creencia en que dicha proposición es verdadera, no tiene por qué estar incluida en la propia proposición. Aún más, usualmente no hay tal inclusión. Análogamente, una explicación no necesita incluir sus fundamentos. Tampoco es usual que los incluya [Scriven 1962, p. 67]. Esto supone una apelación a las circunstancias usuales en que aparece una explicación, lo cual es un supuesto fundamental de toda la propuesta de Scriven. La propuesta A identifica la explicación con el *explanans* y no exige la inclusión de todos sus fundamentos, pero observemos que la justificación para esta simplificación marca una tendencia hacia la modelación de los aspectos básicos en que pragmáticamente aparece la explicación. La idea básica a la que apunta la argumentación²⁸ de Scriven, puede reconstruirse mediante dos *desideratas*: **minimalidad y enfoque pragmático** en la reconstrucción de la explicación.

²⁷ Es decir, que cada una de sus premisas es verdadera.

²⁸ Tanto en el caso de su crítica como en sus aportaciones más positivas.

Pretender incluir todos los fundamentos es imposible. Esto se debe, como señala Scriven, a que no tiene sentido dar una fundamentación completa de una explicación, que sea ajena al contexto. La razón es que un fundamento es siempre en contra de alguna duda específica o desacuerdo y hay un número infinito de dudas o desacuerdos. Por tanto, habrá un número infinito de fundamentos para la explicación [Scriven 1962, p. 67-8]. Por supuesto, con este nuevo argumento Scriven ataca la posibilidad de tener una explicación completa (con *todos* los fundamentos) sobre un fenómeno y no arremete directamente en contra de la posibilidad de incluir sólo algunos de estos fundamentos²⁹. En primera instancia, el argumento fundamenta bien que un modelo no podría capturar todos los fundamentos del *explanans*. Sin soslayar lo anterior, desde mi punto de vista, hay una dirección más remota a la que intenta dirigirse el argumento. En el fondo, la idea de Scriven es mostrar la importancia del contexto en la clarificación de las explicaciones. Así, puede bien suponerse que con este argumento Scriven intenta hacer tres cosas más: a) prepara el terreno para mostrar el papel del contexto, encarnado en las dudas o los desacuerdos, en los fundamentos de las explicaciones; b) apoyar con esto su intención de distinguir mejor la explicación de los fundamentos de la explicación y con ello, c) relativizar al contexto las condiciones que debería cumplir algo para ser considerado una explicación.

El fortalecimiento de las implicaciones a-c anteriores requerirá para Scriven reforzar la idea de que el contexto tiene un rol importante en los fundamentos de la explicación. En este mismo artículo de 1962 intentará demostrar que el contexto determina el valor que asignamos a ciertos fundamentos de una explicación y las expectativas que tenemos acerca de éstos cuando exigimos una explicación. Para ello argumentará en favor de dos afirmaciones³⁰: i) que el requerimiento de deducibilidad de D-N (a partir de leyes y condiciones iniciales) no es suficiente para determinar cuándo

²⁹Pero indirectamente este argumento podría usarse (no está claro que Scriven lo use en esta forma), aunque débilmente, en contra de la inclusión de los fundamentos en la explicación: Si no todos los fundamentos pueden contemplarse en la explicación entonces ¿Cómo elegir los que se van a contemplar?

³⁰Esta es una reconstrucción de una parte de su argumentación de las páginas 67-70.

estamos ante la verdadera explicación del fenómeno, y ii) que la subsunción legal³¹ del modelo D-N no es necesaria: hay casos de explicación que no la cumplen. El argumento en apoyo de i es el caso del puente al que me referí antes, cuando describía la forma en que apoya la crítica 4 contra el modelo D-N de Hempel y Oppenheim. i le permite a Scriven mostrar cómo el contexto, a saber, lo que sucedió realmente, le resta valor al requerimiento de deducibilidad de D-N como fundamento adecuado de la explicación. Scriven construye otro ejemplo para apoyar ii: Supongamos que golpeamos accidentalmente una botella de tinta y esto provoca que se derrame sobre la alfombra. Ante este hecho, dice Scriven, estamos en disposición de explicar a nuestra esposa el hecho de que la alfombra tiene una mancha de tinta mediante una descripción de los hechos que causaron la mancha. Podemos aducir que íbamos a tomar el cigarrillo y accidentalmente golpeamos la botella, lo que provocó a su vez que ésta cayera en la alfombra y, así, que la tinta se derramara sobre ella produciendo esa horrible mancha. Lo importante en este ejemplo es que, según Scriven, no se necesita apelar a ninguna ley. La propuesta de Scriven es que una descripción de los hechos que causalmente provocaron al *explanandum* es suficiente (minimalidad) para considerarla como el *explanans* de una explicación³². Las leyes no son, en este contexto cotidiano, dice Scriven, necesarias para construir la explicación. El contexto produce ciertas expectativas respecto de los fundamentos que esperamos encontrar en el *explanans*. Scriven dice, finalmente que, hablando del otro ejemplo, ante la duda de qué provocó la caída del puente, apelar a una ley como fundamento no garantiza la solución correcta de la duda, por lo que no es suficiente para resolver esta última y, además, ante la duda de qué provocó la mancha en la alfombra, la apelación a una ley no es necesaria. Será importante tener en mente con esta parte de la argumentación Scriven está extendiendo su enfoque pragmático: no sólo supone que una modelación de la explicación debería apegarse al uso de las explicaciones, sino que también sugiere que algunos elementos pragmáticos, es decir, algunos elementos del contexto que podrían afectar el significado de una explicación, deberían ser

³¹Es decir, el requerimiento de que el *explanans* contenga al menos una ley, lo que expusimos como la condición 1 de Hempel. Véase, *supra*, 1.1.1., El Modelo D-N.

³²Además, Scriven piensa que a veces no es posible citar las leyes que subyacen a tales hechos.

contemplados a la hora de diseñar esta modelación. Deja abierta la discusión, más allá de su pragmatismo inicial, sobre si algunos elementos pragmáticos deberían contemplarse para modelar la explicación.

Más allá de que Scriven pueda o no tener razón sobre su evaluación del requisito de legalidad, de un modelo de explicación, sus argumentos dependen fuertemente de la inscripción del fenómeno explicativo en un contexto dado. Además, si bien es cierto que Scriven logra dar razones en favor del papel importante que el contexto juega en la explicación, sus argumentos particulares dependen del supuesto de que la caracterización de la explicación debe buscarse en la forma en que realmente los sujetos pretendemos explicar un hecho, es decir, en su pragmatismo inicial. Esta es la clase de pragmatismo al que apuntan principalmente sus disquisiciones.

Scriven desprende algunas consecuencias importantes de su argumentación. Si los fundamentos de una explicación no tienen que aparecer dentro de la propia explicación entonces es posible distinguir algunas propiedades nucleares de la explicación de algunas que son más accidentales, aunque no menos importantes. Así por ejemplo, la explicación puede contener proposiciones falsas y no por ello dejaría de ser explicación. De ello surgen naturalmente algunas categorías mencionadas por Scriven que expresan posibles componentes adicionales de una explicación [Scriven 1962, p. 69-70]³³:

(A')

Si la explicación tiene proposiciones falsas, podemos llamarla "explicación incorrecta"³⁴;

Si no establece conexiones reales³⁵ entre los hechos explicativos y el hecho a explicar, podemos llamarla "explicación inadecuada";

³³No consideré aquí todas las variantes que pueden ser subsumidas bajo estos nuevos términos propuestos por Scriven.

³⁴Para Scriven dos explicaciones en conflicto (con *explanantes* mutuamente contradictorios) deben seguir, ambas, siendo consideradas como explicaciones. En este sentido Hempel y Oppenheim, dice Scriven, estudiaron la explicación correcta, no la explicación [Scriven 1962, p. 63].

Si no satisface las expectativas de su contexto, "explicación irrelevante".

El primer enunciado de *A'* es particularmente interesante. Para Scriven dos explicaciones en conflicto (con *explanantes* mutuamente contradictorios) deben seguir, ambas, siendo consideradas como explicaciones. Lo que sugiere una solución más natural para el problema de la ambigüedad epistémica. En este sentido Hempel y Oppenheim, dice Scriven, estudiaron la explicación correcta, no la explicación [Scriven 1962, p. 63].

Hay, sin embargo, un dilema que Scriven tendrá que resolver: por un lado, una tendencia hacia la minimalidad en la modelación y, por otro, una tendencia a incorporar elementos del contexto, ¿Lo último no acaba con la tentativa de minimalidad³⁶? Coherentemente, Scriven usa la implicación *b* que mencionamos antes para resolver el problema, es decir, distinguir entre explicación y sus fundamentos. La idea de separar la modelación de la explicación, y disponer también de una colección de rasgos para calificarla, también contribuye a nuestra atribución de una tendencia a minimizar la modelación de explicación. La conciliación de su tendencia a incorporar elementos pragmáticos con su tendencia a simplificar el modelo de explicación puede entenderse si pensamos que estos elementos pragmáticos quedan, por así decirlo, "fuera" de la modelación. El factor contexto, aunque es clave en esta forma pragmática de modelación, no sería problemático si lo mantenemos fuera del esquema de modelación principal. Así, Scriven, en el fondo, se encuentra criticando la empresa misma de hallar una modelación general para la explicación.

Desde esta perspectiva, la propuesta *B* de Scriven, a saber, la incorporación de las nociones de entendimiento y de contexto, se clarifica. El contexto juega un papel muy importante en la caracterización de la explicación. Scriven afirma explícitamente que la evaluación de una explicación no puede ser independiente de su contexto [Scriven 1962,

³⁵Como en el caso del puente.

³⁶Debe tenerse en cuenta que Scriven no usa ni el término pragmático ni el de minimalidad, en realidad son conceptos que uso para clarificar su propuesta.

p. 72]. Queda entonces por ver la justificación de la inferencia que tiene el entendimiento en la explicación.

La noción de contexto en Scriven parece estar amalgamada con la de conocimiento. La de entendimiento, por el contrario, parece distinguirse de esta última. Su idea de entendimiento aparece bosquejada en la nota 17 de su artículo [**Scriven 1962, p. 73**]³⁷:

"[...] El entendimiento es aproximadamente la percepción de las relaciones y, por ello, puede ser interpretado como un proceso que localiza el fenómeno enigmático en un sistema de relaciones. [...] No deducimos nada; nuestro entendimiento viene porque *vemos* el fenómeno por lo que es, y estamos *en posición* de hacer otras inferencias a partir de este reconocimiento."

Para Scriven, la noción de entendimiento puede ayudar a elucidar la de explicación. Un ejemplo de ello, según Scriven, es la pertinencia que tiene (el entendimiento) en la semejanza existente entre las llamadas por Hempel y Oppenheim "explicaciones de significado" y las explicaciones que se ajustan al modelo D-N. Por ejemplo, Scriven aduciría que la explicación de una demostración también consiste en una especie de *localización* de la proposición a demostrar en un sistema más general de relaciones con otras proposiciones.

La insistencia sobre esta noción y sobre las relaciones que guarda con el contexto y con las expectativas que tenemos de una explicación señalan una nueva manera de abordar el problema de la elucidación de la explicación científica. Esta novedad incluye, como he descrito en este apartado, la consideración de la dimensión pragmática de la explicación (a través de **A**, **A'** y **B**), la recuperación de las relaciones no lógicas entre

³⁷[...] understanding is roughly the perception of relationships and hence may be conveyed by any process which locates the **puzzling phenomenon** in a system of relations. [...] **We deduce nothing**; our understanding comes because we *see* the phenomenon for what it is, and are *in a position* to make other **inferences** from this realization."

explanans y *explanandum* (aunque débilmente, mediante la crítica 2, y más fuertemente a través del ejemplo del puente y la dinamita en 4) y el examen crítico de la idea de que la relación entre estos componentes es caracterizable en términos de una relación inferencial deductiva (por 3 y 4).

Así, el tipo de racionalidad al que apuntan las disquisiciones de Scriven, tiende más bien a flexibilizar los modelos clásicos, mediante la eliminación de la restricción de legalidad, la contextualización vinculada a la recuperación de relaciones no deductivas, y la apertura a modelos de explicación no basados en relaciones deductivas. Tal flexibilización, sin embargo, podría hacer más complejo el modelo resultante, pues está vinculada a la incorporación de elementos pragmáticos, a la caracterización de rasgos calificativos de la explicación modelada y a una contextualización general, respecto de la cual debería evaluarse si conviene su incorporación al modelo o si sería más prudente mantener al margen, como parece sugerir Scriven.

1.3 La Tensión Hempel-Coffa

A diferencia de Michael Scriven, quien tiende a pensar la explicación como dependiente de ciertos contextos, Alberto Coffa pretende eliminar elementos de dependencia contextual. Mientras el objeto de análisis en [Scriven 1962] es el modelo D-N de explicación, el de [Coffa 1974]³⁸ es el modelo I-S. En este apartado describiré los principales puntos de análisis que hace Coffa en este artículo.

En su artículo de 1974, Coffa dice que pretende hacer un análisis de la naturaleza de las diferencias entre ambos modelos (D-N, I-S) y rechazar la perspectiva hempeliana de entender las explicaciones inductivas [Coffa 1974, p. 56].

³⁸ Coffa, Alberto J., "Hempel's Ambiguity", en *Synthese*, 28, 1974, pp. 141-63. Usaré la paginación de su reimpresión en [Ruben 1993].

Básicamente, para Coffa, la distinción entre el modelo D-N y el modelo I-S puede verse como una distinción entre los niveles epistémico y óntico. El modelo D-N supone que las explicaciones pueden ser verdaderas y, con ello, que pueden establecer relaciones objetivas reales entre los hechos contenidos en el *explanans* y el hecho descrito por el *explanandum*. En contraste, el modelo I-S supone explicaciones, dice Coffa, en un nivel no óntico, sino epistémico. Una explicación I-S postula relaciones en el nivel de nuestras expectativas acerca de los vínculos entre los hechos involucrados en la explicación.

Para entender mejor esta diferencia, puede acudir a la noción de relativización epistémica. Coffa empieza distinguiendo entre expresiones epistémicas y expresiones no epistémicas. Las primeras refieren al conocimiento, las segundas no. La idea de confirmación es una idea que requiere de una referencia a un conjunto particular de conocimientos. Así, también la idea de explicación D-N bien confirmada puede referirse a un conjunto particular de conocimiento. Bajo el análisis de Hempel, una explicación D-N bien confirmada para un conjunto particular K de conocimientos puede entenderse como una que *es racional tomar, en K, como una explicación D-N verdadera*. Este es un sentido en el cual podemos entender la propiedad de ser **epistémico**. Sin embargo, dice Coffa, después de encontrarse con el problema de la ambigüedad epistémica los resultados de Hempel no otorgan este sentido de "epistémico" a las explicaciones inductivas. Las explicaciones inductivas no pueden ser tomadas racionalmente como explicaciones verdaderas para un conjunto de conocimientos K puesto que no tiene sentido hablar de explicaciones I-S verdaderas: son epistémicas no confirmacionales. A este hecho Coffa le llama la "**tesis de la relatividad epistémica de la explicación inductiva**" [Coffa 1974, p. 73-4].

Obsérvese que he hablado de explicaciones inductivas y de explicaciones I-S. Como Coffa distingue, la ambigüedad epistémica es una propiedad que se atribuye a definiciones de explicación inductiva cuando producen inconsistencias inductivas [Coffa

1974, p. 59]. Abordando el problema de la ambigüedad epistémica, puede existir otra forma de entender las explicaciones inductivas que sea diferente de su entendimiento como explicaciones I-S. Para Coffa la diferencia consistiría en rechazar la tesis de la relatividad epistémica para las explicaciones inductivas.

El camino que lleva a Hempel desde la ambigüedad epistémica a la tesis de la relatividad epistémica atraviesa, según Coffa, las siguientes suposiciones [**Coffa 1974, p. 66**]:

- 1.- Una definición de explicación inductiva es inadmisibles si sufre de ambigüedad.
- 2.- La única forma de mejorar la situación consiste en introducir la cláusula de la máxima clase de referencia (RME).
- 3.- No hay una cláusula de este estilo que a) no haga referencia al conocimiento y b) que garantice la ausencia de ambigüedad.

RME obliga a que tengamos todo el conocimiento disponible antes de establecer relaciones explicativas entre dos hechos. Sin embargo para Coffa RME de Hempel es una versión "epistemologizada" de un requerimiento que originalmente era óntico: La demanda por referir a todos los aspectos relevantes al fenómeno *explanandum*. La forma en que podemos manejar RME nos obliga a relativizar el conocimiento. En realidad, lo que hace que la demanda contenida en RME, dice Coffa, "sea factible es la ignorancia más que el conocimiento" [**Coffa 1974, p. 70**], pues tal vez si tuviéramos más datos al respecto del caso de explicación no podríamos explicar los hechos que explicamos. Sin embargo, si nuestras explicaciones son tan sensibles al aumento de información, para Coffa es natural preguntarse: "¿Qué hay en este argumento inductivo que lo hace una explicación de su última fórmula? ¿Qué razón pudiera alguien tener para decir que es un explicación de su conclusión?" [**Ibid., p. 71**]. Esto hace que Coffa vea a RME como un principio que es, a la larga, inútil.

En las preguntas anteriores Coffa se encuentra pisando el terreno de un problema que será muy importante en la discusión posterior sobre la explicación: el problema de saber qué tipo de relaciones legitiman la relación explicativa entre dos hechos (o entre sus descripciones). Llamaré a esto el "problema de las relaciones de pertinencia. No obstante, también señala otro problema importante, que le hace pensar que la modelación de Hempel está errada: un aumento de información podría impedir que infiriéramos el *explanandum*.³⁹ Para Coffa, el que nuestra aceptación de un argumento como una explicación sea sensible a la nueva información hace que sospechemos si se trataba realmente de una explicación. Su idea de explicación está nítidamente comprometida con un ideal de adecuación con la realidad. Pero este ideal parece cegar a Coffa respecto del comportamiento real de nuestras explicaciones en Ciencia. Si optamos por el diseño de un modelo de explicación, con un enfoque pragmático, en el sentido tan sólo del pragmatismo inicial de Scriven, tendríamos que reconsiderar si deberíamos o no, como una cuestión justamente de adecuación a la realidad, capturar este aspecto de las explicaciones: su sensibilidad al cambio en la información de nuestro contexto.

Su idea de explicación claramente es diferente a la de Scriven. Para Coffa las explicaciones tienen que ser ópticamente correctas, no pueden reducirse a meras especulaciones a nivel epistémico esto es un resultado en una dirección contraria a la que apunta el ejemplo del puente de Scriven: para Coffa una explicación tendría que satisfacer las relaciones que en realidad se dan en el fenómeno respecto del cual se construye la explicación; en cambio, para Scriven, la misma idea de que el argumento del puente no corresponde con la realidad, lo hace pensar que la explicación no debería contener sus propios fundamentos y, así, a postular argumentos que sean explicaciones aunque no sean explicaciones adecuadas.

La idea de que un argumento I-S en realidad no modela una explicación viene reforzada, en Coffa, además, por el hecho de que la interpretación de la premisa general

³⁹ Esta es una de las características que enfatizaré en esta investigación.

que expresa la relación probabilística en un argumento I-S es interpretada por Hempel desde el punto de vista frecuentista de las probabilidades. Tal interpretación, dice Coffa, no sólo de esta premisa sino de la relación de derivación en un argumento I-S, hace que Hempel se concentre en la solución del problema de la clase de referencia, pues este problema es tradicional en este enfoque de la teoría de la probabilidad [Coffa 1974, p. 61].

Coffa piensa que el mundo posee características que pueden ser determinísticamente responsables para otras **si y sólo si** podemos definir un modelo de **explicación inductiva verdadera**. Para Coffa, entonces, Hempel no cree en estas características dada la relativización epistémica y su interpretación frecuentista del argumento inductivo. Estas características son las que deben ser descritas en una explicación inductiva, y no precisamente las relaciones estadísticas. Es decir, no precisamente las relaciones de probabilidad frecuentista. Una explicación de este tipo (frecuentista) no tendría, en este sentido, ningún interés; pues no nos diría nada acerca del mundo [Coffa 1974, p. 71].

La **relevancia estadística** no es entonces una relación deseable para modelar la explicación inductiva. Más fuertemente, Coffa piensa que la relevancia estadística no es ni necesaria ni suficiente para determinar si una clase de referencia es la clase apropiada [Coffa 1974, p. 75].

Coffa opina que la relación entre el *explanans* y el *explanandum* en una explicación debe establecer alguna **pertinencia nómica** que ocurra realmente [Coffa 1974, p. 74]. Desde esta perspectiva, bajo los supuestos hempelianos la existencia de explicaciones inductivas desaparece. Debemos entonces, concluye Coffa, rechazar algún supuesto hempeliano. La solución puede ser un enfoque óntico y no epistémico de RME. Coffa sugiere con algunas reservas que esto podría efectuarse mediante la concepción de probabilidad como propensiones [Ibid., 76].

La crítica de Coffa pone de relieve varios aspectos de la explicación, que es importante recapitular. La tesis de la relatividad epistémica supone que la explicación es construida en un plano epistémico y no, como parece que debería ser el caso, en uno óntico. Debemos, dice Coffa, buscar definiciones alternativas de explicación inductiva que rechacen la tesis de la relatividad epistémica. Según Coffa, los argumentos I-S no corresponden con nuestra expectativa racional de lo que es una inferencia que modela la explicación. RME revela la falta de pertinencia entre los elementos de una explicación. Las relaciones de pertinencia entre el *explanans* y el *explanandum* no pueden modelarse postulando relaciones estadísticas. Estas no son la clase de conexiones reales que una explicación debe exigir.

Consecuentemente, la tendencia de Coffa es más bien hacia la eliminación de contextualizaciones para la explicación, pero no solo eso, sino también a la idea de que una relación probabilística basada en propensiones puede ayudar a sustituir nuestra idea de explicaciones que asumen relaciones de inferencia más débiles que la deductiva. El móvil de la tendencia de Coffa es evitar la contextualización epistémica de las explicaciones, es decir, el elemento que habría que evitar es la relativización, pues implica un rasgo indeseable, la pérdida de inferencias debido al aumento de información en el contexto sería como asumir que el mundo no posee características que son determinístamente responsables para otras. Así, la propuesta de Coffa se dirige a reconstruir algunas explicaciones basándose en propensiones, porque con ello puede incorporar lo que Scriven llamó “adecuación”: la referencia a relaciones realmente ocurridas en el fenómeno. Coffa pretende entonces modelar la explicación “real”, en el sentido de que algo no puede ser una explicación si las relaciones que atribuye al fenómeno modelado no son reales, es decir, una explicación es siempre ónticamente adecuada en el sentido que vimos con Scriven. A diferencia de Scriven, entonces, Coffa apunta a un tipo de racionalidad más dura, una racionalidad que intenta eliminar relativizaciones, desaparecer repercusiones de cambio de información, y que supone una

representación determinista sobre las relaciones, ópticamente entendidas, del fenómeno explicativo.

1.4 Otros Modelos de Explicación Científica

Hasta ahora he presentado los dos modelos fundamentales defendidos por Hempel: el modelo D-N y el modelo I-S. Además he resumido las principales tesis contenidas en dos de los artículos críticos de los modelos: [Scriven 1962] y [Coffa 1974]. Por un lado, la tensión Hempel-Scriven es una discusión que pone sobre la mesa de análisis la adecuación del rigor lógico del modelo D-N. Por otro lado, y en un sentido contrario, la tensión Hempel-Coffa es una discusión sobre la demanda de fundamentos racionales objetivos para el modelo I-S de explicación. Podemos decir que la primera tensión surge por una propuesta para la relajación de un sentido de racionalidad "dura" y que, en cambio, la segunda se debe a una propuesta que se dirige hacia un sentido de racionalidad "dura". El debate posterior sobre la explicación científica puede analizarse mediante el uso de estas dos tensiones como postes indicadores para comprender la discusión. En general los modelos y teorías sobre la explicación propuestos posteriormente intentaron debilitar la idea de explicación deductiva y objetiva, o fortalecer la de explicación inductiva y contextualizada. Además, ambas tensiones pueden mostrar que la discusión sobre la explicación involucra también un debate sobre si la racionalidad científica, vista a través del caso de la explicación, debería suponer relaciones determinísticas, de modo que un modelo de explicación debería ser una representación de las relaciones ópticas, reales, del fenómeno, o más bien, por otra parte, no debería tener tal supuesto y concebir esta representación ubicada en un plano más epistémico, relativizado, e involucrando elementos pragmáticos.

Una pregunta que surge naturalmente después del análisis de estas dos tensiones es si es posible integrar a un esquema argumentativo, o bien si es posible establecer en un

conjunto de condiciones, algunos elementos que puedan dar lugar a una mejor representación de la explicación científica. Por un lado, la crítica de Scriven contra la suficiencia y la necesidad de un modelo deductivo de explicación, nos invita a pensar si será prudente seguir confiando en una **representación argumental** de la explicación. Queda la duda, entonces, de si un modelo argumental de la explicación puede dar cuenta de éste fenómeno de la actividad científica. Por otro lado, después de la crítica de Coffa, uno puede pensar si habrá forma de establecer de una vez y para siempre un conjunto de relaciones fácticas determinísticas tales que constituyan el núcleo de la relación requerida entre los sucesos a los que hacen referencia las oraciones de una explicación. La pregunta en este caso, más precisamente, es la de si es posible caracterizar contexto-independientemente las **relaciones relevantes** que debería cumplir una explicación.

En este apartado presentaré tres modelos más cuya propuesta creo que esclarece los elementos en juego en el debate sobre la explicación a partir de los modelos clásicos: el modelo de Relevancia Estadística (S-R), de Wesley C. Salmon; el modelo Nomológico-Deductivo de la Explicación Probabilística (D-N-P), de Peter Railton; y el modelo que llamaremos "Erotético-Contextual" (E-C), de Bas van Fraassen. La intención de presentar estos modelos es la de observar, con mayor detenimiento, los problemas que están involucrados en una **representación inferencial** de la explicación científica. La exposición de cada modelo será muy breve. En cada uno me detendré un poco para mencionar algunas tesis importantes subyacentes en la concepción del modelo, que nos llevarán a entender mejor el tipo de supuestos que subyacen al debate y las dificultades que tendría que enfrentar una **modelación argumental** de la explicación en el contexto de la Filosofía de la Ciencia. Los tres modelos que veremos guardan una relación importante con la tentativa de construir modelos argumentales de la explicación y muestran rasgos analizables desde los elementos mencionados cuando describimos las tensiones Hempel-Scriven y Hempel-Coffa.

1.4.1 Salmon

El modelo (S-R) de Salmon posee tres principales propiedades: es probabilístico, no argumental y pretende reflejar relaciones objetivas causales. En sus primeras formulaciones el modelo no satisface completamente la última pretensión. Por ello, en su libro [Salmon 1984b]⁴⁰, Salmon intentó hacer un análisis sobre la causalidad que lo llevaría a una mejor elucidación de la explicación. En este último texto Salmon consideró que los desarrollos que había pretendido que fueran un modelo de la explicación científica en [Salmon 1971]⁴¹ (cuando apareció por primera vez), constituían más bien una fracción, aunque muy importante, para un modelo de explicación. Así, él llamó a estos desarrollos "base S-R para una explicación".

De cualquier forma, los textos de Salmon dicen mucho acerca de su idea de explicación e influyeron el curso de la discusión en este tópico. La propia base S-R contribuye mucho al esclarecimiento de la explicación. Es imprescindible enfatizar que Salmon rechaza que un modelo de explicación pueda ser argumental. Su modelo de explicación está pensado como un conjunto de datos acerca de los factores que son estadísticamente relevantes al hecho que se pretende explicar.

Para Salmon es problemático construir el esquema de una explicación estadística mediante la estructura de un argumento. El problema básico que él ve es que el requerimiento de alta probabilidad para la obtención de la conclusión frecuentemente no se cumple en las explicaciones⁴². Examinaré esto con más detalle más adelante.

⁴⁰Salmon, Wesley C., "Scientific Explanation and the Causal Structure of the World", Princeton University Press, 1984. La paginación referirá en adelante a [Salmon 1993].

⁴¹Salmon, Wesley C., *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, 1971. Usaré para su referencia [Salmon 1988].

⁴²Salmon 1988, *op. cit.*, p. 78

Así, la explicación para Salmon posee necesariamente un conjunto de datos que incluyen factores estadísticamente relevantes acerca del *explanandum*. Otra característica peculiar es que la relevancia estadística de los factores incluidos en una explicación puede ser negativa. Su idea de relevancia estadística se define como sigue: Un factor *C* es relevante al *explanandum B* en las circunstancias *A*, si se cumple la siguiente condición [Salmon 1993, p. 95]:

$$P(B/ A \wedge C) \neq P(B/ A)$$

La idea más clara que puede desprenderse de los textos de Salmon⁴³, acerca de esta relevancia que, obsérvese, es posiblemente negativa, es que una explicación completa debería incluir tanto los factores que contribuyen a la ocurrencia del *explanandum* como los factores que actúan en contra de tal ocurrencia. Ambos tipos de factores interactúan entre sí modificando la probabilidad de la ocurrencia del *explanandum*. El resultado de la combinación de ambos tipos de factores debería tan sólo cambiar la probabilidad del *explanandum* y en el caso de aumentarla no necesariamente hacerlo con un alto grado probabilístico [Salmon 1993, pp. 106-8].

Expondré en seguida la base S-R para una explicación. Considérese el siguiente ejemplo [Ibíd., p. 99]⁴⁴: Roberto ha cometido un acto delictivo y nos preguntamos por la explicación de este hecho. Nuestro contexto es que Roberto es un adolescente. La pregunta entonces es ¿Por qué este adolescente cometió este acto delictivo?

La base S-R puede construirse mediante un procedimiento preciso. El procedimiento consta de 8 pasos [Salmon 1993, pp. 98-102]. Describiré en cada paso lo concerniente al ejemplo anterior.

⁴³Por ejemplo en [Salmon 1993, p. 108].

⁴⁴El ejemplo es una reconstrucción del de [Greeno, 1971]: Greeno, "Theoretical Entities in Statistical Explanation" en Roger C. Buck y Robert S. Cohen (eds.) *PSA 1970*, D. Reidel Publishing Co., Dordrecht, 1971.

(Paso 1) Seleccionamos una clase de referencia apropiada.

B representa la clase general de los actos delictivos, una clase de referencia adecuada para B en este *explanandum* es la clase de los adolescentes. Designemos esta clase mediante A . Con esta clase estimaremos las probabilidades antecedentes⁴⁵: $P(B/A)$.

(Paso 2) Seleccionamos una partición B_1, \dots, B_m del *explanandum* B sobre la clase A . Esto significa que la partición de B debe ser consecuente con el hecho de que A es la clase de referencia. En este ejemplo, los actos delictivos pueden clasificarse de cierta manera para el caso de los adolescentes. Por ejemplo la partición de B en: B_1 , la clase de delitos por infracciones menores, B_2 , la de delitos por ofensas mayores. Pero quizá si la clase de referencia es la clase de los adultos, la clasificación de los actos delictivos sea distinta.

(Paso 3) Apelamos a un conjunto de factores estadísticamente relevantes, representados por las clases C_1, \dots, C_s . Esta será la partición del *explanans*. Y realizamos una partición de A en combinación con C como las que siguen: $A \wedge C_1, \dots, A \wedge C_s$.⁴⁶ Estas combinaciones se llaman "celdas". El conjunto de factores C del *explanans* puede ser en este ejemplo las siguientes clases: la de los hombres, la de las mujeres, los judíos, los protestantes, los católicos, los ateos, la gente de posición económicamente alta, de posición económicamente baja, los tímidos, los arrojados, los rebeldes, etc.

(Paso 4) Estimamos la lista de probabilidades antecedentes y posteriores.

Antecedentes:

$$P(B_i/A) = p_i$$

⁴⁵"prior probabilities".

⁴⁶ Acá la conjunción es en realidad una intersección en las clases.

para toda i ($1 \leq i \leq m$)

Posteriores:

$$P(B_i/A \wedge C_j) = p_{ij}$$

para toda i y toda j ($1 \leq i \leq m$) y ($1 \leq j \leq s$).

(Paso 5) Requerimiento de Homogeneidad: toda celda debe ser homogénea en relación a la partición del *explanandum* B_i . Esto es, ninguna de las celdas en la partición es susceptible de subdividirse en algunas otras celdas que sean también estadísticamente relevantes para algún miembro de la partición del *explanandum*. Este requisito es análogo al RME de Hempel, nos obliga a tener todos los factores relevantes.

(Paso 6) Estimamos el tamaño relativo de las celdas, como sigue:

$$P(C_j/A) = q_j$$

La razón de este paso es que en ocasiones algunos factores C_j pueden ser separadamente relevantes en forma positiva para un miembro de la partición del *explanandum* B_i y, sin embargo, en combinación pueden ser relevantes negativamente. Es necesario entonces saber las estimaciones separadamente; pues esto permitirá estimar la probabilidad de B_i dada la combinación de varios factores C_j .

(Paso 7) Requerimiento de Máxima Homogeneidad: La partición del *explanans* debe tener, para todo $i \neq k$, probabilidades tales que $p_{ji} \neq p_{jk}$. Lo cual intenta ser una restricción en contra de que se repitan celdas con la misma relevancia. Por ejemplo restringe usar el vector sexo si la relevancia estadística respecto de un tipo de acto delictivo es igual para hombres que para mujeres. Esto nos obliga a quedarnos sólo con los factores necesarios para la estimación.

(Paso 8) Determinamos a cuál celda pertenece el individuo al que se refiere el *explanandum*.

Es importante precisar algunas ideas generales:

1) Para Salmon es importante tener a la mano todos los factores relevantes al *explanandum*. Sin embargo, él piensa que esto no es factible si no consideramos al *explanandum* respecto de una circunstancia particular. Todo *explanandum* es entendible sólo respecto de una circunstancia particular que le da sentido [Salmon 1993, p. 95-7]. Esto es tomado en cuenta en el Paso 1 de su modelo.

2) El requisito de *homogeneidad* es un requisito que refiere a relaciones "objetivas". Esto es, no relativizadas epistémicamente [Salmon 1984, p. 103]. Esta es una diferencia importante respecto a RME de Hempel. La intención de Salmon en este respecto está dirigida en dirección de Coffa, hacia la eliminación de elementos epistémicos. Sin embargo, es importante notar que la idea de relaciones objetivas, en el caso de Salmon no necesariamente se compromete con que la explicación capture los sucesos que realmente ocurrieron, es decir, no necesariamente estaría comprometido con lo que Scriven llamaría una explicación adecuada. Piénsese que puedo asegurar que los datos estadísticos sobre el derrumbe de puentes⁴⁷ se refieran a relaciones objetivas en el mundo, no excluye la posibilidad de que esos datos no correspondan con lo que realmente le sucedió al puente.

3) Para Salmon, el modelo de explicación tipo D-N es definitivamente inadecuado. La propuesta de Salmon es un intento por modelar todo tipo de explicación en términos probabilísticos. Esto puede entenderse como un movimiento hacia el

⁴⁷ Me refiero al ejemplo de Scriven sobre el puente. Véase *supra*, sección 1.2., *La Tensión Hempel-Scriven*.

extremo de Scriven en la tensión Hempel-Scriven, en el sentido de que la probabilidad permite explicaciones más débiles que no podrían admitirse en el modelo D-N: explicaciones de baja probabilidad. No obstante, en la forma en que Salmon aborda la explicación estadística, hay también un elemento que claramente lo sitúa como un seguidor del ideal de objetividad buscado por Coffa: el requerimiento de homogeneidad, que es un claro intento por hacer más objetivas las explicaciones estadísticas hempelianas, pues no se trata de los datos disponibles, sino de la asunción de que se tienen a la mano todos los factores relevantes. Por otro lado, la idea de que hay explicaciones con un bajo grado de probabilidad y el rechazo de la asunción argumental, pueden entenderse como propuestas de debilitación de los requerimientos generales de una explicación, consecuentemente, como propuestas en la dirección de Scriven. Salmon, al hacer entrar en juego ambas propuestas, abrirá caminos alternativos para clarificar la noción de explicación.

4) Debe notarse, también, que de acuerdo a su paso 6 Salmon, como consecuencia de su uso del cálculo de probabilidades, introduce la idea de que podría haber factores que, solos darían al *explanandum* más probabilidad que si se acompañaran de otros factores. Puede entenderse esto como si la probabilidad del *explanandum* dependiera de una función que es **no monotónica**: que cuando se aumentan argumentos a la función, el resultado puede cambiar. Si el modelo de Salmon fuera argumental, quizá tendría que establecer un límite de baja probabilidad tal que si se traspasase, el *explanandum* ya no sería inferible por ese conjunto de premisas. Tal como funciona la propuesta de no monotonicidad a la que hice referencia en la Introducción. Sin embargo, Salmon no podría sostener algo así pues no puede haber argumentos con baja probabilidad.

Finalmente, Salmon piensa que ningún modelo podrá ser completo si no incluye un análisis de la causalidad. La presencia de la causalidad en la explicación es uno de los

factores que lo llevaron a pensar que el modelo S-R de explicación debería tomarse más bien como siendo apenas una base para una explicación estadística. Por ello dedicará muchos esfuerzos en elucidar la noción de causalidad. La necesidad de una noción de causalidad en Salmon está estrechamente ligada a su ideal de búsqueda de relaciones objetivas, la explicación tiene que versar sobre relaciones claramente establecidas en el mundo, esto posibilitaría, y quizá éste sea el objetivo, la oportunidad de tener explicaciones adecuadas en el sentido que hemos visto en Scriven. Así, su incorporación de una teoría de la causalidad parece un movimiento en el sentido de Coffa.

1.4.2 Railton

No obstante lo anterior, Salmon también incorpora por lo menos dos elementos a la explicación, que lo acercan a la posición de Scriven: la baja probabilidad y la no monotonidad. Si Salmon es radical (en contraste con los modelos clásicos de Hempel) cuando supone que la relevancia estadística contiene casos de relaciones de baja probabilidad, Railton lo es aún más: con su modelo D-N-P, en [Railton 1978]⁴⁸, este autor intenta capturar no sólo casos de baja probabilidad para el *explanandum* sino además casos de *explananda* azarosos. La importancia de poder representar este tipo de casos está directamente relacionada con la tentativa de dar cuenta de los casos de hechos objetivamente azarosos que contempla la física cuántica.

Railton piensa que hay casos de explicación no argumental y que su modelo representa algunos de estos casos. La dificultad de representar casos de *explananda* azarosos mediante un argumento se implica del conflicto de derivar algo que puede no suceder, así tenga alta o baja probabilidad de ocurrir respecto de ciertas circunstancias.

⁴⁸ Railton, Peter, "A Deductive-Nomological Model of Probabilistic Explanation", en *Philosophy of Science*, 45, 1978, pp. 206-226. Referiré a la paginación de su reimpresión en [Pitt 1988].

Para Railton, el modelo D-N-P representa la explicación no como un argumento sino como una "forma de dar cuenta"⁴⁹. La idea general es la siguiente: normalmente los científicos ofrecen explicaciones deductivas o bien bosquejos de explicaciones deductivas. Una explicación deductiva tiene limitaciones para explicar cierto tipo de *explananda*, a saber, los *explananda* que representan procesos indeterministas. Puede verse que la preocupación de Railton tiene un presupuesto que es contradictorio con la propuesta de Coffa, a saber, que es posible que a una explicación le subyazcan relaciones objetivas no determinísticas, cosa que Coffa no consideró en su crítica contra el modelo clásico I-S. Podría esperarse entonces que el intento por modelar relaciones de este tipo llevaría a Railton a la consideración de estructuras argumentativas no deductivas, tal vez probabilísticas, al componer su modelo de explicación. Por el contrario, Railton opta por preservar las deductivas. Railton construirá un modelo de explicación no argumental pero mediante relaciones deductivas y algunos complementos dirigidos a dos objetivos: que sus premisas intenten elucidar los *mecanismos* explicativos del *explanandum* y que sea capaz de capturar mecanismos indeterministas.

La tendencia deductivista de Railton es consecuencia de su idea de que los argumentos inductivos son la razón de que las explicaciones estadísticas de Hempel exijan tanto un alto grado de probabilidad como ser relativizadas epistémicamente [Railton 1978, p. 123]. La forma en que Railton intentará prescindir de ambas exigencias es poniendo el énfasis sobre los mecanismos que provocarían la ocurrencia del fenómeno a explicar. Railton piensa que la descripción de dichos mecanismos es, esencialmente, lo que constituye una explicación [Ibid., 124] independientemente de si el fenómeno descrito por el *explanandum* ocurre o no. Cuando hay baja probabilidad de que ocurra el fenómeno usualmente generamos descripciones de mecanismos causales que tienen, según Railton, capacidad explicativa. La baja probabilidad lleva a Railton a admitir explicaciones inadecuadas en el sentido de Scriven. Por ello Railton elimina el requisito de alta probabilidad en la explicación. El énfasis sobre los mecanismos aludidos significa para

⁴⁹Puede interpretarse el término "account" que usa Railton, como una forma particular de *dar cuenta* de un fenómeno. Más allá del modelo D-N-P, Railton no clarifica dicho término en su artículo.

Railton interpretar los enunciados probabilísticos como las *tendencias objetivas de los fenómenos* en vez de interpretar a estos enunciados como meras afirmaciones estadísticas acerca de los datos de que se dispone acerca de un fenómeno, una estrategia para eliminar, como Coffa, la relativización. Esto le permite explicar, desde un punto de vista objetivo, hechos azarosos. El costo de admitir la baja probabilidad no resulta muy alto, pues no está obligado a admitir ninguna clase de relativización, la inadecuación de la explicación no es producto de diferencias entre los datos de que disponemos, sino que se debe a una indeterminación intrínseca al estudio de ese fenómeno. Así, interesantemente, Railton, mientras intenta explicar hechos azarosos y admite explicaciones que implican baja probabilidad, lo cual parecería una tendencia hacia una debilitación en la racionalidad de la explicación, su idea de las relaciones involucradas en la explicación lo conduce, sin embargo, más a una posición como la de Coffa que a un acercamiento a Scriven.

Una consecuencia de la posición de Railton, como también lo fue de Salmon, es que no puede admitir que la explicación se modele como un argumento. La explicación no puede ser un argumento pues no admitiríamos argumentos con baja probabilidad. A pesar de ello, Railton admite que una explicación, aunque no es un argumento, contiene un argumento.

El modelo de Railton es un compuesto que contiene, entre otras cosas, un esquema de argumento deductivo y un "*addendum*". El *addendum* lo define como una expresión parentética que aclara que el evento descrito por el *explanandum* del argumento sucedió de hecho. Si el *explanandum* es un hecho azaroso entonces el argumento deductivo no podría dar como conclusión que necesariamente el hecho ocurrirá. La pertinencia del *addendum* es consolidar la relación del argumento con el hecho. Railton considera que la el argumento involucrado es muy importante para la explicación, pero es necesario explicitar cómo se relaciona con la idea de que el hecho a explicar ocurra.

El caso que Railton presenta en su artículo de 1978 es el de la explicación de la desintegración de una partícula atómica. Railton presenta primero el caso concreto de explicación y después proporciona una versión generalizada del esquema de explicación. Los presentaré en este mismo orden.

El caso concreto es el siguiente [Railton 1978, p. 125]. Supongamos que β_E es un valor de probabilidad sumamente baja, $[t_0-t_n]$ es un intervalo cualquiera de tiempo y θ es el largo de un intervalo de tiempo. Podemos tener la siguiente condición universal:

(1)

Todo núcleo de radioelemento E tiene la probabilidad β_E de emitir una alfa-partícula durante cualquier intervalo de tiempo $[t_0-t_n]$ de largo θ , a menos que esté sujeto a radiación ambiental.

Supongamos que cada instancia de (1) respecto de algún tipo particular de radioelemento es una ley. Así, podemos construir el siguiente argumento para el caso del tipo de radioelemento U y para el caso particular u :

(2)

(a) Todo núcleo de U tiene la probabilidad β_U de emitir una alfa-partícula durante cualquier intervalo de tiempo $[t_0-t_n]$ de largo θ , a menos que esté sujeto a radiación ambiental.

(b) u era un núcleo de U en el tiempo t_0 , y no estuvo sujeto a ninguna radiación ambiental antes o durante el intervalo $[t_0-t_r]_\theta$, que es un intervalo de largo θ .

(c) u tuvo la probabilidad β_U de emitir una alfa-partícula durante el intervalo $[t_0-t_r]_\theta$.

(2) sólo está ofreciendo una explicación de la probabilidad que tiene u de emitir una alfa-partícula. Pero esto puede convertirse en una explicación probabilística de la desintegración⁵⁰ de la partícula u si se le añaden algunas cosas. La explicación para Railton sería como se indica a continuación:

(3)

(3a) Una derivación de (2a) a partir de (1).

(3b) La inferencia D-N (2).

(3c) Un *addendum* al efecto de que u se alfa-desintegró de hecho durante el intervalo $[t_0-t_r]_{\theta}$.

Railton defiende que (3) es una explicación [Railton 1978, p. 126]. Él entiende a (3) como una forma de dar cuenta de un fenómeno. Las razones son: primero, la derivación de (2a) muestra que, con base en una estructura universal, hemos derivado la posibilidad del evento *explanandum*; segundo, además de (3), *nada más puede decirse* acerca del mecanismo que origina el evento *explanandum*. Lo cual significa para Railton que no hay factores desconocidos que determinen el fenómeno *explanandum* y que por lo tanto lo que se aduce para explicarlo es una descripción de los mecanismos por los que ocurriría el fenómeno. Ahora bien, Railton aclara que (3) no explica por qué *tuvo que tener* lugar el *explanandum* ni por qué *podría esperarse* que tuviera lugar, en el sentido de que no nos permite decir que el *explanandum* tendría que suceder *necesariamente* ni nos permite decir que podría esperarse la ocurrencia del mismo. Railton aclara que el *explanandum* es un evento azaroso y sumamente improbable. Por ello, más bien (3) explica "por qué la desintegración tuvo lugar de manera improbable" [Idem]. Lo cual parece querer decir que por qué el *explanandum* tuvo lugar a pesar de ser improbable que tuviera lugar. Vale la pena notar que esta idea de explicación de Railton se ubica en el contexto de una

⁵⁰Decir que un núcleo de un radioelemento emitió una alfa-partícula puede verse como equivalente a decir que ocurrió una alfa-desintegración en tal núcleo.

búsqueda para “dar cuenta” del fenómeno, que ésta es pues, la idea que desde su punto de vista elucida la noción intuitiva de dar cuenta de un fenómeno.

Railton escribe que la forma en que (3) explica lo que he mencionado en el párrafo anterior es demostrando que existe una pequeña pero definida posibilidad de que ocurra la desintegración y que, por azar, la posibilidad se realizó de hecho [Railton 1978, p. 126-7].

La caracterización formal del modelo D-N-P es la siguiente lista de fórmulas A-D contenidas en (3') [Railton 1978, p. 127-8]:

(3')

(A)

$$\forall t \forall x [Fx_t \rightarrow P(Gx_t) = p]$$

que se lee: "En cualquier tiempo, algo que es F tiene la probabilidad p de ser G".

(B)

$$Fe_{t_0}$$

"e es F en el tiempo t_0 "

(C)

$$P(Ge_{t_0}) = p$$

"e tiene la probabilidad p de ser G en el tiempo t_0 "

(D)

[Ge_{t₀}/¬Ge_{t₀}]

"(e llegó / e no llegó, a ser G en el tiempo t₀)"⁵¹

En este caso D es el *addendum* del cual (3c) era una instancia. Obsérvese que C está dado como consecuencia de A y B. Esto sugiere que cuando Railton habla de la D-N inferencia (3a), en (3), está hablando del proceso inferencial y no del esquema de argumento. Tal esquema no aparece en (3').

En otro lugar del texto, Railton intenta dejar más claro el *explanandum* que (3) explica [Railton 1978, p. 129]⁵²:

(3) explica el particular *hecho acerca* del evento observado que nos proponemos explicar, a saber, que una alfa-desintegración con energía no especificada (ni dirección, ni *momentum* angular, etc.) tuvo lugar en el núcleo *u* durante el intervalo de tiempo en cuestión.

Es importante notar que (3a) indica una caracterización teórica del mecanismo que subyace a la alfa-desintegración y que esto es fundamental para el modelo D-N-P. Railton piensa que una explicación debe referir a relaciones de pertinencia que sean "objetivas", en el sentido de que sean conexiones reales de los objetos mencionados por la explicación. Esto, como dije antes, tiene relación con su enfoque de las relaciones probabilísticas. Para Railton, la manera correcta de representar los enlaces causales

⁵¹ La diagonal corresponde a, alternativamente, llegó o no llegó.

⁵²"(3) explains the particular *fact about* the event observed that we set out to explain, namely, that an alpha-decay with unspecified energy (or direction, or angular momentum, etc.) took place at nucleus *u* during the time interval in question". Un defecto que debe señalarse, aun cuando sólo sea escuetamente, es la complicación excesiva para caracterizar el explanandum involucrado en el caso de explicación que le sirve de argumento a favor de su modelo. Railton tiene que hacer muchas aclaraciones acerca de lo que realmente se encuentra explicando su ejemplo y, a pesar de ello, permanece cierta oscuridad al respecto.

necesarios en algunas explicaciones [Railton 1978, p. 120]⁵³ es mediante una interpretación probabilística basada en propensiones [Ibid., p. 131].

Dos consecuencias claras son: a partir de que intenta incluir el modelo D-N y de que agrega el *addendum* como parte de la explicación, se libra del requerimiento de alta probabilidad postulado para las explicaciones I-S. Y a partir de su interpretación de las probabilidades, su modelo parece no tener que recurrir a ningún tipo de relativización epistémica. El modelo D-N-P se ubica en el plano óntico de las relaciones explicativas y no en el plano epistémico y esta es una tendencia que corresponde a la tensión Hempel-Coffa.

Así, la idea de capturar una explicación probabilística en una estructura deductiva y la interpretación de las probabilidades como propensiones, constituyen una tentativa claramente tendiente a la propuesta de Coffa. Sin embargo, la intención de capturar explicaciones en las que se involucra de manera fundamental un bajo grado de probabilidad y características azarosas de la ocurrencia del *explanandum*, constituye más bien una debilitación de la noción de explicación respecto del modelo I-S. La pretensión de Railton puede describirse como un intento de rigorización de la noción de explicación a la vez que como una búsqueda por ampliar el dominio de aplicación de la misma. En este intento por ampliar su dominio pareciera incluir aspectos que debilitan la racionalidad supuesta en la explicación. Por otro lado es importante enfatizar que Railton, como Coffa, insiste en la importancia que tiene tomar en cuenta los mecanismos mediante los que los hechos se encuentran realmente relacionados (adecuación en el sentido de Scriven) [Railton 1978, p. 120-1], aunque, a diferencia de Coffa, esto lo lleva no a la reconstrucción de una definición de explicación inductiva, que abandona la característica de ser estadística, sino más radicalmente, al acto de extirpar el carácter inductivo de las explicaciones. Si ha de preservarse una argumentación, tendrá que ser deductiva.

⁵³Dado que él piensa que pueden existir otro tipo de conexiones objetivas entre hechos involucrados en una explicación.

Notemos entonces que, en primera instancia, podemos incorporar la propuesta de Railton a por lo menos dos líneas de argumentación susceptibles de esgrimirse en contra de la idea de que podemos modelar la explicación como una argumentación. La primera de ellas, basada en los problemas sobre la suficiencia y la necesidad de la relación argumental en el modelo D-N de explicación (con los contraejemplos mencionados de Scriven). La segunda, fundada en la idea de que podemos modelar explicaciones con baja probabilidad (que nos ha traído Salmon). Sobre esta línea de argumentación podemos distinguir dos argumentaciones diferentes: la de Salmon y la de Railton. Uno puede usar, también, la existencia de explicaciones tipo I-S, como evidencia de que las explicaciones D-N no son necesarias. Y usar la existencia de las explicaciones de baja probabilidad como evidencia para mostrar que las I-S no son necesarias. La suficiencia de ambos modelos clásicos (D-N e I-S) puede refutarse mostrando que los factores de adecuación o la mención de mecanismos causales objetivos, son necesarios para identificar algo como una explicación. Construido de esta manera, una conclusión natural puede ser que la explicación no puede modelarse mediante un argumento.

1.4.3 Van Fraassen

Salmon y van Fraassen han sido dos de los filósofos que más han contribuido a la discusión sobre la explicación científica. Ambos no sólo han propuesto críticas y modelos de explicación, sino que también han elaborado textos que elucidan los problemas generales, sistematizan las tesis alternativas y conectan las consecuencias de este tema con otros problemas de la filosofía.

El modelo E-C de van Fraassen es no argumental, lo cual es de entenderse dado el contexto general de la discusión filosófica sobre el asunto. **[Van Fraassen 1980]**⁵⁴ aporta mayores argumentos que los propuestos por Salmon para la crítica contra la idea de que un modelo argumental pueda elucidar la explicación. Su argumentación está ligada a todo

⁵⁴ **[Van Fraassen 1980]** Van Fraassen, Bas C., *The Scientific Image*, Oxford University Press, 1980. La traducción al castellano es **[Van Fraassen 1996]**: *La Imagen Científica*, Trad. de Sergio Martínez, Col. Problemas Científicos y Filosóficos Paidós/IIF-UNAM No. 2, Paidós, México, 1996. Usaré la versión en castellano para las referencias.

el desarrollo del tema de la explicación, que realiza en su libro de 1980. Una explicación, para van Fraassen es algo que no puede elucidarse si no se analiza su dimensión pragmática. Una cosa sobre la que Scriven está claramente de acuerdo, según lo que hemos visto. Decir que A explica B es decir que A explica B respecto de una teoría aceptada T [Van Fraassen 1980, p. 126]. La crítica contra la idea de que la explicación se modele como un argumento se encuentra mezclada con la demostración de que la explicación es contexto-dependiente. Así, hemos de anotar una línea más de argumentación en contra de la representación argumental de la explicación, además de la que puede constituirse a partir de lo dicho por Railton: el argumento esgrimido por van Fraassen, basado en la contexto-dependencia de las explicaciones. Van Fraassen intentará demostrar esto fundamentalmente mediante tres vías: la idea de que las leyes son contexto-dependientes, la idea de que la noción de causalidad también lo es, y la propuesta de que una explicación no tiene sentido si no se ubica en cierto contexto particular. Para van Fraassen, como para Scriven, la inclusión de leyes no es un requisito para ser una explicación. Su argumentación es mucho más compleja que la que presenta Scriven, pero no la incluiré en este trabajo pues pienso que la discusión de si una explicación es o no un argumento o si puede modelarse mediante uno, es independiente de la discusión sobre si la explicación debe incluir leyes o no. Profundizaré sobre esto en el capítulo siguiente. La idea de van Fraassen de considerar que la relación de causalidad es contexto-dependiente es argumentada con base en la idea anterior, es decir, depende de la idea de que las leyes lo sean, pero además, de ejemplos prácticos de explicación en donde no se apela a leyes, por lo menos esto no se hace de manera explícita. Una cosa que vale la pena notar en este punto es que si la idea de que la relación causal es contexto-dependiente fuera verdadera, tendría consecuencias importantes en la concepción de la racionalidad científica desde la perspectiva filosófica, sobre todo considerando que la relación causal, como requisito primordial de las explicaciones, es subyacente a toda la discusión filosófica sobre la explicación. Ahondaré sobre estas cuestiones en el capítulo siguiente.

Para van Fraassen una explicación es una *respuesta* a una pregunta-porqué [Van Fraassen 1980, p. 167]; La elección de este tipo de preguntas (pregunta-porque) para referir su caracterización de la explicación no se encuentra justificada explícitamente pero, aunque sea sólo metodológicamente, es una elección sensata si se quieren discutir directamente tanto las carencias y tropiezos de la propuesta Hempel-Oppenheim como las ventajas del modelo E-C sobre dicha propuesta. La *respuesta* a este tipo de preguntas se distingue de otros tipos de cosas a los que se ha querido atribuir la noción de explicación en que contiene una dimensión intencional [Van Fraassen 1980, p. 189]. La intencionalidad de las respuestas se encuentra determinada por la ubicación de éstas en un contexto.

El modelo de van Fraassen incluye una caracterización formal. Para ello utiliza dos áreas de investigación [Van Fraassen 1980, p. 171-2]: la pragmática formal y la lógica de las preguntas⁵⁵. La primera noción en el modelo de van Fraassen, después de la idea de que la explicación es una respuesta a una pregunta, es la de *tema*. El *tema* de una pregunta es la proposición principal que la pregunta supone como verdadera. Así, en la pregunta "¿Por qué las placas fotográficas se velaron?" existe el presupuesto de que la proposición "las placas se velaron" es verdadera, ésta proposición es el tema de la pregunta.

Para entender el modelo es importante también entender la noción de "clase de contraste". La clase de contraste es una clase de hechos que podrían estar vinculados con las mismas circunstancias con las que se produjo el *explanandum*. Si P es la pregunta de la que debe dar cuenta la explicación entonces debe haber una clase X de hechos tal que la respuesta "debe aducir información que *favorezca* a P en contraste con (es decir, "en vez de a") otros miembros de X" [Van Fraassen 1980, p. 160]. Para van Fraassen la clase de contraste es un elemento muy importante de las explicaciones pues ésta determina el contexto en el que la respuesta debe ser dada. Por ejemplo, supongamos la siguiente

⁵⁵Conocida también como "lógica erotética".

pregunta ¿Por qué María salió a pasear con Juan? puede interpretarse de distintas formas. Una de ellas podría ser ¿Por qué María salió a pasear con Juan en vez de pasear con Luis o conmigo? Otra interpretación podría ser ¿Por qué María salió a pasear con Juan en vez de dedicarse a estudiar para su examen? Formalmente, tenemos dos clases de contraste asociadas a la pregunta: la clase que contiene las proposiciones "María salió a pasear con Juan", "María salió a pasear con Luis" y "María salió a pasear conmigo". Y, por otro lado tenemos la clase de contraste que contiene las proposiciones "María salió a pasear con Juan" y "María se dedicó a estudiar para su examen". Consecuentemente, la clase de contraste va a determinar el tipo de respuesta que demos a la pregunta y puede suceder que una respuesta que tome en cuenta la primera clase de contraste no será una buena respuesta en caso de haber tomado en cuenta la segunda clase de contraste.

Finalmente, un tercer concepto esencial para el modelo de van Fraassen es el de pertinencia⁵⁶. Así, podemos notar dos elementos importantes en la propuesta de van Fraassen: lo que llama "pragmática formal" y la exploración de la relación de pertinencia, esta última en el caso de van Fraassen es también un intento de incorporación de elementos pragmáticos. El modelo E-C depende crucialmente de la noción de pertinencia. Para que algo pueda ser una respuesta adecuada debe cumplirse una relación de pertinencia entre alguna proposición de la respuesta y el tema, tomando en cuenta el contexto de la pregunta. En el modelo de van Fraassen tomar en cuenta el contexto de la pregunta se traduce en tomar en cuenta la clase de contraste en relación con un contexto, una precisión de valía. Es claro que para van Fraassen la relación de pertinencia es contextual en el sentido de que depende directamente de la clase de contraste asociada con la pregunta. Al ser formulada la pregunta, desde esta perspectiva, se sugiere ya implícitamente su clase de contraste correspondiente y, con ello, la relación de pertinencia exigida a cualquier candidato a respuesta.

⁵⁶"Relevance".

Más formalmente el modelo E-C puede describirse de la siguiente manera [Van Fraassen 1980, p. 176-83]⁵⁷:

Ante una pregunta Q como "¿Por qué P_k ?" tenemos algunos componentes destacables: una clase de contraste X, un tema P_k y una relación de pertinencia R.

La pregunta puede ser representada mediante la tupla: $\langle P_k, X, R \rangle$, donde X se define por

$$X = \{P_1, \dots, P_k, \dots\}$$

Los presupuestos de Q son los siguientes:

- (a) P_k es verdadero.
- (b) cada P_j en X es falso si $j \neq k$.
- (c) Hay al menos una proposición verdadera A que mantiene una relación R con $\langle P_k, X \rangle$

Los presupuestos y la pregunta se encuentran relativizados a un contexto de conocimiento K. Para la explicación científica el contexto puede suponerse único y representaría el estado de la ciencia en un momento dado.

Una respuesta E a la pregunta $\langle P_k, X, R \rangle$ se expresa como sigue:

" P_k en contraste al resto de X porque A"

E debe cumplir con (a), (b), y con las siguientes condiciones:

⁵⁷ Para otra presentación concisa del modelo véase [Salmon & Kitcher 1987]: "Van Fraassen on Explanation", *Journal of Philosophy*, 84, 1987, pp. 315-30. Mi presentación formal está basada en la de ellos y en el texto de van Fraassen.

(r1) A es verdadera

(r2) A mantiene R con $\langle P_k, X \rangle$, (se cumple P_k porque A).

La relativización a un conjunto de conocimientos K se debe a que la interpretación de la relación de pertinencia exigida tendrá que tomar en cuenta, además de la clase de contraste, un *conocimiento de fondo* asociado al surgimiento de la pregunta. Si la pregunta "¿Por qué esta hoja de papel recubierta con cianuro de platino resplandece?" se hace de manera formal y por primera vez en, por ejemplo, el laboratorio de Wilhelm Roentgen en 1895, habrá cierto conjunto de ideas presupuestas distinto del que puede asumirse que habrá si la pregunta hubiese surgido en un laboratorio de rayos X a principios del siglo XXI. En este último caso puede que el problema supuesto por la pregunta sea simplemente "¿Quién dejó el aparato encendido?".

Como puede notarse, gran parte del peso del modelo de van Fraassen descansa sobre la relación R de pertinencia. Van Fraassen piensa que la relevancia estadística no es ni suficiente ni necesaria y para mostrar esto hará recordar algunos ejemplos clave contruidos por otros filósofos imbuidos en el tema de la explicación [Van Fraassen 1980, p. 137-9]. Van Fraassen opta por considerar la relación causal como una relación muy importante en la explicación científica. Sin embargo, el modo como van Fraassen incorpora la relación causal no obedece a una tendencia en el sentido de Coffa, sino de Scriven: es una incorporación de la causalidad como elemento contextual clave para la elucidación de cada explicación particular. Van Fraassen no está asumiendo relaciones objetivas a las que uno pueda aludir en la explicación.

Deben aclararse, entonces, dos cosas en este respecto: 1) En el modelo E-C no siempre R es una relación causal; y 2) su noción de relación causal no es "objetiva" en el sentido en que la pensaron Salmon y Railton sino contexto-dependiente, es, en palabras de Coffa, epistémica. Para van Fraassen la relación causal es completamente dependiente del contexto del conocimiento científico [Van Fraassen 1980, p. 156].

Debido a su intento de incorporación de los aspectos pragmáticos de una explicación, el modelo E-C es una buena tentativa de avance en el sentido de Scriven. Para van Fraassen la relación de pertinencia aparece completamente desligada del ideal de objetividad que buscaban Salmon y Railton. De manera ingeniosa, parte del contexto es capturado en van Fraassen mediante la tripla que contiene la clase de contraste. Por ello, el modelo de van Fraassen representa, quizá, la tentativa más decidida de debilitar las condiciones que constituyen nuestra noción de explicación, lo cual, debe aclararse, se encuentra más allá de las intenciones de Scriven.

Puede verse entonces, que el modelo de van Fraassen clarifica en qué sentido las explicaciones son contexto-dependientes. Lo son porque están ligadas a una pregunta de cierto tipo y esa clase de preguntas no pueden entenderse si no se ubican en cierto contexto particular. No se puede responder adecuadamente a una pregunta por-qué si no se sabe en qué contexto se ubica la pregunta y este contexto puede caracterizarse mediante lo que él llama “clase de contraste”. Una respuesta adecuada, es decir, una explicación adecuada, es aquella que cumple con alguna relación de pertinencia respecto de la asunción de la pregunta y la clase de contraste en un contexto dado. Y, naturalmente, la relación de pertinencia es entonces también contexto-dependiente. Así, lo que puede dar cuenta de esta clase de fenómeno en la Ciencia, las explicaciones, no es un análisis de los argumentos, sino un análisis de las preguntas y la adecuación de sus respuestas. De este modo, además de la idea de que un argumento no puede modelar una relación de baja probabilidad, que él toma de Salmon (aunque también aparece en Railton), la manera en que van Fraassen contribuye a esta idea de un modelo no argumental es a través de una argumentación positiva a favor de las preguntas como elementos más inmediatamente pragmáticos desde los cuales podemos caracterizar las explicaciones.

En este capítulo hemos revisado los problemas a los que se enfrentaron los modelos clásicos de la propuesta Hempel-Oppenheim. En esta revisión, analizamos dos de las primeras críticas, a saber, la de Coffa y la de Scriven, y propuse que estas críticas pueden servir de postes indicadores, vistas como dos tensiones respecto de la caracterización de una racionalidad explicativa: una que puede considerarse la de una racionalidad, digamos, “dura”, que tiende a una eliminación de los elementos epistémicos y pragmáticos a favor de rasgos objetivos de las relaciones subyacentes en las explicaciones; otra que tendería a incluir esta misma clase de elementos mientras que prescindiría de los rasgos objetivos buscados por los primeros. Esto puede verse como parte de una tensión, más general, entre concebir la explicación en términos epistémicos, dependiente de contextos o de algo como marcos conceptuales, como una propuesta teórica, y, en cambio, concebirla como algo dependiente únicamente de rasgos bien determinados en el mundo. En este marco de discusión, analizamos tres modelos alternativos: el de Salmon, el de Railton y el de van Fraassen. Estos modelos, además de que dos de ellos son de los modelos con mayor repercusión filosófica (el de Salmon y el de van Fraassen), resultan interesantes sobre todo para el problema general de si la explicación puede modelarse como un argumento, en el contexto de la tensión general mencionada en las líneas anteriores.

En el siguiente capítulo, profundizaremos sobre un conjunto de problemas relacionados con este problema general, el problema de una modelación inferencial de la explicación. Un análisis de estos problemas nos pondrá en una mejor posición para precisar la estructura de un modelo formal que intente responder, en este terreno, a las exigencias prescritas por la Filosofía de la Ciencia al tema de explicación científica.

CAPÍTULO 2

LOS PROBLEMAS CLÁSICOS PARA LA MODELACIÓN ARGUMENTAL DE LA EXPLICACIÓN: INFERENCIA Y PERTINENCIA EXPLICATIVA

Introducción

El desarrollo de nuevos modelos de la explicación no es gratuito. Las tensiones Hempel-Scriven y Hempel-Coffa muestran básicamente el choque de dos tipos de intuiciones acerca de la explicación: una que intenta involucrar elementos pragmáticos y otra que intenta mantener cierta idea de objetividad independiente del contexto. A lo largo de la discusión sobre el tema fueron identificándose algunos problemas más precisos. Los problemas se encuentran tan intrincadamente relacionados que es difícil hacer una exposición sistemática de ellos sin dejar la impresión de que no se hicieron algunas conexiones entre los mismos, que pueden ser importantes. Ordenarlos implicará algún criterio de selección. El criterio en el que estoy interesado es el de su relación con la posibilidad y pertinencia de una representación inferencial de la explicación científica.

En este capítulo examinaré algunos de los problemas clásicos mencionados en el capítulo anterior. Comenzaré con dos grupos de problemas clásicos planteados en la crítica contra los modelos tradicionales (D-N e I-S). El primer grupo se constituye de problemas contra la relación de inferencia de los modelos, el segundo, de problemas contra la idea de que los modelos capturen correctamente la relación de relevancia en la explicación y, en este sentido, pueden ser importantes a la hora de determinar las condiciones de inferencia de las estructuras propuestas por los modelos de explicación. Luego describiré muy brevemente otro par de problemas relacionados con los anteriores

que, aunque menos típicos, son muy importantes sobre todo en relación con la deducción: el problema de la sobredeterminación epistémica causal y el problema de la elección causal, sección 2.1.5. En el examen de ambos grupos de problemas procuraré plantear con precisión las cuestiones más importantes y argumentaré a favor de una solución tentativa para ellas. Al mismo tiempo trataré de ir dejando clara la idea de explicación que ayudaría a construir un modelo adecuado de explicación.

Como dejé claro en la introducción, en todo momento el tipo de explicación sobre la que desarrollaré la discusión será la explicación que se exige mediante una pregunta-porqué y cuyo *explanandum* es un evento singular.

2.1 Desarticulando Dos Problemas Clásicos

2.1.1 El Supuesto Argumental

2.1.1.1 El problema

Entiendo "**Supuesto Argumental**" (SA) como el supuesto de que la explicación es susceptible de ser adecuadamente modelada mediante argumentos. SA podría acompañarse de otros dos supuestos:

SA1) la explicación x es susceptible de modelarse mediante argumentos pues la relación explicativa que constituye a x tiene similitudes con una, o es una clase de, relación inferencial.

SA2) la modelación de una supuesta explicación x , mediante un argumento, no necesariamente nos compromete con tesis sobre la existencia de estructuras

inferenciales de x pero sirve como criterio para decidir si x es realmente una explicación o bien para decidir si x no lo es.

De los modelos expuestos en la sección precedente, sólo D-N e I-S guardan una clase de SA. Hempel parece que mantuvo más bien SA2 que SA1. Sus aseveraciones en **[Hempel & Oppenheim 1948]** y **[Hempel 1965]** acerca de la explicación y acerca de los esquemas que la modelan no indican ningún supuesto que tenga que ver con SA1. Scriven y Salmon son críticos directamente de SA2. Ellos critican que sea suficiente o que sea necesario, para decidir ya sea que algo es o que no es una explicación, verificar una relación argumentativa (y sus condiciones de adecuación) entre el *explanans* y el *explanandum*. Por su parte, las críticas de Railton y van Fraassen se encuentran dirigidas, de forma más vaga, contra SA en general. Es plausible que una refutación de SA2 refute también SA1. Por estas razones, mi análisis de SA supondrá que el punto de discusión es SA2 y no SA1. La estructura general de la crítica contra SA se efectúa sobre dos flancos: contra la derivación deductiva y contra la derivación probabilística.

Scriven combate la derivación deductiva como criterio para decidir si algo es una explicación, es decir, Scriven critica SA2. Los contraejemplos presentados para refutar la suficiencia de la derivación deductiva reúnen dos características: $\neg B$) que el *explanans* confiere apoyo deductivo al *explanandum* y, $\neg A$), que sin embargo esos contraejemplos no serían considerados intuitivamente como explicaciones. De modo simplificado, la reconstrucción del contraejemplo de Scriven⁵⁸ puede esquematizarse como sigue:

(Puente)

Todo puente dinamitado en una de sus arcadas se destruye parcialmente

El puente G fue dinamitado en una de sus arcadas

El puente G se destruyó parcialmente.

⁵⁸Véase, *supra*, capítulo anterior, sección 1.2, *La Tensión Hempel-Scriven*.

Sin embargo, añade Scriven, puede suceder que la destrucción parcial de G haya sido provocada no por su dinamitación sino por una fractura en una de sus arcadas, a pesar de haber sido efectivamente dinamitado en una de sus arcadas.

Los ejemplos de Scriven contra la necesidad de la derivación deductiva, que presenté antes⁵⁹, el caso del simbolismo en una danza tribal y el caso de la explicación de una teoría o de una prueba, no son adecuados contraejemplos al modelo D-N de explicación. Pues ambos ejemplos son en realidad acerca de explicaciones-qué y no de explicaciones-porqué⁶⁰. Para el caso de la refutación de la necesidad de la derivación deductiva, se reúnen dos características para los contraejemplos: A) que son considerados intuitivamente como explicaciones y, B), que el *explanans* no confiere apoyo deductivo al *explanandum*. En realidad Hempel no propone la derivación deductiva como único caso de explicación-porqué, pues propone el modelo I-S. Observemos que cualquier ejemplo de explicación probabilística I-S que no tenga probabilidad $r = 1$ es un contraejemplo a la necesidad del requerimiento de derivación deductiva. Por ello, la crítica contra la necesidad de una derivación deductiva no debe entenderse como una crítica contra la teoría de la explicación defendida por Hempel⁶¹. Consecuentemente, el énfasis de la crítica contra Hempel fue puesto sobre la suficiencia de ese tipo de derivación, la deductiva. De cualquier forma, un ejemplo contra la necesidad de la derivación deductiva, y por tanto contra SA2, puede ser el siguiente argumento I-S:

(Rec)

Px: a x se le aplica penicilina.

Sx: x tiene una fuerte gripa.

Rx: x se recupera de la gripa.

j: Juan.

⁵⁹Los que apoyan el inciso (2) en el apartado 1.2, *La Tensión Hempel-Scriven*.

⁶⁰Además de que, en retrospectiva, trivialmente no son contraejemplos adecuados, por el hecho de que Hempel propone en realidad dos alternativas de modelación: Los modelos D-N e I-S.

⁶¹ Puesto que la Teoría de Hempel no es sólo la parte donde se refiere al esquema D-N.

$$\begin{array}{l}
 P(Rx / Sx \wedge Px) = .8 \\
 Sj \wedge Pj \\
 \hline
 Rj
 \end{array}
 \quad [.8]$$

El segundo flanco de la crítica es contra la derivación probabilística del modelo I-S. Los contraejemplos tienen características similares a los esgrimidos contra la derivación deductiva. Contra el caso de la suficiencia de la derivación probabilística se construyen contraejemplos con dos características: B') que el *explanans* confiere apoyo derivacional probabilístico al *explanandum* y, -A'), que no son considerados intuitivamente como explicaciones. Salmon es uno de los principales críticos de SA sobre el flanco de la derivación probabilista. [Salmon 1965]⁶² opone, en relación a *Rec*, el siguiente ejemplo⁶³:

(Vitamina)

- Cx : x toma fuertes dosis de vitamina C.
- Sx : x tiene gripa.
- Rx : x se encuentra recuperado al cabo de un par de semanas.
- j : Juan

$$\begin{array}{l}
 P(Rx / Sx \wedge Cx) = .8 \\
 Sj \wedge Cj \\
 \hline
 Rj
 \end{array}
 \quad [.8]$$

La probabilidad frecuencial de *Rj*, dado *Sj* y *Cj*, es alta. Pero, en realidad, como lo señala Salmon, la mayoría de los casos de gripa son casos de recuperación en un par de

⁶²"The Status of Prior Probabilities in Statistical Explanation", *Philosophy of Science* 32, 1965, pp. 137-146.

⁶³El ejemplo originalmente se propuso en [Salmon 1965]. La que presento enseguida es mi reconstrucción formal del ejemplo mencionado en [Salmon 1984b].

semanas. Puede entonces considerarse dudoso que el factor importante para la explicación de la recuperación de Juan sea la vitamina C.

Para el caso de la necesidad de la derivación probabilística, las características de los contraejemplos son: A') que son considerados intuitivamente como explicaciones y, -B'), que el *explanans* no confiere apoyo derivacional probabilístico al *explanandum*. El ejemplo clásico de explicación que cumple con ambas características es el de la paresia⁶⁴.

(Paresia)

La paresia sólo se desarrolla en personas que han padecido sífilis pero sólo un bajo porcentaje de los que han padecido sífilis, la desarrollan. Supongamos que José tiene paresia. Una explicación para el hecho de que José tiene paresia es apelar a que él tuvo sífilis.

Sin embargo el *explanans* que dicta que José tuvo sífilis y que la probabilidad de que alguien tenga paresia dado que tuvo sífilis es muy baja, no puede ponerse en términos de una derivación probabilística al estilo de Hempel, y plausiblemente a ningún estilo común de inferencia probabilística. Esto se debe a que, si se trata de una probabilidad baja, esto no cumpliría con el requerimiento de alta probabilidad en la derivación y en general se asume algún índice de alta probabilidad para tales inferencias. Así, aunque normalmente se supone una relación causal entre el hecho de haber tenido sífilis y el de padecer de paresia y en este sentido en general se supone una relación explicativa entre ambos hechos, no puede esta relación expresarse en términos de una relación argumental probabilística.

Estos dos ejemplos constituyen una de las líneas de argumentación que Salmon esgrime en contra de SA2. Otra línea de argumentación de Salmon es relacionar el problema de si debemos mantener SA2, dado el problema planteado por Carnap en

⁶⁴Se menciona en [Salmon 1984b]. También en [Salmon 1984c]: Salmon, W. C., *Four decades of Scientific explanation*, University of Pittsburgh, 1984; p. 59.

[Carnap 1950]⁶⁵ acerca de si deberíamos tener reglas de aceptación para la lógica inductiva. En [Salmon 1971] propone que el problema puede ilustrarse muy bien con la paradoja de la lotería, que consiste en lo siguiente: dado que ninguno de los boletos que tienen la posibilidad de ganarla son altamente probables, entonces cualquier caso es rechazable. Sin embargo, sabemos que alguno de los boletos se cumplirá, así, ninguno se cumple pero alguno se cumple. El problema que Salmon enfatiza consiste en que la inferencia de la conclusión de un argumento I-S es dependiente de un determinado límite de alta probabilidad, es decir, el rechazo o la aceptación de la conclusión es dependiente de una regla que establezca este límite. Por ello, como siempre hay un determinado número de boletos mediante el cual la probabilidad de cada uno de los mismos esté por debajo del límite preestablecido de alta probabilidad, siempre habrá forma de calificar a todos los boletos como casos rechazables: aumentando el número de boletos. En el caso de una explicación sucede algo similar. La nueva evidencia puede provocar que la probabilidad de la inferencia disminuya tanto que quede por debajo del límite preestablecido. De este modo, el *explanandum* tendría que ser rechazado pues no cumpliría con la regla. La idea básica de Salmon es como sigue. Carnap dice que no puede haber reglas de aceptación en lógica inductiva, dada la paradoja. Un argumento probabilístico supondría una regla de aceptación, pero como hemos visto no es posible una regla tal, entonces no es posible un argumento de esta clase. Para Salmon el error consiste en poner el efecto de los datos probabilísticos en términos de aceptación o rechazo, esto es, en términos de un esquema inferencial del cual dependa la aceptación o rechazo. Lo mismo pasaría con la explicación: se trata de la aceptación o rechazo del *explanandum*. Estas son las otras razones por las que justifica la elección de un modelo no argumental de explicación.

Railton complica más el problema. Como vimos en el capítulo anterior, en [Railton 1978] se construye el modelo D-N-P, que no sólo no es argumental sino que se pretende que sea capaz de subsumir explicaciones de *explananda* con baja probabilidad y azarosos

⁶⁵Carnap, Rudolf, *Logical Foundations of Probability*, University of Chicago Press, Chicago, 1950.

La dificultad para SA, contenida en la idea de que algunos de los *explananda* son azarosos es que, podamos aducir *explanans* que le confieran baja o alta probabilidad, no tenemos apoyo suficiente para determinar que sucederán. Esto implica que en ocasiones no podemos tener un argumento para la ocurrencia del *explanandum* aún a pesar de que el *explanandum* efectivamente haya sucedido. Railton no niega la posibilidad de que haya explicaciones modelables mediante SA, lo que asegura es que hay casos en los que esto no es posible. No obstante, Railton hace una defensa de la forma argumental deductiva de las explicaciones en la práctica normal de los científicos. Así, para Railton el modelo D-N-P intenta recuperar esa intuición argumental en una estructura que no es un argumento, sino una forma de *dar cuenta* del fenómeno.

Railton intenta realizar dicha recuperación a través del *addendum*, que nos asegura la ocurrencia del *explanandum*. En el caso del núcleo de uranio, presentado en la sección 1.4.2, una vez que sabemos que el *explanandum* realmente tuvo lugar entonces el argumento que nos lleva a la aceptación de su baja probabilidad cobra sentido como parte de una explicación del hecho azaroso. Esta es la manera en que funciona una forma de dar cuenta según la opinión de Railton.

Para un examen de la crítica contra SA es suficiente con que consideremos lo que he expuesto al respecto hasta ahora. Los detalles que confirmarían que van Fraassen es también un crítico de SA no los expondré en esta investigación. La intervención de van Fraassen en el debate sobre la explicación es mucho más importante en relación con el problema de la pertinencia, que abordaré después, en este mismo capítulo. Las críticas de Scriven, Salmon y Railton nos servirán de punto de partida para un análisis del problema general de la posibilidad de una representación inferencial de la explicación científica.

Tenemos entonces varias dificultades en contra de SA:

- 1) La derivación deductiva no es ni necesaria ni suficiente.
- 2) La derivación probabilística no es ni necesaria ni suficiente.
- 3) Entender la probabilidad en términos de un argumento es inadecuado.
- 4) Modelar bajo SA es inadecuado por lo menos para ciertos *explananda*, a saber, los eventos que tienen baja probabilidad de ocurrir y que son azarosos.

Las dificultades podrían concatenarse, si uno quisiera reconstruirlo de este modo, en el siguiente argumento:

*Cualquier derivación propuesta para sostener SA no es necesaria ni suficiente para una explicación. Además, la idea de SA asociada a datos probabilísticos supone entender la probabilidad en términos de aceptabilidad, lo cual es inadecuado. Finalmente, la condición de **alta expectabilidad** del explanandum (afirmación 4 anterior) que parece subyacer a toda propuesta de SA hace que algunos hechos no puedan entenderse como casos de explananda cuando aparentemente lo son. Debido a todos estos resultados, SA es inadmisibile.*

2.1.1.2. El Análisis del Problema o El Supuesto de Alto Grado de Certeza

Debo decir que las afirmaciones 1 y 2 me parecen suficientemente fundamentadas. De cualquier modo, un acercamiento al contenido de la afirmación 4 nos hará comprender mejor el alcance de ellas y el error de las afirmaciones 3 y 4.

La **alta expectabilidad** en la explicación puede formularse como sigue: una explicación debe dar razones para que el *explanandum* parezca un fenómeno muy esperable. La conexión de SA con este supuesto de alta expectabilidad se traduce en el supuesto de que toda relación de derivabilidad debe estar fundamentada en un **alto grado de probabilidad**; es decir, en el supuesto de que en un argumento las premisas deben conferir un alto grado de probabilidad a su conclusión.

Si bien es verdad que el supuesto de alta expectabilidad en la explicación resultará adecuado para muchos *explananda* interesantes, no es necesario que SA incluya condiciones que capturen un supuesto como éste. La línea de pensamiento que conecta SA, como se postuló tradicionalmente, con el supuesto de alto grado de certeza en la inferencia podría reconstruirse como haré en los siguientes párrafos.

Para Hempel y Oppenheim la idea de modelar la explicación estuvo asociada con algunas ideas restrictivas. Una de ellas es que las explicaciones científicas deben ofrecerse en el marco de ciertas formulaciones generales, a saber, las leyes científicas. Otra, que la relación de pertinencia entre dos clases de fenómenos debe ofrecer, como una de sus características principales, la posibilidad de predecir fenómenos, de aquí la importancia de la simetría entre explicaciones y predicciones.

Estas ideas configuran, juntas, una noción de explicación rígida. La base para la capacidad predictiva de las leyes es, moderada y escuetamente, su capacidad de proveer una alta expectabilidad para un fenómeno dada la aparición de otro. Así, intentando seguir una línea uniforme de pensamiento, es natural que Hempel y Oppenheim intentaran que la capacidad predictiva de casos particulares a partir de leyes se reflejara también en una explicación de este tipo de casos, pues las explicaciones suponen leyes y esto hace que podamos establecer una simetría entre explicaciones y predicciones. De acuerdo a una perspectiva en la que la explicación y la predicción tienen el mismo estatus, como funciones principales de la Ciencia, la explicación debería ofrecer entonces también una alta expectabilidad para el *explanandum*. Así, un fenómeno resultaría explicado sólo mediante otros fenómenos que aseguraran esta alta expectabilidad.

Obsérvese también que, puesta en estas condiciones la relación entre *explanans* y *explanandum*, su versión en términos aléticos significaría que dada la verdad del *explanans*, debería esperarse *eficazmente* la verdad del *explanandum*. Desde esta

reconstrucción racional de los motivos para el diseño de los modelos clásicos, cobra un sentido muy intuitivo que una explicación pueda expresarse en términos de una relación inferencial. La relación entre el *explanans* y el *explanandum* es una relación expresable en términos de un argumento. El primer proyecto de modelación basado en SA se encontraría, en esta reconstrucción, asociado con el supuesto de alta expectabilidad. Para capturar este supuesto, la modelación de la relación de pertinencia en una explicación puede entenderse adecuadamente como una relación inferencial entre el *explanans* y el *explanandum*⁶⁶.

Un supuesto de esta forma de arribar a SA para la modelación de la explicación es la tesis de la simetría. Una consecuencia, la suposición de una fuerte dependencia del proyecto de modelar la explicación como un argumento, sobre la idea de que se logre una formulación precisa del concepto de ley científica. Esta última suposición llega a ser una de las preocupaciones centrales de la modelación de la explicación desde el punto de vista de los modelos clásicos.

Ahora bien, sea o no plausible una reconstrucción como la anterior, para decidir lo cual convendría exponer una investigación histórica más rigurosa, la estrategia de Hempel y Oppenheim para modelar la explicación presupuso originalmente una clase de inferencia clásica. Desde esta perspectiva, una relación argumental supone que el conjunto de premisas debe ofrecer fundamentos para la creencia en la conclusión del argumento. Sin embargo, si consideramos sólo el tipo de inferencia clásica entonces no podemos entender dicha fundamentación si no es bajo el supuesto de que del conjunto de premisas verdaderas se deriva irrestrictamente la verdad de la conclusión, el esquema debe ser capaz de transferir, con toda seguridad, una probabilidad absoluta, o mejor dicho, el esquema debe ser válido. Sin embargo, dado el desarrollo de la Lógica contemporánea, el modelo D-N representa, entre otras cosas, una forma muy específica de capturar el supuesto de alta expectabilidad.

⁶⁶Como se ha visto, la relación inferencial en los modelos de Hempel tiene otras condiciones que he dejado de lado en este análisis, pero que son independientes de él.

Es interesante que consideraciones similares a las que he descrito en los párrafos anteriores pueden aplicarse también a los argumentos I-S. Aunque la probabilidad absoluta no aparece en los argumentos I-S, la aspiración a modelar la alta expectabilidad sigue presente. En términos generales, en un argumento explicativo (D-N o I-S) la relación de pertinencia entre el *explanans* y el *explanandum* se refleja en el hecho de que el segundo se deriva del primero por lo menos con un alto grado de probabilidad. La forma clásica de entender una relación inferencial, la cual es coherente con este ideal de expectabilidad, es lo que probablemente llevó a Hempel a construir un modelo argumental inductivo-probabilístico (I-S) que se encuentra basado en un alto grado de probabilidad para el *explanandum*. El grado de sorpresa de Hempel acerca de las diferencias entre el modelo D-N y el modelo I-S, descritas en el capítulo anterior, parecería así directamente proporcional al grado en que Hempel mantiene el supuesto de alto grado de probabilidad idealizado por el modelo D-N; y los ajustes que Hempel intentó hacer al modelo I-S (el requisito de máxima especificidad) estarían motivados, de igual forma, por este mismo supuesto.

La derivación de un argumento I-S está fundamentada en una alta probabilidad del *explanandum*. Sin embargo, si hemos descubierto el fondo de la perspectiva de los modelos clásicos, los contraejemplos a favor de la baja probabilidad revelarían que la relación de pertinencia modelada entre el *explanans* y el *explanandum* no debería estar restringida a una alta expectabilidad. Bajo SA2, el supuesto de alto grado de probabilidad en los argumentos modela, por el contrario, la alta expectabilidad. La derivación deductiva y la derivación por alta probabilidad supondrían así, equivocadamente, una alta expectabilidad. Se nota así la pertinencia de las aseveraciones 1 y 2 anteriores: la deducción y la probabilidad no son relaciones ni necesarias ni suficientes en la modelación de la explicación. De esta forma, D-N e I-S resultan inadecuados.

Una consecuencia de mirar de este modo a los modelos clásicos es, también, que el punto crucial de la crítica contra SA que sostienen Scriven, Salmon y Railton puede verse como basado en una crítica contra el supuesto de alta expectabilidad. La crítica señala atinadamente un presupuesto que no debería efectuarse en el caso de las explicaciones. Por tanto, a la luz de estos autores, SA2 no puede ser una forma adecuada de clarificar la explicación científica, pues los argumentos no podrían, en ambas versiones, probabilística o deductiva, representar algunos casos legítimos de explicación.

El argumento parece exitoso. Observemos sin embargo, que en el argumento del párrafo anterior, hay una salvedad. Se refiere, como hemos dejado claro antes, al tratamiento clásico de la relación inferencial: deducción y alta probabilidad. No todo tipo de derivación ha sido contemplado en los modelos defendidos por Hempel. Así, la conexión del supuesto de alto grado de probabilidad con SA podría no ser necesaria. Si lográramos construir un modelo de explicación como argumento cuya relación inferencial no estuviera fundamentada en el supuesto de alto grado de probabilidad, la sucesión asumida entre una crítica contra el supuesto de alta expectabilidad y una crítica contra el proyecto de SA para modelar la explicación, se vería socavada.

Las críticas contra la necesidad y contra la suficiencia de la inferencia deductiva están basadas en, por un lado, la inferencia probable y, por otro, la falta de pertinencia del *explanans* respecto del *explanandum*. Como he descrito, críticas análogas se efectuaron contra la inferencia probable, basadas en la baja probabilidad (para el caso de la necesidad) y la falta de pertinencia (para el caso de la suficiencia). Estas dos críticas análogas constituyen la doble estrategia de ataque contra SA, a veces en la forma de una crítica más concreta contra SA2. La crítica contra SA (construida por Salmon y Railton entre otros) parece suponer que toda clase de relación inferencial debe estar basada en el supuesto de alto grado de probabilidad. Por lo tanto, ambas vías de ataque (1 y 2) tienen sólo cierto alcance: la derivación deductiva y la probabilística. Atacar estas dos clases de derivación no debe verse como una refutación del supuesto argumental (SA), si es que

podemos construir y defender una noción de derivación que no implique un alto grado de probabilidad.

En efecto, podemos construir argumentos cuya relación inferencial no esté basada en dicho supuesto de alto grado de probabilidad. Una forma de hacerlo podría ser utilizando un tipo de lógicas contemporáneas que están usándose de manera importante para el diseño de inteligencia artificial: las lógicas no monotónicas⁶⁷, lógicas que tienen la propiedad de retractar inferencias que describí en la introducción.

No sólo Salmon y Railton se vieron envueltos en esta forma de razonamiento que los lleva de las críticas 1 y 2 a una crítica contra SA. Van Fraassen interpretó en [Van Fraassen 1980]⁶⁸ el problema de sostener SA básicamente como el problema de pretender fundar la explicación sobre la idea de tener *bases firmes para creer*⁶⁹. En este texto se señala que la identificación de una explicación puede hacerse mediante el criterio de verificar si ofrece bases firmes para creer en el *explanandum*. Dicha verificación daría origen, según el autor, a la idea de modelar la explicación como un argumento (o como he apuntado, a SA), pues si el candidato a explicación cumple con ciertos esquemas argumentales ofrece las bases firmes en cuestión. Esta es una de las ventajas interesantes de un modelo argumental de explicación científica, que captura esta intuición básica de una explicación. Pero el hecho de ver refutada la suficiencia y la necesidad de esta suerte de criterio de "ajuste argumental" para los casos de la derivación deductiva y la probabilística, sirvió a van Fraassen para justificar su abandono de SA. No obstante,

⁶⁷Especialmente [Reiter 1980]: Reiter, Raymond, "A Logic for Default Reasoning", en *Artificial Intelligence*, Vol. 13, 1980, pp. 81-123. Reimpreso en Matthew L. Ginsberg, (comp.), *Nonmonotonic Reasoning*, Morgan Kaufmann Publishers, E.U.A. 1987. Una breve descripción general de las lógicas no monotónicas puede consultarse en [Minker 1991]: Minker, Jack, "An Overview of Nonmonotonic Reasoning and Logic Programming, *First International Workshop on Logic programming and Nonmonotonic Reasoning*, Washington, D. C., Julio 23, 1991. Como he mencionado, por ejemplo en la introducción de esta tesis, mi artículo [Gaytán 2005] y mi tesis de maestría describe una forma de aplicación de la inferencia *default* de Reiter, en el caso del problema de las salvedades.

⁶⁸Van Fraassen, Bas C., *The Scientific Image*, Oxford University Press, 1980. La traducción al castellano es *La Imagen Científica*, Trad. de Sergio Martínez, Col. Problemas Científicos y Filosóficos Paidós/IIF-UNAM No. 2, Paidós, México, 1996, p. 133. Usaré la versión en castellano para las referencias.

⁶⁹Influido por algunas reflexiones posteriores de Hempel en su *Philosophy of Natural Science*.

siguiendo mi línea de argumentación, también las bases firmes para creer parecen entenderse, por lo menos cuando se ponen en términos inferenciales, bajo el supuesto de alto grado de probabilidad. Notemos por ahora que una explicación tal vez no necesite proporcionar *bases firmes*, sino tan sólo *bases para creer*. Si esto es admisible, la refutación de SA por la vía contenida en [Van Fraassen 1980] parece tener el mismo destino que las anteriores: no es un golpe contundente contra SA sino contra cierta versión particular de SA. Bajo la consideración de otro tipo de argumentos cuyo ofrecimiento de *bases para creer* en la conclusión no sea dependiente de dicho supuesto, la no necesidad del abandono de SA se hace, por lo menos, posible. Con ello, la representación inferencial de la explicación también aparece como posible.

En las líneas precedentes he argumentado que no es necesario abandonar SA para modelar la explicación, esto es, que la explicación puede adecuadamente ser modelada mediante argumentos. Si hemos de tener éxito en esta línea de pensamiento, hemos de sostener i) que hay alguna relación inferencial más débil que puede adecuadamente modelar la explicación y ii) que esta relación puede entenderse como una representación argumental de la explicación, es decir, como un caso de SA. Como expliqué en la introducción, en mi tesis de maestría mostré que hay sistemas lógicos que contienen una relación inferencial que es capaz de dar cuenta del problema de las salvedades de Hempel. Dije además que esta clase de relación inferencial no estaba comprometida con una alta probabilidad en la inferencia. Así, para mostrar i, es preciso que además de la propuesta de una relación inferencial debilitada, mostremos la adecuación de esta clase de relación inferencial para la modelación de la explicación científica. Como anticipé también en la introducción, haría falta una adaptación de este modelo inferencial, que dé cuenta de los problemas suscitados en el contexto de la discusión sobre el tema en la Filosofía de la Ciencia. En el siguiente apartado sobre el problema de la ambigüedad epistémica, desarrollaré la idea del fondo filosófico necesario para construir una propuesta que aborde los problemas que desde la Filosofía de la Ciencia habría que tomar en cuenta en una modelación de la explicación científica. Después, en el subsiguiente

apartado, trataré de fundamentar ii, la idea de que esta clase de inferencia, una de baja probabilidad con retractación de inferencias (en términos muy generales es decir, con no monotonicidad) puede verse como una forma de representación argumental.

2.1.2 La Ambigüedad Epistémica

2.1.2.1 La Reacción Estándar

El problema de la ambigüedad epistémica, tal como lo caracterizó Coffa, consiste en que las explicaciones de tipo I-S producen inconsistencias inductivas. Como dije antes⁷⁰, ante el resultado de inconsistencias inductivas, Hempel recurrió a dos elementos: el requerimiento de máxima especificidad (RME) y la relativización epistémica.

Para Coffa, el problema consiste en que la definición de explicación inductiva de Hempel, a saber, I-S, no es la adecuada, pues conduce a inconsistencias inductivas. Una definición adecuada no tendría que tener esta consecuencia. Según Coffa, el planteamiento erróneo contenido en I-S puede estarse produciendo por la interpretación estadística de las relaciones de probabilidad. Una interpretación más adecuada podría ser la de propensiones. Además para Coffa no es necesario que una restricción del tipo de RME tenga que exigir la relativización epistémica. En contraste con Coffa, Railton opinó que el problema de la ambigüedad epistémica se produce cuando intentamos hacer encajar la explicación en un esquema inductivo. No es un problema de la definición de explicación inductiva, sino de la interpretación de una explicación estadística en términos de un esquema de razonamiento inductivo.

La distinción entre Railton y Coffa puede entenderse mejor si partimos de otra distinción: la distinción entre inducción y probabilidad estadística. Hempel dejó clara esta distinción cuando presentó su modelo I-S junto con el modelo D-S. Para Hempel la

⁷⁰Véase, *supra*, *El Modelo I-S y las Diferencias entre ambos Modelos*.

relación de probabilidad estadística se refiere a relaciones entre los eventos del mundo y la relación de inducción es una relación de derivabilidad que se efectúa entre proposiciones. En una explicación la propiedad de ser probabilística es independiente de la de ser inductiva. Sólo que la relación de inducción es expresada por Hempel en términos probabilísticos. Así, mientras que Coffa se encuentra preservando la explicación inductiva y deshaciéndose de la propiedad de probabilidad estadística, Railton se encuentra preservando la explicación probabilística y deshaciéndose del esquema inductivo. Coffa no cree que existan explicaciones estadísticas, tanto como Railton no cree que existan explicaciones inductivas. Ambos, debe aclararse, oponen la interpretación en propensiones a la interpretación estadística de la probabilidad. De modo que ambos conciben la posibilidad de explicaciones probabilísticas pero interpretadas mediante propensiones. Es claro que para ambos autores, las dos características indeseables de las explicaciones I-S son RME y la relativización epistémica.

La relativización epistémica lleva a Coffa a negar que los argumentos de tipo I-S sean explicaciones. RME también es inadmisibles para Coffa:

En efecto, entonces, es la ignorancia, más que el conocimiento, lo que hace que el principio de máxima especificidad parezca una demanda viable. [...] un incremento en el conocimiento podría dejarnos sin ninguna explicación en absoluto.⁷¹

Esto lo lleva a pensar que RME es, a la larga, inútil. A partir de ello Coffa se pronuncia por el intento de construir una idea de explicación probabilística que tenga bases objetivas más firmes, por ejemplo, basada en propensiones.

Una forma de entender la tensión Hempel-Coffa en un sentido más amplio, es tratando de determinar el rumbo de lo que propone, como si fuese una tendencia general.

⁷¹"In effect, then, it is ignorance, rather than knowledge, that makes the maximal specificity principle look like workable demand. [...] an increase in knowledge could leave us with no explanation at all.", Coffa, Alberto J., "Hempel's Ambiguity" en Ruben, David-Hillel (ed.), *Explanation*, Oxford University Press, E.U.A, 1993, p. 70. Fuente original: "Hempel's Ambiguity", en *Synthese*, 28, 1974, pp. 141-63.

La tensión Hempel-Coffa se mueve, dije antes, hacia un tipo de racionalidad "dura". Una teoría de la explicación heredera de la tensión invita a concebir la explicación como una entidad infalible. Algo sobre lo que no se puede dar marcha atrás.

La cita de Coffa que hemos presentado compromete la noción de explicación con una especie de **ideal de buena explicación**. Gran parte de las dificultades identificadas para la modelación de la explicación se deben a una concepción de este tipo: una concepción que considera a la explicación como algo infalible. Concebir el ideal de que algo es una explicación si y sólo si es una buena explicación es un recurso pragmático que puede ayudarnos a esclarecer tanto los problemas en contra de SA como el problema de la ambigüedad epistémica. Este ideal puede ser entendido por lo menos en dos sentidos: dicta que algo es una explicación si y sólo si es *una* buena explicación, y, en el caso más extremo, dicta que algo es una explicación si y sólo si es *la* explicación. La posición de Coffa parece buscar la segunda versión de este ideal y, por tanto, mantiene también la primera versión.

Coffa, al comprometer su teoría de la explicación con el ideal de buena explicación, cancela considerar explicaciones a los argumentos I-S. Sin embargo, este compromiso también le impide apreciar algo que podríamos llamar "las virtudes de la ignorancia". Esto puede ejemplificarse en la cita anterior. Para Coffa una explicación I-S no tiene nada que decirnos acerca del mundo. Esto significa que no estamos seguros de que lo que nos dice es verdadero ni si corresponde con las relaciones reales entre los eventos involucrados. La posibilidad de un cambio de parecer ante nueva evidencia en el *explanans* hace que Coffa niegue la calidad de ser explicación a los argumentos I-S. Esto es producto de su idea de que una definición de explicación tiene que capturar lo que se entiende por buena explicación.

El caso de Railton es parecido. Las razones por las que Railton desacredita los esquemas inductivos e intenta extirparlos de toda noción de explicación obedece también

al ideal de buena explicación. A diferencia de Coffa, Railton piensa que la inducción es la responsable directa de las consecuencias indeseables del modelo de Hempel: RME y la relativización epistémica. Para Railton no hay posibilidad de construir una explicación en términos de inducción. Railton enfatiza, sobre todo, la importancia de que el *explanans* proporcione relaciones (mecanismos) que *efectivamente* se den entre los fenómenos. Tras esta subordinación de la idea de explicación a la mención de relaciones que correspondan realmente a las relaciones entre los hechos involucrados, se encuentra el ideal de buena explicación. Railton piensa que el nivel epistémico en que se encuentra concebida una explicación I-S es el problema. Así, necesitamos ubicar la explicación en el nivel ontológico. Pero, además, Railton supone un esquema deductivo como un razonamiento que además de ser deductivo parte de premisas verdaderas. Esta puede ser la razón por la que Railton prefiere entonces mantener los esquemas deductivos dentro de la explicación aunque no interpreta la explicación como un esquema de argumento.

Observemos, antes de proseguir, un detalle que puede clarificar el rumbo de la investigación. Ambos autores, Coffa y Railton, hacen importantes aportaciones relacionadas con las explicaciones que están involucradas con probabilidad pero suponen para ello el ideal de buena explicación en un sentido fuerte, pensando a la buena explicación como aquella que corresponde a los hechos, lo que podríamos llamar “explicación adecuada” en el sentido de Scriven⁷², cuando intenta distinguir entre explicación y fundamentos de la explicación. Por supuesto, una explicación adecuada es una explicación que además es correcta también en el sentido de Scriven⁷³, pues si una explicación es correcta quiere decir que las proposiciones de su *explanans* son también verdaderas, esto es, ¿de qué otro modo podría estarse refiriendo a los sucesos realmente acaecidos? Scriven propone estas dos calificaciones de explicación (y una tercera, “explicación irrelevante”, que vimos junto con éstas en el capítulo anterior pero que trataremos después, en la sección 2.1.5. “Causalidad, Contexto y Explicación”) como un intento para distinguir la explicación de sus fundamentos. Así las cosas, uno podría

⁷² Véase, *supra*, sección La Tensión Hempel-Scriven.

⁷³ *Idem*.

preguntarse si será posible mostrar los rasgos mínimos de una explicación sin que ello implique ni siquiera si estamos ante una explicación con componentes verdaderos, ya no digamos los reales, es decir, ¿Podemos distinguir algo como una *explicación mínima* respecto de una explicación que además de mínima sea una explicación correcta y, a esta última, distinguirla de una explicación adecuada?

La idea de evaluar bajo un ideal de buena explicación las propuestas que hemos estudiado, puede colocarnos en una mejor posición para estudiar el problema de la elucidación formal de la explicación si tomamos en cuenta estas distinciones de Scriven. ¿Qué diferencia hay entre suponer que una buena explicación es una explicación que corresponde con los hechos (el sentido fuerte del ideal que hemos descrito) y otras formas de suponer que si algo es una explicación, tiene que ser una buena explicación?

Si extendemos esta forma de pensar los problemas alrededor de la modelación de la explicación científica, podemos obtener incluso mayor luz. Puede considerarse, por ejemplo, que la propuesta de Hempel para dar solución al problema de la ambigüedad epistémica también supone un ideal de buena explicación, pero por el hecho de que prefiere la mejor explicación posible y no cualquiera que cumpliera algún canon mínimo de lo que puede considerarse una explicación. Veámoslo detenidamente. El problema que Hempel percibe en su modelo I-S tiene sentido si se supone que se quiere obtener una forma para discriminar entre una familia de argumentos I-S, el argumento que corresponda más a la realidad. Aunque esta no es la clase de supuesto fuerte del que parten Coffa y Railton, puede verse como correspondiendo a una versión débil del ideal de buena explicación: a Hempel le interesa la mejor explicación, aunque la mejor explicación en este caso no pueda (aunque se desee) equipararse con una restricción a las explicaciones adecuadas. Para ello Hempel elabora su RME, aunque lamenta las consecuencias de relativización que RME produce en las explicaciones I-S. Observemos que si Hempel no estuviera interesado, en el fondo, en hallar el argumento que *realmente* representa *la* explicación de los eventos involucrados, no se vería urgido a construir una

restricción como RME. Esto también se refleja en su comparación de los argumentos I-S con los argumentos D-N. Si bien Hempel es mucho más moderado que Coffa y que Railton respecto de los tipos posibles de explicación, pues tiene a los argumentos I-S por explicaciones, todavía se resiste a darles el mismo crédito que les ha otorgado a los argumentos D-N. Lo epistémico, en Hempel, es concebido como un mal necesario. Así, lo que hemos descrito apunta a que Hempel también asume el ideal de buena explicación, entendido como hemos propuesto con las distinciones de Scriven.

Hay un detalle más que es importante notar, relacionado con la forma en que Hempel concibe el problema de la ambigüedad epistémica. Hempel tiene muy claro que también pueden ocurrir cambios en las premisas de un argumento D-N. En **[Hempel & Oppenheim 1948]** se menciona que ante nueva evidencia es posible que las premisas no sean consideradas ya como verdaderas y, en tal caso, la explicación del evento no puede considerarse más la que contiene las premisas falsas⁷⁴. Lo que indica es que la principal preocupación respecto de los argumentos I-S, para Hempel, no puede ser la posibilidad de cambio en el valor de verdad de las premisas, él es consciente del fabilismo en el que puede estar involucrada la empresa científica. La preocupación de Hempel empieza cuando se percata de las inconsistencias inductivas a las que lo llevan las explicaciones I-S. Para Hempel, este resultado de las explicaciones I-S, es la violación de un requerimiento fundamental de todo esquema válido: no conducir a contradicciones. Si esto le pasa a los esquemas I-S, las conclusiones que de ellos se obtienen no pueden ser confiables.

Vayamos entonces al fondo del problema de las inconsistencias inductivas. El problema de la ambigüedad epistémica es la incertidumbre de si la conclusión de un esquema I-S es confiable o no. Pero la preocupación de Hempel se centra en la confiabilidad de la inferencia y no en la posibilidad de cambio en las premisas. El problema es más bien que aún a pesar de concebir que todos nuestros datos iniciales en el

⁷⁴**[Hempel & Oppenheim 1948]** Hempel, Carl, G. y Oppenheim, Paul, "Studies in the Logic of Explanation", en Pitt, Joseph, *Theories of Explanation*, Oxford University Press, U.S.A, 1988, p. 11. Fuente original: "Studies in the Logic of Explanation", *Philosophy of Science*, vol. 15, pp. 567-579, 1948.

explanans son verdaderos, ante nueva evidencia es posible que nuestra conclusión se venga abajo, defecto del cual no adolecen los argumentos D-N. Esto se debe a que la inferencia de un esquema I-S no confiere una certeza absoluta a su conclusión. El verdadero problema de Hempel parece ser que, aún supuesta la verdad de las premisas de un argumento I-S, no es posible asegurar la conclusión. De modo que a Hempel le es más preocupante la debilitación de la noción de derivabilidad que la falsedad del *explanans*. Por una parte, esto muestra lo profundamente amalgamada que se encuentra su idea de explicación con su idea de derivación deductiva: Hempel juzga el valor de las explicaciones I-S a través de un ideal de buena explicación asociado al modelo D-N. Además, en consecuencia, su ideal de buena explicación mantiene el supuesto de alto grado de certeza en la derivación, idealizado en D-N. Por otra parte, además de que una explicación I-S nos coloca ante un tipo de derivación que carece de algunas cualidades apreciables, nos coloca ante un tipo de derivación que tiene una incómoda cualidad: que, en cierto sentido, puede producir distintas conclusiones alternativas contradictorias. Los ejemplos *Rec* y \neg *Rec* son dos esquemas que pueden coexistir en un mismo marco de conocimientos. Si creemos en ambos esquemas, con un mismo conjunto de datos iniciales tendremos dos alternativas de consecuencia que son mutuamente contradictorias, a saber, que Juan se recupera y que Juan no se recupera.

Esta consecuencia es catastrófica para Hempel. Él percibe la consecuencia como equivalente a la inconsistencia del conjunto de conocimientos obtenido a partir de dichos datos iniciales. Lo cual implica que, si no se resuelve este problema, las explicaciones I-S son completamente inútiles. Ahora bien, mientras que la ambigüedad epistémica muestra estos dos problemas, el de la debilitación de la relación inferencial y su consecuente búsqueda de un límite a sus consecuencias y, el problema de las inconsistencias, la solución de Hempel, a saber RME y la relativización epistémica, son producto de los problemas más la asunción de otro factor: un ideal de buena explicación. Bajo este ideal de buena explicación, debería buscarse una forma de discriminar entre los argumentos I-S el argumento que corresponda más con la relación que guardan los eventos en la realidad.

Por ello Hempel se dedicó a buscar algún criterio que le sirviera para ello. El resultado es RME. Para Hempel, la forma racional en que uno debería abordar el problema, es buscando descalificar alguna de las dos conclusiones en conflicto con base en la suposición de que disponemos de toda la evidencia pertinente al *explanandum*. RME discrimina entre las conclusiones discriminando entre los argumentos.

El camino que lleva a Hempel a la relativización epistémica comienza con la suposición de que disponemos de toda la evidencia pertinente al *explanandum*. Hempel está consciente de que esto no será posible más que idealmente. En la práctica, la evidencia disponible generalmente es incompleta y adquirida en distintos estadios. Así, en un momento de nuestro conjunto de conocimientos un argumento I-S se encontrará apoyando cierta conclusión pero en un momento posterior es posible que ya no pueda seguir apoyándola. La nueva evidencia anulará el argumento anterior. Esta es la forma práctica en que se puede tomar RME. Si se descubre nueva evidencia, RME puede anular ciertas conclusiones que dependen de argumentos I-S. Un caso como este sólo puede concebirse si hay distintos estadios de conocimiento. Aquí nace la relativización epistémica: es un intento por ordenar los momentos de la inferencia, de modo que podamos discriminar la mejor posible explicación, de entre las generadas por los distintos estadios de conocimiento. A pesar de que la explicación I-S se ha planteado en un nivel epistémico, se busca la mejor explicación a nivel epistémico. Lo cual es un intento por preservar en I-S un ideal semejante al ideal de buena explicación preservado en el modelo D-N.

2.1.2.2 Sugerencia de una Nueva Perspectiva

Hemos establecido, mediante nuestro análisis de las críticas 1 y 2, planteadas en la sección 2.1.1., que la plausibilidad de un modelo argumental de explicación depende de dos cosas: i) la posibilidad de una relación inferencial que no esté comprometida con la alta probabilidad y ii), que sea plausible asumir una relación inferencial de este tipo como

una relación argumental. Hemos dicho que la i) parcialmente puede fundarse en la idea de que existen relaciones de derivabilidad, estudiadas por la Lógica, que pueden mantener relaciones de baja probabilidad, como probé en el caso de las soluciones del problema de las salvedades de Hempel. La otra parte de este fundamento debe constituirlo una adecuación de esta relación inferencial no monotónica, específicamente, inferencia por default, a los problemas surgidos en la discusión en la Filosofía de la Ciencia. En este camino, también hemos establecido, mediante el análisis sobre la ambigüedad epistémica, que una perspectiva que abandone cierto ideal de buena explicación puede desvanecer algunos problemas relacionados con la modelación inferencial de la explicación. En realidad, ambos casos (i, ii) establecidos pueden verse como consecuencia del abandono de cierto ideal de buena explicación. Esto es así pues podríamos ver el caso de la inferencia altamente probable como constituyendo el ideal mencionado y a la inferencia no comprometida con esto, como casos de explicación no necesariamente buena. Una vez relacionados ambos casos, pensemos detenidamente si una perspectiva filosófica diferente podría aportar la clave para los problemas relacionados con la representación inferencial de la explicación.

En primer lugar, no resulta completamente claro qué sería el abandono de un ideal de buena explicación, puesto que no resulta claro qué es exactamente lo que distinguiría una explicación de una buena explicación. Aún en el caso de que tomáramos como frontera de la distinción justamente la frontera entre explicación y fundamentos de la explicación (de Scriven), la distinción no estaría clara. En contra de una tal distinción podría preguntarse, por ejemplo, si la mención de una relación causal es o no parte de la definición de explicación o sólo de sus fundamentos. Si es considerada parte de la definición o no, puede ser algo muy arbitrario.

No obstante, podemos pensar esto desde un punto de vista más pragmático. No intentemos definir de una vez y para siempre la noción de explicación o buena explicación, sino que intentemos dar con una definición de ello que pueda servirnos para ciertos

efectos. Un criterio posible sería dar una definición que pudiera dar cuenta de los diferentes problemas planteados en la discusión de la explicación en Filosofía de la Ciencia, lo cual es nuestro principal interés. Más específicamente, se trata de intentar una definición que dé cuenta del problema de la representación inferencial de la explicación científica, tal como se le ha presentado en este ámbito filosófico. Esto nos invita a cuestionarnos si es posible una representación mínima de la explicación particular, que pueda tener por lo menos algunas de las ventajas de reivindicar la intuición de que una explicación puede representarse argumentalmente. Ulteriormente, quizá, una exploración del problema desde esta perspectiva, podría contribuir a una definición realista de la explicación, es decir, a una teoría de la explicación tal como es asumida en ciertos contextos.

Dada la reflexión anterior, propondré como una distinción inicial, una diferencia entre explicación y buena explicación basada en la distinción scriveana explicación-fundamentos de la explicación. Aunque es una distinción propuesta estipulativamente, es una distinción que, como se demostrará después, resultará útil en el análisis de la noción de explicación en el contexto de su discusión en Filosofía de la Ciencia. Hay que enfatizarlo, esto debe considerarse una propuesta tentativa de distinción para hallar un modo de dar cuenta de los problemas clásicos relacionados con la modelación argumental de la explicación. Una distinción así, sería la de pensar a la explicación como independiente de la verdad de los miembros de su *explanans*, e independiente de si es adecuada o no. Una explicación que cumpla con la verdad y la adecuación sería una buena explicación. exploremos esta distinción inicial.

Si abandonamos el ideal de buena explicación, basado en la distinción de Scriven a la que me he referido en el párrafo anterior, el grado de problematización que implica la ambigüedad epistémica se atenúa. Abandonar dicho ideal implica no considerar la verdad del *explanans* como una propiedad necesaria de las explicaciones y tampoco considerar la adecuación de las explicaciones como uno de sus rasgos mínimos.

Una vez que la importancia por construir explicaciones con *explanantes* verdaderos no es considerada crucial para catalogar algo como una explicación, podemos concebir dos argumentos en conflicto como explicaciones. Esto se sigue incluso en el caso de que ambos *explanantes* sean verdaderos pues podrían no generar ambas explicaciones adecuadas. Un problema generado, entonces, es la contradicción de las proposiciones *explanantes* de dos argumentos en conflicto. Abordaré este problema en el siguiente capítulo.

Una posición así, para abordar la ambigüedad epistémica acepta de entrada una posición epistémica sobre la explicación. La idea de que la explicación no necesita tener un *explanans* verdadero, que algo es una explicación a pesar de que no cumpla con RME y que cualquiera de las consecuencias en conflicto puedan ser parte de una explicación, debería conducirnos naturalmente a una forma de entender la explicación como una entidad epistémica, en contraste con la idea, por ejemplo de Railton, de que la explicación debería representar relaciones nómicamente relevantes y verdaderas que sean satisfechas por el caso real. El costo filosófico de una posición que abandone esta clase de ideal de buena explicación es, así⁷⁵, que la explicación pueda entenderse como una tentativa de elucidar la ocurrencia de un fenómeno, no como algo que describe cómo realmente es que ocurrió dicho fenómeno, sino sólo entenderse como una forma de dar cuenta de él. La asunción de una noción mínima de explicación, basada en la distinción de Scriven, nos lleva naturalmente a una perspectiva epistémica para entender las explicaciones: a una teoría que considera las explicaciones como objetos de una teoría que intentan dar cuenta de fenómenos, pero que en tanto teoría, pueden fallar por lo menos de dos maneras: en su corrección y en su adecuación.

Así las cosas, sería importante considerar que las explicaciones I-S puedan ser afectadas por la adquisición de nueva evidencia. Toma una nueva forma de entenderse la

⁷⁵ Como lo expliqué más brevemente en mi tesis de maestría.

idea de que los esquemas de argumento I-S deberían poder ceder el paso a otros cuando aparece nueva evidencia. Esta forma de discriminar entre los argumentos resultará de importancia capital. Pues si la explicación es tomada como una propuesta que no está comprometida con este ideal de buena explicación, convendrá construir un mecanismo de representación para la falibilidad de las explicaciones. Un mecanismo de esta naturaleza debería dejar abierta la posibilidad de que algunas conclusiones puedan dejar de inferirse debido a la aparición de nueva evidencia. Un tipo de derivación más débil que la derivación deductiva puede tener estas características. Como dije en la introducción, he mostrado algunas de estas características de una forma de derivación de esta clase en la tesis de maestría, donde abordo el problema de las salvedades de Hempel.

La aceptación de la debilitación de la relación de derivabilidad es también mucho menos problemática si nos olvidamos de este ideal de buena explicación y si continuamos con una perspectiva epistémica. Aún más, aparece como algo que captura mejor la dinámica de nuestros estados de conocimiento. En general no tenemos absoluta certeza de lo que aducimos en nuestros *explanantes*. Y tal conocimiento contenido en el *explanans* muchas veces está sujeto a cambio. De este modo la inducción y, en general, las relaciones de derivabilidad más débiles que la derivabilidad deductiva, cobran mayor importancia. Esto nos lleva a una mayor fundamentación de, la idea de que una relación inferencial más débil pueda representar la explicación científica.

Hay que percatarse de que la relativización epistémica, a pesar de que Hempel haya llegado a ella por la vía de las explicaciones I-S y sus problemas, es una propiedad que no es peculiar de las explicaciones I-S. Si observamos detenidamente, no se ve con claridad que sea necesario rechazar la relativización epistémica en el caso de las explicaciones D-N. Podría ser completamente adecuado extender la relativización a las explicaciones D-N. Una razón para que las explicaciones D-N sean relativizadas epistémicamente es la omisión del requerimiento de que los miembros del *explanans* sean verdaderos. Dejando de lado este ideal de buena explicación, una explicación D-N no

depende de la verdad de estos elementos. Por supuesto, esto no provoca que la explicación D-N pierda su capacidad de asegurar la verdad de la conclusión una vez que se suponen verdaderos los miembros de su *explanans*: las diferencias entre la no deducción y la deducción permanecen intactas.

Así, tanto a Coffa y Railton como a Hempel, el ideal de buena explicación les impide apreciar lo que podríamos llamar “las virtudes de la ignorancia”. Desde la perspectiva que he construido de a poco en lo que va de este capítulo, la noción de explicación debe contener la posibilidad de que el *explanans* pueda ser falso y que pueda también fallar para explicar al *explanandum*.⁷⁶ La *representación racional de la ignorancia*, en este caso, de las explicaciones como son postuladas, independientemente de las restricciones extraepistémicas, permite un acercamiento más fiel a las explicaciones tal y como a menudo se nos presentan. Por una parte, impide las limitaciones para considerar como explicaciones a candidatos de ellas aún cuando no tenemos acceso a la confirmación absoluta de su *explanans*. Por otra parte, la *representación racional de la ignorancia* puede también incluir una forma de concebir la relación entre el *explanans* y el *explanandum*, en la que la aceptación del último no se encuentre subordinada al supuesto de alto grado de certeza. Finalmente, la *representación racional de la ignorancia* además permite una concepción de la explicación como un elemento dinámico de la actividad científica, pues la estrategia de relativización epistémica con la que puede ser abordada, podría permitir la representación de lo que puede llamarse “cambio explicativo”: el cambio en considerar como verdaderos o falsos a miembros del *explanans* y las consecuencias de esto en la teoría; o bien el cambio en la consideración de que las explicaciones sean o no adecuadas, y sus consecuencias en la teoría.

Esto conduce al último punto que abordaré en este apartado. Una consecuencia de una perspectiva epistémica de esta naturaleza, para abordar la explicación es que el

⁷⁶Un caso que ilustra el problema de la no correspondencia, mencionada también en la nota anterior, es el de *Puente* de Scriven. La falla en una de las arcadas y la bomba, ambas son ciertas, pero intuitivamente uno pensaría que sólo la falla (que es la única causa) explica el derrumbe del puente.

contexto tiene un papel esencial en la representación de la explicación. La relativización epistémica de Hempel, no es sólo una forma de solucionar un problema, sino la representación de un elemento inherente a toda explicación particular, sea ésta representada deductivamente o no: el contexto.

Una perspectiva epistémica de la explicación, entonces, puede reunir varias características señaladas críticamente en contra de los modelos clásicos. Nos conduce a una noción de explicación independiente de la verdad, de su adecuación y de la alta probabilidad. Se trata de una noción mínima de explicación que podría ser representada por una relación de inferencia no monotónica que incorporara el contexto y que pudiera representar casos de baja probabilidad. La adecuación final de esta clase de representación inferencial en la que hemos intentado ubicar un modelo de explicación i) dependerá también de un análisis de su relación con el contexto. Por ahora abordaré su relación con ii), la idea de que una relación inferencial de este tipo pueda o no constituir una representación argumental de la explicación.

2.1.3 Supuesto argumental y No monotonicidad

2.1.3.1. La Deducción en las Propuestas de Explicación

En la introducción mencioné dos problemas importantes de la tentativa de modelar, desde una perspectiva deductivista, la explicación científica: los problemas de idealización de la deducción y el de trivialización de la deducción. En el primero de ellos, Hempel mismo muestra que un modelo deductivo de la explicación no sería posible, debido a las salvedades implicadas en toda explicación. En el segundo, Orayen demuestra que hay cierto tipo de explicaciones que no pueden ser vistas como casos claros del modelo D-N. En el caso de Hempel, la conclusión para él es que no es posible reconstruir formalmente la explicación ni siquiera en una propuesta probabilística. Orayen opta por sugerirnos la incorporación de elementos pragmáticos en la modelación de la explicación

o bien tomar en cuenta aspectos intensionales de la Lógica. Como hemos visto en los párrafos precedentes, la idea de sostener el supuesto argumental (SA) bajo un formato probabilístico también a tenido sus problemas en el caso del modelo I-S.

La idea de un camino de modelación alternativo subyace, aunque no muy explícitamente, en la literatura sobre la explicación. El alejamiento respecto de los modelos clásicos y la búsqueda de caracterizaciones que abandonen SA y se encuentren basadas en probabilidades es una muestra de ello. Esto constituye un alejamiento general de una perspectiva deductivista del análisis de la explicación.

Ya Scriven anunciaba lo inadecuado de una perspectiva como ésta para la elucidación de la explicación, cuando decía que el peor error de la propuesta Hempel-Oppenheim era su sobre-simplificación basada en los esquemas deductivos. El propio descubrimiento hempeliano de las explicaciones I-S puede verse como el choque contra una realidad inesperada en la exploración de los argumentos explicativos. Como era de suponerse, la reacción de Hempel fue interpretar esta nueva realidad en el sentido viejo dominado por su ideal deductivo. Esta es la razón de que el requerimiento de máxima especificidad junto con la relativización epistémica hayan sido tomados tácitamente por Hempel como debilidades del modelo.

La mayoría de los filósofos de la explicación fueron sensibles a la incomodidad de la perspectiva deductivista. Pero no sólo a esto, sino también a la idea de que hay otras formas de abordar el problema. Railton, van Fraassen y Salmon lo fueron.

Railton se propone capturar explicaciones de fenómenos azarosos o improbables. Aunque este objetivo lo lleva a abandonar SA, ello muestra un claro interés por investigar el lado indeterminista de la explicación. Van Fraassen identifica el supuesto de que una explicación debe ofrecernos bases firmes para creer y consecuentemente decide también no fundar su investigación sobre SA.

Hay otro aspecto que hace muy interesante la investigación de van Fraassen. En su análisis sobre la causalidad para elucidar la explicación, llega a la consecuencia de que la elucidación más viable de la misma es mediante contrafácticos, como lo hace David Lewis. Para van Fraassen, una elucidación de la causalidad nos llevará a la conclusión de que esta última es una noción contextualmente dependiente. Su forma de demostrarlo es probar la dependencia contextual de los contrafácticos. Uno de sus ejemplos es el siguiente: supongamos que a Francisco le gustan mucho las mujeres, si Francisco fuera mujer ¿Le gustaría mucho los hombres o sería lesbiana? Van Fraassen piensa que para decidir la solución a una pregunta como ésta necesitaríamos recurrir a un contexto particular. En cada contexto se consideran fijas algunas creencias y se consideran mudables otras. Si el contexto mantiene fija la idea de que a Francisco le gusta la gente del sexo opuesto al suyo, y mutable la idea del gusto por las mujeres entonces la respuesta es que le gustaría mucho los hombres. Pero si lo fijo y lo mutable es invertido en otro contexto la respuesta es que sería lesbiana⁷⁷. Para van Fraassen estos ejemplos muestran que la noción de causa, basada en contrafácticos, es dependiente del contexto. De aquí van Fraassen parece decidirse por una relación de pertinencia relativizada contextualmente, que abordaremos después en este mismo capítulo.

Lo importante para la discusión de este apartado es la idea de que algunas de nuestras inferencias pueden estar determinadas por un juego entre lo que mantenemos fijo y lo que estamos dispuestos a considerar mutable. Si lo que hemos decidido considerar mutable ha sido, a su vez, inferido mediante otras premisas, esta última inferencia no puede ser considerada deductiva.

Van Fraassen está consciente de que las propiedades de un condicional contrafáctico son distintas de las de un condicional deductivo. Según van Fraassen, como

⁷⁷[Van Fraassen 1980] Van Fraassen, Bas C., *La Imagen Científica*, Trad. de Sergio Martínez, Col. Problemas Científicos y Filosóficos Paidós/IIF-UNAM No. 2, Paidós, México, 1996, p. 149. Fuente original: *The Scientific Image*, Oxford University Press, 1980.

en los condicionales coloquiales, en los contrafácticos subyace en el antecedente una cláusula *ceteris paribus* (es decir, todo lo demás permaneciendo igual). Un efecto de esto, afirma van Fraassen, es que no es posible aplicar en ellos el refuerzo del antecedente⁷⁸ Esto es, si en un condicional contrafáctico aumentamos información en el antecedente original puede ser que el consecuente ya no pueda inferirse del antecedente incrementado. La implicación de los contrafácticos no obedece a una relación de “suficiencia-necesidad”⁷⁹.

Aun cuando van Fraassen no se da cuenta de la importancia que puede tener esto para una elucidación de la explicación basada en SA, parece claro que este tipo de inferencias están estrechamente relacionadas con la estructura interna de la explicación. No es gratuito que van Fraassen se haya encontrado con ellas de paso por su análisis de la explicación⁸⁰.

Como se ha visto en las secciones precedentes, Salmon ha sido uno de los principales promotores contra la perspectiva deductivista para elucidar la explicación. Además parece estar muy consciente de la estrecha relación que tienen el tipo de inferencias no deductivas de las que habló van Fraassen, con nuestra idea de explicación. En esta misma medida parece más consciente que van Fraassen, del problema que implica partir de una perspectiva deductivista para el análisis de la explicación. Para Salmon, una distinción fundamental en el estudio de la explicación es la distinción entre lógica deductiva y lógica inductiva. En uno de sus artículos señala cómo la última, no cumple con transitividad, contraposición ni refuerzo del antecedente. Leyes que cumplen normalmente las inferencias deductivas. Salmon agrega:

⁷⁸[Van Fraassen 1996] (1980), *op. cit.*, p. 145.

⁷⁹*Idem.*

⁸⁰Hay otras investigaciones de la explicación en las que también se intentó reconstruir un tipo de inferencia como éste.

Es esta básica disimilitud entre las relaciones deductivas y las inductivas (probabilísticas) la que da origen a lo que Hempel llamó *la ambigüedad de la explicación inductiva-estadística*.⁸¹

Sin embargo como puede notarse por la cita, Salmon no puede concebir una lógica inductiva que no sea probabilística. Y, como he dicho en este mismo capítulo tampoco puede concebir una relación inferencial inductiva que no se encuentre basada en el supuesto de alto grado de certeza (o de bases firmes para creer, como las entiende van Fraassen).

En contraste, dados los problemas que hemos analizado, y la forma en que quizá podrían solucionarse, me interesa sobre todo proponer, como lo hice cuando partí del problema de las salvedades en la Introducción de este trabajo, que una propuesta epistémica claramente no deductivista puede ser una opción natural para modelar la explicación científica. Intentaré explorar la explicación sin abandonar SA pero desde una perspectiva no deductivista. No obstante, la esperanza de obtener con ello una buena elucidación de la explicación no está basada únicamente en este cambio de perspectiva. Como he dejado claro también se encuentra basada en supuestos acerca de la naturaleza epistémica de la idea intuitiva de la explicación.

Por otro lado, he expuesto varios problemas para modelar la explicación. Primero, a través de contraejemplos, principalmente de Scriven, Salmon y Railton, se mostró un ataque esgrimido contra cualquier propuesta de modelación de la explicación que involucre la asunción argumental (SA). Segundo, mediante las críticas de Salmon y Railton vimos las dificultades de sostener un modelo probabilístico que descansa sobre SA (ambigüedad epistémica). Tercero, exploramos la posibilidad de una inferencia no

⁸¹"It is *this* basic disanalogy between deductive and inductive (or probabilistic) relations that gives rise to what Hempel called *the ambiguity of inductive-statistical explanation*". [Salmon 1984b] Salmon, Wesley C., "Scientific Explanation and the Causal Structure of the World" en [Ruben 1993] Ruben, David-Hillel (ed.), *Explanation*, Oxford University Press, E.U.A. 1993, p. 90. Fuente original: "Scientific Explanation and the Causal Structure of the World", Princeton University Press, 1984.

deductivista sino más bien no monotónica para representar la explicación, a partir del análisis de un ideal de buena explicación y la consecuente perspectiva epistémica desde la cual abordar una idea mínima de la explicación. La exploración de una relación inferencial no monotónica vino a raíz del caso descrito en la introducción, sobre el problema de las salvedades de Hempel. Un caso igualmente contundente que el de Hempel en contra de la posibilidad de una representación inferencial deductiva es el propuesto por Raúl Orayen que por razones de espacio no podré describir en este trabajo. En los primeros dos casos, el de SA y el de la ambigüedad epistémica, ofrecí un análisis de los problemas y una sugerencia de solución. El tercer caso, el de las salvedades y el de Orayen, puede verse como solucionados a partir de una inferencia por *default*⁸².

Lo que haré enseguida es clarificar las razones que di antes y ofrecer algunas razones más para consolidar las sugerencias en los casos primero (SA) y segundo (Ambigüedad Epistémica). Esto me servirá para argumentar a favor de ii), que es posible ver una inferencia no monotónica del tipo que hemos pensado al final de la sección anterior (es decir, no monotonicidad que pueda representar baja probabilidad), como una representación argumental de la explicación.

2.1.3.2. Un Modelo Argumental

Cuando analicé el problema de la asunción argumental dejé claro mi acuerdo con los siguientes puntos:

- 1) La derivación deductiva no es ni necesaria ni suficiente.
- 2) La derivación probabilística no es ni necesaria ni suficiente.

Mientras que manifesté mi desacuerdo con las afirmaciones:

⁸² Este caso muestra de modo muy agudo los problemas de una representación deductiva al mismo tiempo que apunta brevemente a una incorporación de elementos pragmáticos en un modelo de la explicación.

- 3) Entender la probabilidad en términos de un argumento es inadecuado.
- 4) Modelar bajo SA es inadecuado al menos para ciertos *explananda*, a saber, los eventos que tienen baja probabilidad de ocurrir y que son azarosos.

Por supuesto, el tipo de derivación al que se refieren 1 y 2 es una derivación deductiva y una probabilística, las que suponen respectivamente el modelo D-N e I-S de Hempel. 3 es más general, se refiere a todo tipo de derivación probabilística.

La idea general de mi análisis en las secciones anteriores fue como sigue. Mostré que la crítica contra SA, que es la que sostiene a 1 y 2, está dirigida contra los tipos de derivación supuestos en los modelos clásicos D-N e I-S. Ambos modelos mantienen el supuesto de alto grado de certeza. Por supuesto, no todo tipo de derivación queda agotado en estos dos casos. Existen casos de derivación, dije, que no están comprometidos con el supuesto de alto grado de certeza. De modo que las afirmaciones 1 y 2 valen pero la afirmación 4 se refuta. 3 se refuta si existen casos como los que acabo de mencionar pero que sean probabilísticos.

Ahora bien, mi argumento descansa sobre la idea de que existen casos de derivación que no están comprometidos con el supuesto de alto grado de certeza. Esta clase de derivación podría constituir un modelo de explicación que mantenga SA al menos en su versión SA2. Mostré esta clase de inferencia en mi tesis de maestría, cuando intenté representar mediante inferencias por default, los casos de explicaciones con salvedades.

Si consideramos el límite en la referencia de 1 y 2, la demostración de no necesidad o de insuficiencia de cada tipo de derivación que se proponga para modelar la explicación y que sea distinto de los aludidos por 1 y 2, es una cuestión que habrá que resolver en el caso particular. No podemos, entonces, aducir a 1 y 2 como razones para la imposibilidad *a priori* de un modelo de explicación basado en SA. Así, el punto crucial para determinar si es posible este tipo de modelación es si 4 es verdadera.

En relación a 4, estoy de acuerdo en la existencia de explicaciones cuyo *explananda* tiene baja probabilidad. Pero no coincido en que de esto se siga la inadecuación de cualquier modelo que suponga SA.

La razón más fuerte para rechazar 4 es que de hecho existen tipos de derivación que no suponen un alto grado de certeza. Son tipos de derivación no basados en probabilidad, por lo que la afirmación 3 no les afecta. Por ejemplo, los tipos de derivación de la lógica del razonamiento por *default* de Reiter⁸³. Un esquema de estos no tiene ninguna restricción que lo constriña a admitir sólo premisas que confieran alto grado de certeza a su conclusión.

Ante la existencia de estos esquemas, un crítico de SA quizá procedería primero a objetar que exista una identificación entre esquema de derivación y argumento. En efecto, una primera distinción entre estas dos cosas es que el primero, el esquema, es considerado la estructura de un argumento pero no el argumento mismo. Sin embargo, este no es el punto; al menos algunas instancias del esquema son argumentos del tipo que está en cuestión: argumentos que no confieren alto grado de certeza a su conclusión. Más bien, apoyar la no identificación argumentos-esquemas puede interpretarse como apoyar la afirmación de que si bien a todo argumento le corresponde un tipo de derivación no a todo tipo de derivación le corresponde un argumento. Esto es tanto como distinguir argumentos de la realidad respecto de propuestas lógicas de argumentos (tipos imaginarios de derivación). Es suponer que hay tipos de derivación admitidos y tipos no admitidos.

Ahora bien, lo que llamemos "argumentos reales" es una cosa muy oscura. Tal vez no haya tipos de derivación supuestos en un discurso sino tan sólo relaciones de apoyo entre proposiciones. Salmon, mientras elimina esta clase de argumentos probabilísticos

⁸³Que mencioné brevemente en la introducción.

mantiene una relación de apoyo estadístico entre *explanans* y *explanandum*. Sin embargo, independientemente de si los tipos de derivación propuestos por la lógica contemporánea están realmente presentes en el discurso, pueden funcionar como modelaciones de fracciones del discurso. Este es el sentido en que algunos de los que apoyamos SA pensamos en tales tipos de derivación (esquemas) como argumentos.

Si Salmon defendiera esta idea de una no identificación esquema-argumento entonces Salmon se encontraría criticando SA1 más bien que SA2. La única forma de conectar el rechazo de esta idea de identificación con la idea de si es posible un modelo argumental de la explicación es suponiendo que SA sólo puede sostenerse mediante SA1. Si esto fuera así, ante la idea de que no hay argumentos reales que no estén comprometidos con un alto grado de certeza, las explicaciones tampoco podrían ser argumentos de este tipo. Como SA1 implica que la explicación debería tener estas características reales, se seguiría entonces que no es posible SA basado en S1.

Por el contrario, SA basado en SA2 queda inmune a esta crítica. La razón es que SA2 es independiente de si los argumentos reales son o no como algunos tipos de derivación propuestos por la lógica contemporánea.

Otra posible objeción a la existencia de esta clase de tipos de derivaciones es que la afirmación 3 valga no sólo para casos de probabilidad sino para cualquier caso de argumento no deductivo. Observemos que 3 es una posible consecuencia de la paradoja de la lotería. La idea es suponer, en cualquier tipo de argumento no deductivo un grado de certeza determinado. Ahora, si los argumentos no deductivos incluyen esquemas de derivación entonces hay una contradicción con los hechos, pues entonces los argumentos no supondrían un grado de probabilidad determinado ya que, por lo menos, esto no vale para los esquemas de derivación a los que he aludido antes, los esquemas *default*. Por otro lado, si éstos no están incluidos en la clase de los argumentos no deductivos entonces la crítica se reduce a aducir que no hay argumentos reales con baja probabilidad, pero

entonces la crítica sólo puede sostenerse contra SA1 pero no contra SA2, pues el que no existan realmente argumentos así no significa que no podamos modelarlos con esquemas así.

Por otro lado, afirmar 3 es ya de por sí extraño. Como no hay razones para aceptar sólo parcialmente el argumento de Salmon, basado en la paradoja de la lotería, el argumento de Salmon, tendría que aceptarse también para el caso de cualquier argumento probabilístico. No obstante, esto significa no concebir como argumentos ningún caso de probabilidad, esto es, significa que no tenemos argumentos probabilísticos sino sólo relaciones de apoyo probabilístico. La idea es no admitir que son argumentos. Pero la idea intuitiva de un argumento es la de ofrecer razones para pensar hacer verdadera o aumentar la probabilidad de la verdad de una tesis dada. Y es esto justamente lo que ocurre cuando mantenemos las relaciones de apoyo probabilístico entre diferentes tesis. ¿Por qué entonces no llamar a esto argumentos? La intención quizá viene del rechazo a aceptar como fundadas las consecuencias de un argumento no deductivo. Pero aquí de nuevo salta a la vista la presencia de una racionalidad más rígida de lo que uno podría esperar para la actividad de argumentar. Algo que puede dar cuenta de por qué tendemos a comprometernos con un ideal de buena explicación: la tendencia a aproximarse al ideal deductivista de racionalidad. La siguiente objeción trata sobre las consecuencias de mantener un modelo argumental no deductivo.

Una objeción independiente a la idea de la existencia del tipo de esquemas que he estado mencionando, es rechazar SA dado que su aceptación nos conduciría a la derivación de consecuencias no suficientemente fundamentadas. Al margen de que he mostrado que SA2 puede subsistir aún a pesar de que no haya en la vida real argumentos de baja probabilidad, es interesante notar el fondo de una crítica como ésta. La crítica va dirigida a recordarnos que en argumentos no deductivos las inferencias se realizan a partir de premisas insuficientes. En muchos casos, quizá, no estaríamos de acuerdo en que tales derivaciones se llevaran a efecto. Para decirlo simplídicamente, puede haber

explicaciones de este tipo, pero no argumentos de este tipo, aceptar SA es modificar erróneamente nuestra idea de argumento.

El problema de fondo en esta crítica consiste en considerar las afirmaciones obtenidas a partir de argumentos no deductivos de forma diferente que las afirmaciones obtenidas a partir de argumentos deductivos. Creo que la desarticulación de esta objeción pasa por un análisis epistémico de la similitud entre esquemas deductivos y no deductivos. Hemos visto, como consecuencia del análisis que hemos hecho del tratamiento clásico de la explicación (D-N e I-S), que ambos pueden fallar en la verdad de sus premisas, que ambos pueden ser inadecuados y, sobre todo, que ambos pueden tratarse como relativizados epistémicamente. En realidad, la suposición de sus diferencias en estos aspectos, vienen de la adopción acrítica de una tradición general en el tratamiento de los razonamientos no deductivos. Por ejemplo se piensa que las condiciones de adecuación o corrección de los no deductivos son contextuales y que no ocurre así con los razonamientos deductivos. No me detendré en este paso de la argumentación, pero dado el análisis que he elaborado respecto de las explicaciones D-N e I-S, pienso que puede verse la implausibilidad de un tratamiento diferenciado para ambas clases de argumentos. Sin embargo, daré enseguida un argumento positivo a favor de que pueda considerarse la existencia de argumentos de baja probabilidad. La idea intuitiva de un argumento es, de nuevo, la idea de ofrecer razones para la probabilidad (ésta es absoluta en el caso deductivo) de una tesis. Cuando ofrecemos razones que hacen probable, aunque con baja probabilidad, una tesis, estamos cumpliendo esa idea intuitiva de un argumento. Podemos sin embargo pensar que el argumento de baja probabilidad es un mal argumento, podemos incluso en ciertos contextos rechazar esa clase de argumentos, pero no tenemos razones fuertes para negarles un estatus mínimo de argumentos. De nuevo, una salida a esta forma de pensar consiste en una distinción a la Scriven: una forma de separar la noción mínima de argumento, de su noción fortalecida.⁸⁴

⁸⁴ Estas consideraciones tienen repercusiones interesantes en el tratamiento de argumentos en contextos reales: sus condiciones de corrección (o “validez” para el caso deductivo) principalmente. Una posición así he promovido en el caso de la Teoría de la Argumentación y el Pensamiento Crítico.

Obsérvese por último que el caso de Salmon es muy ilustrativo en el sentido contrario. Salmon abandona SA y, sin embargo, preserva cierta clase de relaciones de apoyo entre el *explanans* y el *explanandum* en su modelo. Históricamente, modelos adecuados para la elucidación de las relaciones de apoyo han sido desarrollados por la Lógica, específicamente, a través de los esquemas de derivación. Esto es una razón a favor de aceptar que éstas relaciones de baja probabilidad pueden considerarse como legítimas relaciones inferenciales, aunque menos buenas que las deductivas. Un modelo argumental de explicación responde a las intuiciones básicas esgrimidas en este apartado.

Todas estas consideraciones son razones en contra de la necesidad del abandono de SA como una estrategia viable para modelar la explicación. Pero también están dirigidas a defender que si encontramos una relación inferencial adecuada para modelar la explicación, aún en el caso de razonamientos de baja probabilidad, podemos decir legítimamente que tenemos una representación inferencial de la explicación, es decir, son razones en apoyo de ii).

2.1.3.3. Un Modelo No Monotónico

La forma plausible de asumir SA para modelar la explicación está relacionada con la idea de una concepción epistémica de la explicación. Un aspecto de la explicación que es interesante y que sale a la luz en el análisis de los contraejemplos dirigidos en contra de los modelos clásicos, es que no es claro el tipo de influencia del contexto en la identificación de explicaciones. Cuando determinamos que un caso de explicación sirve de contraejemplo para un modelo estamos suponiendo que lo que contraponemos es una explicación. En el caso *Puente* de Scriven⁸⁵ podríamos tener dos candidatos a *explanans* del evento de que el puente se destruyó parcialmente, el caso de la fractura en una de las arcadas y el caso de la dinamitación de una de sus arcadas. Tal vez nuestro criterio pueda basarse en la idea de que una explicación siempre refiere a la causa real o, en un modo

⁸⁵Véase, *supra*, sección 1.3.

distinto, nuestro criterio se base en que una explicación sólo tiene que proporcionar causas posibles. En este caso no sólo nuestra concepción de la explicación nos marca la pauta para considerar uno u otro o ambos *explanantes* como adecuados para constituir una explicación del evento en cuestión, pero el contexto también se encuentra interviniendo. En el primer caso por ejemplo todo depende de cuál factor es el que de hecho causó al evento a explicar. Al parecer, el contexto al que puede asociarse un candidato a explicación cambia nuestra idea acerca de si es o no una explicación, pero no se ve con claridad cómo sucede esto.

Si el contexto cambiara, tal vez cambiaría nuestra expectativa respecto del candidato a explicación. Esto significa que el contexto puede cambiar la relación que suponemos que ocurre entre el supuesto *explanans* y el evento a explicar. Llamaré a esta relación "relación *explanans-explanandum*". El tipo de relación que más plausiblemente parecen mantener un *explanans* y un *explanandum* en una explicación es la relación de apoyo.

Tener claro cómo funciona la relación de apoyo entre el *explanans* y el *explanandum* implica clarificar la interacción entre la relación de apoyo entre ambos y el contexto.

En la introducción general expliqué lo que se entiende por no monotonicidad. Es una relación entre un conjunto de información y una noción de derivación. Si el conjunto se expandiera, cabría la posibilidad de que algunas de las consecuencias de esa noción de derivación se perdieran. Una propiedad como ésta la tiene la noción de derivación del modelo I-S de Hempel. Si el conocimiento en K aumenta entonces es posible que algunas de las conclusiones avaladas por el estado de conocimiento anterior ya no estén avaladas por el nuevo estado⁸⁶. Podemos pensar también en la no monotonicidad en un sentido más general. O más claramente, podemos hallar propiedades similares a la no

⁸⁶Piénsese en la conclusión de que Juan se recupera de la gripa y la de que Juan no se recupera de la gripa. Véase, *supra*, secciones 1.2 y 1.3.

monotonicidad en otros ámbitos. Una propiedad como ésta la posee el tipo de relación de apoyo establecida entre el *explanans* y el *explanandum* en el modelo S-R de Salmon. Si los datos de la base probabilística aumentan puede darse el caso de que la relación de apoyo cambie. Podría por ejemplo hacerse al *explanandum* más probable o menos probable. Esto es una propiedad análoga a la propiedad de no monotonicidad que he descrito, sólo que Salmon no la consideraría una relación inferencial.

Como he dicho, parte de la interacción entre el contexto y la relación *explanans-explanandum* es la posibilidad del cambio en ésta última debido a un cambio en el contexto. Esta es también una relación similar a la relación de no monotonicidad. Modelar la interacción entre el contexto y la relación *explanans-explanandum* implica incluir la propiedad de no monotonicidad o alguna similar. Un modelo adecuado de la explicación debe ser capaz de capturar los casos en que un cambio en el contexto produce un cambio en la relación *explanans-explanandum*. El cambio podría producir un cambio en la relación inferencial, debilitándola o fortaleciéndola, pero también podría producir una pérdida de la relación *explanans-explanandum*, el conjunto de proposiciones que constituyen el *explanans* podría ya no poder mantener esa relación con el *explanandum*.

He propuesto que un modelo de explicación basado sobre SA es posible y que resultaría familiar a la relación de apoyo supuesta en varios de los modelos anteriormente presentados⁸⁷. Sin embargo, un modelo formal argumental que estuviera a tono con la posición epistémica que he sugerido y con la necesidad de capturar la interacción entre el contexto y las explicaciones tendrá también que ser un modelo no monotónico.

Los modelos R-S e I-S son modelos que incluyen cierta clase de no monotonicidad. Sin embargo, no capturan detalladamente la interacción entre el contexto y las explicaciones. La no monotonicidad en estos modelos es un resultado accidental de su intento por capturar bien las explicaciones. Son modelos que, contrariamente a lo que

⁸⁷Al menos en D-N, I-S, S-R y D-N-P.

hemos considerado aquí, intentan erradicar algunos de los efectos de la no monotonicidad. Como vimos, se aborda la ambigüedad epistémica como un problema de la explicación y no sólo como un problema técnico de modelación, se estipulan restricciones como RME de Hempel o el requerimiento de homogeneidad de Salmon que impiden modelar en un mismo conocimiento dos alternativas explicativas excluyentes y se intenta reducir al mínimo la relativización epistémica. Una teoría de la explicación que, por el contrario, promueva la forma epistémica de abordar el problema de su elucidación, intentando caracterizar una noción mínima de explicación, debería realizar un tratamiento de la explicación que, en lugar de reprimir al máximo las consecuencias de la relativización epistémica, como lo es la no monotonicidad, la baja probabilidad, la inadecuación o la falibilidad, las explore y las precise. Sostengo, además, que esto es posible mediante una elucidación formal de los elementos de la explicación, una elucidación lógica que incluya la representación de la interacción de sus componentes.

Lo que intentaré en esta investigación construir un modelo adecuado de explicación mínima construido mediante lógicas *default*. Para que el modelo sea una elucidación fructífera de la explicación deberá tener ciertas propiedades que correspondan a problemas planteados en la discusión filosófica. Por lo pronto hemos defendido que puede ser un modelo argumental, epistémico, contextual y no monotónico. Hay otras características del modelo que debemos, sin embargo, examinar. Éstas tienen que ver con lo que podemos llamar “relación de pertinencia explicativa”, que explicaremos en la siguiente sección.

2.1.4. Las relaciones de pertinencia

Hasta ahora la clase de problemas que he abordado tiene que ver con la relación de inferencia involucrada en el candidato a modelo. Como dije al principio de la última sección, la clase de representación inferencial en la que hemos intentado ubicar un modelo de explicación i) dependerá también de un análisis de su relación con el contexto además

de abordar ii), la idea de si una relación inferencial de un tipo debilitado pueda o no constituir una representación argumental de la explicación. He terminado mi análisis de ii). En el resto de este capítulo exploraré i).

Es en relación justamente con las reflexiones que acabo de apuntar en los últimos párrafos, que es pertinente la revisión de otra clase de problemas: los problemas relacionados con la relación de pertinencia explicativa, una clase de relación que parece involucrar las explicaciones con el contexto. Scriven mencionó también, al realizar la separación entre explicación y fundamentos de la explicación, que una explicación podía ser relevante o no, dependiendo de si atendía las exigencias del contexto en donde fue formulada. Como veremos enseguida, estas exigencias involucran asunciones sobre una relación de pertinencia, supuesta entre el *explanans* y el *explanandum*, que parece escapar, como en el caso del ejemplo *Puente*, de una mera relación inferencial y que parece estar en el corazón de nuestras intuiciones acerca de lo que pensamos que es una explicación. ¿Podemos capturar esto en un modelo argumental?

2.1.4.1. Las Asimetrías

De los contraejemplos que presenté antes, los que atacan la necesidad de la relación de derivación (sea deductiva o probabilística) nos indican que hay relaciones de pertinencia (entre el *explanans* y el *explanandum*) que son suficientes para calificar algo como una explicación y que o bien no requieren de un vínculo derivacional deductivo o bien no requieren vínculo derivacional alguno. Esta consecuencia es ajena al problema de modelar la explicación según hemos visto en la sección precedente. La cuestión de si es posible modelar adecuadamente la explicación mediante un argumento es independiente de si algunos casos de explicación son, efectivamente, argumentos o no. Así, el problema real sería si dichos casos son o no susceptibles de capturarse adecuadamente en términos de un argumento. Esto es lo que entiendo cuando digo que las explicaciones tienen forma de argumento.

En contraste, los contraejemplos que se oponen a la suficiencia de la relación de derivación atañen directamente a la posibilidad de modelar la explicación. Dichos casos nos indican que la derivación deductiva y la probabilística, tal y como fueron presentadas por los modelos clásicos, no capturan adecuadamente las relaciones de pertinencia entre el *explanans* y el *explanandum*. Pues a pesar de que el candidato a explicación cumplía con un vínculo derivacional, esto no era suficiente para calificar al candidato como una explicación. Los casos aludidos son el de *Puente* y el de *Vitamina*. Una consecuencia importante de esta clase de ejemplos es que clarifican el hecho de que las relaciones de pertinencia se encuentran identificadas con la relación de expectabilidad (del *explanandum* en relación con el *explanans*) basada en la relación derivacional. En los modelos clásicos, se intenta fundamentar la relación de pertinencia en una relación derivacional. Desde este ángulo dichos contraejemplos son también un ataque contra la tesis de la simetría entre la explicación y la predicción, pues señalan que no siempre que se cumple con la relación de expectabilidad, lo que *ipso facto* caracterizaría la predicción y que es modelada por el argumento, se cumple con la relación de pertinencia que acredita algo como una explicación.

Existe otra clase de contraejemplos que fueron propuestos sobre todo en relación con el problema de la tesis de la simetría, por lo que uno de sus puntos de ataque es la relación de expectabilidad. Los contraejemplos de esta clase también, como en el caso anterior, son ataques simultáneos contra la idea de relación de pertinencia que son capaces de capturar los modelos clásicos. Y por tanto son, a su vez, ataques contra SA⁸⁸. Normalmente se les ha llamado a estos últimos, "ejemplos de asimetría" o simplemente "asimetrías". En razón de facilitar una distinción general llamaré a los casos referidos en el párrafo anterior ejemplos de "**contrapertinencia**".

⁸⁸ En el mismo sentido en que lo son los casos en contra de la suficiencia.

Los casos de asimetría son una subclase de los casos de contrapertinencia. Son casos de contrapertinencia cuya peculiaridad es que en el *explanans* aparece una fórmula general⁸⁹ mediante la cual es posible establecer la inferencia de alguna de las condiciones iniciales a través de las demás condiciones más el *explanandum*. Presentaré en seguida algunos casos típicos de asimetría.

El ejemplo del péndulo es muy usual. La formulación que presentaré a continuación está basada en [Achinstein 1981]⁹⁰:

(Péndulo)

(P)

El período T de un péndulo simple está relacionado con su largo L por la fórmula $T = 2\pi \sqrt{L/g}$, donde $g = 980\text{cm/s}^2$

Este péndulo es un péndulo simple

Este péndulo tiene una longitud de 100cm.

El período de este péndulo es 2 segundos

(I)

El período T de un péndulo simple está relacionado con su largo L por la fórmula

$T = 2\pi \sqrt{L/g}$, donde $g = 980\text{cm/s}^2$

Este péndulo es un péndulo simple

El período de este péndulo es 2 segundos

Este péndulo tiene una longitud de 100cm.

⁸⁹ La fórmula puede ser o bien un bicondicional o una igualdad.

⁹⁰ Tomada de su "Can there be a model of explanation?" en [Ruben 1993] Ruben, David-Hillel (ed.) *Explanation*, Oxford University Press, EE.UU., 1993, p. 143, fuente original "Can there be a model of Explanation?", en *Theory and Decision*, 13, pp. 201-27, 1981. Usaré para las referencias el texto de Ruben.

Obsérvese que en el caso *Péndulo-P* el *explanandum* es el período del péndulo, mientras que en el caso *Péndulo-I* el *explanandum* ha intercambiado lugar con la longitud del péndulo. El *explanans* de cada caso es distinto. La asimetría está dada por la comparación entre, por un lado, las intuiciones acerca de la diferencia de poder explicativo entre ambos *explanantes* (en relación a sus respectivos *explananda*) y, por otro, el hecho de que ambos casos provienen de la misma fórmula general. Como generalmente se interpreta, el caso *Péndulo-P* contiene un conjunto de premisas que es pertinente para la explicación de su *explanandum*, mientras que el caso *Péndulo-I*, no. Así, se aduce que la longitud del péndulo no puede explicarse mediante su período.

Una posible clarificación de la intuición de por qué *Péndulo-I* falla en sus propósitos explicativos podría formularse de la siguiente manera: en cierto sentido la relación entre el largo L del péndulo no es simétrica respecto del período T del péndulo. Esto se debe a que es posible concebir L actualmente sin T actual (sólo potencial). Esto es, la existencia de L no requiere de la existencia de T . Mientras que, por el contrario, no es posible concebir T actual sin L actual⁹¹.

Esta intuición puede aplicarse también a casos de fórmulas generales bicondicionales. Un ejemplo que contiene una fórmula bicondicional es el que se usa en **[Brody 1972]**⁹² para intentar recuperar una relación de pertinencia más fuerte que la de derivación, relación que por lo demás aparece ya en Aristóteles. El ejemplo que él usa es, de hecho, de Aristóteles. Usaré este ejemplo⁹³ para clarificar el problema.

⁹¹ Esta alternativa de clarificación es producto de una discusión con Raymundo Morado.

⁹² **[Brody 1972]** Brody, B. "Towards Aristotelian theory of Scientific Explanation", en Ruben, David-Hillel (ed.), *Explanation*, Oxford University Press, EE.UU., 1993, p. 116. Fuente original: "Towards Aristotelian theory of Scientific Explanation", *Philosophy of Science*, 39, pp. 20-31, 1972. Usaré para las referencias el texto de Ruben.

⁹³ El ejemplo que escribo a continuación es una formulación alternativa pero equivalente del que usa Brody.

(*Planetas*)

(P)

a) Todos los objetos en el firmamento emiten luz constante si y sólo si están cerca de la Tierra,

b) Los planetas están cerca de la Tierra.

c) Los planetas emiten luz constante.

En este caso podemos decir que de a y b se sigue c, además podemos agregar aparentemente sin ningún riesgo que a y b explican a c. Si intercambiamos una de las premisas con la conclusión, tenemos el siguiente argumento:

(I)

a) Todos los objetos en el firmamento emiten luz constante si y sólo si están cerca de la Tierra,

c) Los planetas emiten luz constante.

b) Los planetas están cerca de la Tierra.

Respecto de *Planetas-I* podemos afirmar sin duda que de a y c se sigue b, pero no podemos afirmar con igual seguridad que a y c explican b. Así, la asimetría surge cuando se observa que, bajo un cierto orden, un conjunto de proposiciones puede parecernos claramente una explicación pero el cambio en el orden de algunos de sus componentes produce un choque contra nuestra intuición cuando se quiere tomar al esquema como una explicación.

Por supuesto, obsérvese que la asimetría no es un problema que ocurra como una incoherencia interna dentro de un modelo de explicación científica. Por el contrario, es un problema que emerge cuando intentamos contrastar la adecuación de un modelo con nuestra intuición de lo que es la explicación.

También pueden formularse asimetrías para la derivación probabilística planteada en I-S. Un ejemplo clásico de contrapertinencia, el del barómetro, puede formularse idealizadamente como una asimetría. La presentación como ejemplo de contrapertinencia para la derivación probabilística puede ser la siguiente:

(Barómetro 1)

C = El barómetro cae.

T = Hay tormenta.

$$P(T / C) = .8$$

C

===== [.8]

T

La caída del barómetro hace probable la inminencia de una tormenta, pero no podríamos decir que la caída del barómetro explica la ocurrencia de la tormenta⁹⁴. Esta última aseveración clarifica la relación entre los casos de contrapertinencia y los casos de asimetría. En los casos de asimetría ocurre también la no suficiencia de la derivación: tenemos una relación derivacional, con ello tenemos una predicción pero no tenemos por esto una relación explicativa.

La formulación asimétrica puede ponerse en los siguientes términos:

(Barómetro 2)

⁹⁴ Este ejemplo es una esquematización del que se examina en [Salmon 1984b], p. 105-6.

(P)

C : El barómetro cae.

T : Hay tormenta.

$$P(T / C) = .8 \wedge P(C / T) = .8$$

T

===== [.8]

C

(I)

C : El barómetro cae.

T : Hay tormenta.

$$P(T / C) = .8 \wedge P(C / T) = .8$$

C

===== [.8]

T

De nuevo, el ejemplo de asimetría del barómetro muestra que a pesar de que la relación derivacional sustenta una predicción, no establece necesariamente lazos de pertinencia explicativa entre el *explanans* y el *explanandum*.

Todos estos contraejemplos de asimetría atacan la suficiencia de la relación derivacional para capturar la pertinencia explicativa. Como puede observarse, atacan sobre todo la relación derivacional deductiva y la inductiva-probabilística. La asimetría también ataca la idea de relevancia estadística de Salmon. En un principio, Salmon, en su [Salmon 1971]⁹⁵, pretendía que tan solo con su propuesta de la relación de relevancia estadística y con un criterio adicional era posible capturar completamente las relaciones de pertinencia explicativa. El criterio adicional era la verificación de la relación de enmascaramiento. El problema puede verse como compuesto de varios elementos.

⁹⁵[Salmon 1971] *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh, 1971.

Designemos C, D, T para un cambio en el barómetro, un cambio en la presión atmosférica y la ocurrencia de la tormenta, respectivamente. Todo ello en el contexto espacio-temporal A. Ante el problema de que C es un factor relevante para la ocurrencia de la tormenta y, sin embargo, no explica la ocurrencia de la tormenta. Así, en [Salmon 1984a]⁹⁶ se dice que D enmascara C desde T. Es decir que:

$$(a) P(T/A \wedge C \wedge D) = P(T/A \wedge D)$$

y que C no enmascara D:

$$(b) P(T/A \wedge C \wedge D) \neq P(T/A \wedge C)$$

La idea es que un cambio en la presión atmosférica es pertinente para la ocurrencia de la tormenta y no lo es, en cambio, el cambio en el barómetro. Nótese que el ejemplo *Barómetro 1* (aunque en este caso esté formulado como argumento para enfatizar su punto contra SA), sin tomar en cuenta ningún criterio de enmascaramiento, basta para derribar la suficiencia de la relación de relevancia estadística como modeladora de la pertinencia explicativa. Sin embargo, en la primera formulación de su teoría, Salmon argumentaría que con la relación de enmascaramiento se puede discernir entre el caso pertinente y el caso no pertinente de explicación.

No obstante, en trabajos posteriores, Salmon consideró que la relevancia estadística, aún con el criterio de enmascaramiento antes mencionado, no es suficiente para modelar la pertinencia. Una forma de ver la insuficiencia de la relación de relevancia estadística en combinación con el criterio que involucra la relación de enmascaramiento es cayendo en la cuenta de que no siempre existe una distinción estadística entre el cambio en el barómetro y el cambio en la presión atmosférica. Es concebible un sistema

⁹⁶ p. 106.

idealmente eficiente en el que la probabilidad de la ocurrencia de la tormenta dado A y C sea también igual a la probabilidad de la tormenta dados A, C y D. Los dos casos serían:

$$1) P(T/A) \neq P(T/A \wedge D)$$

$$2) P(T/A) \neq P(T/A \wedge C)$$

Tales que,

$$3) P(T/A \wedge D) = P(T/A \wedge C)$$

El factor C explicaría T tanto como lo haría el factor D.

Así, los ejemplos de asimetría muestran que la relación de pertinencia explicativa entre el *explanans* y *explanandum* no es adecuadamente capturada ni mediante la relación deductiva ni mediante la inductiva-probabilística ni, finalmente, mediante la relevancia estadística.

2.1.4.2. El Supuesto Causalista de la Explicación

Una diferencia pragmática entre los casos de asimetría y los demás casos de contrapertinencia que he mostrado es que los de asimetría muestran más claramente la falta de pertinencia explicativa en los modelos de explicación. La idea de que un mismo conjunto de proposiciones puede constituir secuencias argumentales explicativas así como secuencias argumentales no explicativas provoca la sensación de que la argumentación no es el punto clave para la modelación de la explicación. Esta fue una de las funciones principales de la propuesta de las asimetrías. La sensación anterior es una importante contribución de las asimetrías. Pero debe tomársele con cautela. Las asimetrías parecen mostrarnos claramente la incapacidad de nuestros esquemas argumentales para distinguir directamente los casos que tienen pertinencia explicativa. Sin embargo, esto no implica la inutilidad de SA para modelar la explicación. Habrá que

enriquecer nuestros modelos para dar cuenta de las asimetrías⁹⁷. A este respecto, refiriéndose a otra clase de problemas, [Orayen, 1976] esboza una salida similar. Con lo anterior no estoy comprometiéndome a evitar ni éste ni todo tipo de asimetría en un razonamiento. Lo que sería interesante resolver es más bien la incapacidad del modelo para discernir, en el nivel formal, si cierto orden está representando o no una explicación, o bien si cierto orden es o no una explicación de tal o cual tipo⁹⁸, éste es el problema que nos interesa.

Desde mi punto de vista, la verdadera importancia de los ejemplos de asimetría reside no en haber mostrado una incompleta adecuación de los modelos clásicos basados en SA, sino en una consecuencia filosóficamente más importante: el hecho de que nos arrojan plenamente a la pregunta por cuáles son las relaciones de pertinencia explicativa o a la cuestión de si es posible determinar un conjunto de relaciones de pertinencia que caractericen una relación explicativa. A esto le llamaré "**el problema de las relaciones de pertinencia explicativa**".

La propia propuesta de los ejemplos de asimetría supone que creemos que tenemos la capacidad de discernir, al menos a un nivel intuitivo, las secuencias argumentales explicativas de las que no lo son. Esta creencia no está suficientemente fundamentada. No obstante, tampoco es una creencia que haya sido edificada en puras ilusiones. Un examen de la forma en que son tratados los ejemplos de asimetría y de la manera en que son usados para defender o atacar distintas tesis acerca de la explicación revela un patrón común de racionalidad. La mayoría de los pensadores clásicos de la explicación sostienen los ejemplos de asimetría en el supuesto de que un *explanans* puede

⁹⁷ Por supuesto, cuando hablo de asimetrías me refiero a este tipo particular de asimetría en un razonamiento, a saber, cuando el cambio del orden entre las proposiciones produce una modificación del tipo de cosa que el razonamiento está capturando. En cierto orden el razonamiento captura a una explicación, en otro aparentemente no. En algunos casos el orden en la secuencia de proposiciones de un razonamiento puede provocar modificaciones de distinto tipo del que he mencionado en este apartado. Estas últimas modificaciones son otro tipo de asimetrías. Las únicas asimetrías que me interesan por ahora son las que he presentado en este apartado.

⁹⁸ Veremos más abajo que mi alternativa de interpretación de las asimetrías no solo distinguir entre explicaciones y no explicaciones, sino también distinguir entre tipos de explicación.

ser mejor que otro si el primero aporta alguna descripción de *la* cadena causal que produjo al *explanandum* y el segundo no. Incluso muchos de los mismos autores sostienen una tesis más fuerte: que todo *explanans* legítimo necesariamente contiene alguna mención de dicha cadena causal.

Observemos los ejemplos que he mencionado en este apartado. En la asimetría *Péndulo*, la longitud tiene una cierta magnitud que es la responsable de que el péndulo tenga cierto período y no viceversa; en el caso *Planetas*, la cercanía de los planetas produce su emisión de luz constante y no al revés; en el caso *Barómetro* (1 o 2), la inminencia de la tormenta es la que produce la caída del barómetro y no la caída del barómetro a la tormenta. En todos estos ejemplos, subyace el supuesto de que una explicación puede ser mejor que otra si sólo la primera contiene en su *explanans* la *descripción de los factores causales* que presuntamente producen al *explanandum*. Llamaré a esto el "**supuesto causalista** de la explicación". Algunos de quienes podemos estar seguros que sostienen por lo menos esta versión débil del supuesto causalista de explicación son Brody, Achinstein, Salmon, Scriven, Coffa, Railton, Lewis, Kim, McCarthy e incluso van Fraassen. La idea que cada uno de estos autores tiene de la noción de causalidad puede, por supuesto, diferir radicalmente de la de otros en la misma lista.

Algunos de estos filósofos han sostenido una versión fuerte del supuesto causalista de explicación. El supuesto en su versión fuerte dictaría que todo *explanans* legítimo debe contener alguna descripción de algunos de los presuntos factores que produjeron al *explanandum*. Entre los que abrazaron esta versión del supuesto están Brody, Lewis y Kim. En cambio Salmon, Achinstein, Scriven, Coffa, Railton, McCarthy y van Fraassen dejan abierta la posibilidad de que haya otro tipo de relaciones de pertinencia explicativa distintas de la relación causal.

En términos de modelación Hempel, en vez de sostener francamente el supuesto causalista de la explicación, prefirió centrar la pertinencia explicativa en la mención de

leyes dentro del *explanans*. No obstante, en el fondo Hempel estuvo, aunque no de acuerdo con el supuesto causalista, sí intentando modelar la relación causal como una relación de pertinencia importante para la explicación. Sólo que la forma en que Hempel intentó modelar esa relación de pertinencia de la explicación fue a través de las leyes. Para Hempel las leyes rescataban, entre otras cosas, la relación causal en la explicación, como aparece en [Hempel 1965]⁹⁹. Por supuesto, lo que lleva a Hempel a construir su modelo de la explicación en esta forma es su idea de causalidad. Para Hempel una relación causal asegura siempre la producción de un tipo de hechos cuando otro tipo de eventos ocurre [Idem]. En este sentido, podemos decir que para Hempel la relación causal es determinista. De manera que una relación derivacional deductiva para la explicación se antoja como adecuada para modelar este aspecto determinista de la relación causal tal y como es concebida por Hempel. Esto es así pues una relación deductiva, como la del modelo D-N, aseguraría pasar de la verdad de la ocurrencia de un caso particular de un tipo de hecho (la causa), a la verdad de la ocurrencia de un caso particular de otro tipo de hecho (el efecto).

Así, en algunos casos de explicación para Hempel subyace una relación causal como relación de pertinencia, aunque la forma de capturarla sea sólo como una relación derivacional. Esta forma de capturar la relación causal produce el efecto de que son capturables otras relaciones de pertinencia por el modelo D-N. La única restricción empírica del modelo es que las relaciones mencionadas por las leyes sean verdaderas. La idea de garantizar relaciones de pertinencia verdaderas entre los fenómenos, lleva a Hempel a no aceptar a las explicaciones I-S como explicaciones objetivas, sino sólo como explicaciones relativizadas a cierto conocimiento. La idea es que las explicaciones I-S no contienen leyes confirmadas sino sólo conjeturadas; volveré sobre esto más adelante.

De cualquier forma, podemos pensar el problema de los modelos clásicos como un problema de adecuación de los mismos. Que Hempel cargue sobre las leyes todo el peso

⁹⁹ [Hempel 1979] (1965) Hempel, Carl G., *La explicación Científica*, Edit. Paidós Argentina, Primera edición, Buenos Aires, 1979, p. 343. Fuente original [Hempel 1965].

de la caracterización de la explicación es totalmente racional. Hempel intuye que tratándose de una explicación causal, si capta las regularidades legales en su modelo, capta todo lo relevante de ese tipo de explicación. No exige la causalidad, sólo que su visión sobre la causalidad es determinista y empirista.

Es interesante cómo Hempel afronta el problema de las asimetrías. Hempel intenta anular la fuerza del contraejemplo del péndulo. Hempel apela a la poca fundamentación, en este tipo de ejemplos, para una distinción clara a nivel intuitivo entre lo que es una explicación y lo que no lo es. Así, en el ejemplo del péndulo Hempel se encuentra dispuesto a aceptar como explicaciones ambos lados de la asimetría: lo que podríamos caracterizar como el lado causal y el no causal. Hempel afirma que la explicación causal no es la única que se ajusta al modelo D-N, también está por ejemplo la explicación de leyes por subsunción teórica, o también la explicación no causal de eventos singulares. Algo que es sumamente interesante es que en este último caso de explicación pone el ejemplo siguiente del péndulo:

El hecho de que un péndulo simple determinado tarde dos segundos para completar una oscilación podría explicarse señalando que su longitud es de 100 centímetros y que el período T (en segundos) de cualquier péndulo simple está relacionado con su longitud L (en centímetros) por la ley: $T = 2\pi(L/g)^{1/2}$, donde g es la aceleración de la gravedad.¹⁰⁰

El ejemplo es interesante pues toma como no causal la relación del largo del péndulo hacia su período (lo que habíamos estado tratando como causal) y, a pesar de ello, está dispuesto a aceptarla como una relación de pertinencia explicativa. Hempel piensa que se trata de una ley de coexistencia. Este tipo de leyes, a diferencia de las llamadas por él “leyes de sucesión”, no describen clases de sucesos (causas) que ocurran antes en el tiempo que otros (efectos). Además afirma que en este tipo de casos no hay

¹⁰⁰ [Hempel 1965], p. 347. Sustituí la notación de las fórmulas y las literales que usa Hempel para que fuera uniforme con el ejemplo *Péndulo*, que mostré antes.

en el sentido común fundamentos claros para considerar una derivación deductiva como explicación o no¹⁰¹.

Esta interpretación, la cual puede ser la interpretación estándar del ejemplo *Péndulo*, puede colocarse fácilmente como un punto en contra de la idea de que subyace un supuesto causalista común en los ejemplos de asimetría. En los otros ejemplos de asimetría, las relaciones de pertinencia se suponen entre sucesos y no entre cualidades de objetos. La dificultad para el ejemplo *Péndulo* consiste en que trata con relaciones entre cualidades de objetos más bien que entre eventos. Sin embargo hay un sentido (que ya anticipé) bajo el cual el caso *Péndulo-P*, y no el caso *Péndulo-I*, puede ser entendido como un caso causal. Esto es así si ampliamos un poco nuestra idea de relación causal o bien si ampliamos nuestra noción de hecho.

Veamos ahora el primer caso de interpretación causal, a saber, cuando ampliamos nuestra noción de relación causal. Observemos primero que intuitivamente la cualidad-longitud del péndulo es, en algún sentido, responsable de la cualidad-período del péndulo. Si bien es cierto que el período y la longitud son variables mutuamente dependientes y que cualquier modificación en una de ellas provoca una modificación en la otra, en términos ontológicos la longitud nos parece más primitiva que el período. Ontológicamente hablando, la longitud es prioritaria respecto del período, podríamos decir, fundamenta ontológicamente al período. Como veremos más adelante, esta relación puede entenderse como un tipo de relación causal. Veamos el segundo caso, es decir, el caso en que ampliamos nuestra noción de suceso. Si bien es cierto que si circunscribimos la relación causal a sucesos no nos será posible considerar como una relación causal la relación que guardan la cualidad-longitud y la cualidad-período, también es cierto que en la explicación no nos encontramos relacionando sólo las cualidades de un péndulo. El *explanandum* en el ejemplo *Péndulo* es la descripción de un suceso y no sólo la descripción de una cualidad. En el caso *Péndulo-P*, el suceso que se describe es que el

¹⁰¹ [Hempel 1965], p. 348.

período del péndulo es de 2 segundos. Toda descripción de una cualidad puede entenderse como la descripción de un suceso si tomamos en cuenta no la propiedad en cuestión sino el acaecimiento de la misma en un objeto. El suceso de que el péndulo tenga una longitud de 100cm. produce el suceso de que su período sea de 2 segundos. Creo que alguna de estas dos maneras de ver una relación causal en los ejemplos del tipo del de *Péndulo*, subyace a la propuesta de las asimetrías en la mayoría de los autores. Esto por un lado, sin embargo, aún así, hay maneras de imaginar que no tenemos una relación causal, sea parecida a la causalidad eficiente o sea por fundamentación, en un lado de la asimetría. Pensemos, que cuando inferimos una cualidad 1 de un objeto empírico no necesariamente lo hacemos a partir de una cualidad 2 que fundamenta ontológicamente la cualidad 1. Lo cual nos coloca en la posición de o bien priorizar la relación causal (entendida en dos formas diversas, por ejemplo) o permitir relaciones que no sean causales, como relaciones de pertinencia explicativa.

Dado lo anterior, ante una asimetría podemos tener dos reacciones: considerar ambos lados como explicaciones o bien sólo el lado causal. Ahora bien, es necesario hacer algunas aclaraciones. El problema de las asimetrías se sostiene sobre la idea de que hay alguna relación que se cumple en uno de los lados de la asimetría y en el otro no. La relación que se supone al proponer las asimetrías es la relación de causalidad. Desde mi punto de vista, la relación de causalidad parece tener sentido sólo entre eventos fácticos. En cambio, no parece tener sentido entre eventos no fácticos, como por ejemplo, entre el evento de que una proposición sea verdadera y el evento de que un sistema formal sea consistente. De manera que los ejemplos de asimetría suponen que la explicación debería cumplir con cierta relación entre los eventos fácticos involucrados. De este modo, el problema de las asimetrías se encuentra en el marco del supuesto de una relación de pertinencia explicativa que vincula eventos fácticos entre sí.

Así, tenemos el siguiente dilema. Si consideramos como explicaciones ambos lados de una asimetría entonces estamos dispuestos a aceptar la existencia de una relación no

causal como una relación de pertinencia explicativa. Si consideramos como explicación sólo al lado causal de la asimetría, hay al menos una relación no causal que no estamos dispuestos a aceptar como una relación de pertinencia explicativa. Si lo primero, a menos que aceptemos caer en una posición trivial, deberíamos tener un criterio para discriminar relaciones no causales que son pertinentes explicativamente de relaciones no causales que no lo son. Si lo segundo, entonces si no creemos que la relación causal es la única de pertinencia explicativa entonces tendríamos que tener un criterio de discriminación entre las relaciones no causales para tener claro cuáles son las de pertinencia explicativa y cuáles no lo son (de lo contrario caeríamos también en una posición trivial). De los que abrazan el supuesto causalista de explicación y no creen que la relación causal es la única de pertinencia explicativa, en cualquiera de los dos cuernos del dilema, dejan completamente obscura la forma de discriminar de entre las relaciones no causales, las que son explicativas de las que no lo son.

La dificultad para escapar al dilema anterior reside en la renuencia a abandonar la idea de que la relación de pertinencia explicativa, en ambos lados de la asimetría, enlaza eventos fácticos. Examinemos el siguiente caso de asimetría:

(CR)

(P)

A) La luz recibida por una galaxia sufre un corrimiento al rojo en el espectro si y sólo si esa galaxia se aleja de nosotros.

B) La galaxia NGC7619 se aleja de nosotros.

C) La luz recibida por la galaxia NGC7619 sufre un corrimiento al rojo.

En este caso podemos decir que de A y B se sigue C, además podemos agregar aparentemente sin ningún riesgo que A y B explican C. Si intercambiamos una de las premisas con la conclusión, tenemos el siguiente argumento:

(CR)

(I)

A) La luz recibida por una galaxia sufre un corrimiento al rojo en el espectro si y sólo si esa galaxia se aleja de nosotros.

C) La luz recibida por la galaxia NGC7619 sufre un corrimiento al rojo.

B) La galaxia NGC7619 se aleja de nosotros.

Respecto de *CR-I* podemos afirmar sin duda que de A y C se sigue B, pero no podemos afirmar con igual seguridad que A y C explican B. Como hemos visto, la asimetría surge cuando se observa que, bajo un cierto orden un conjunto de proposiciones puede parecernos claramente una explicación pero el cambio en el orden de algunos de sus componentes produce un choque contra nuestra intuición cuando se quiere tomar al conjunto como explicación.

Como dije antes, la asimetría no es un problema que ocurra dentro del modelo de explicación científica. Por el contrario, es un problema que emerge cuando intentamos contrastar introspectivamente la adecuación de un modelo con nuestra intuición de lo que es la explicación. Pero tal vez nuestra intuición de lo que es una explicación pueda estar sesgada por el contexto que utilizamos para evaluar el candidato a explicación. Veamos esto detenidamente.

La forma más clara de sentir lo contraintuitivo de *CR-I* es leyendo la expresión "A y C explican B" como "el evento A y el evento C explican el evento B", pues no podríamos aceptar que el alejamiento de las galaxias se debe a que el espectro de la luz que recibimos de ellas sufre un corrimiento al rojo.

No obstante hay también un sentido según el cual A y C sí explican B. Si leemos "A y C explican B" como "la proposición A y la proposición C explican la proposición B" parece que el choque contra nuestras intuiciones desaparece.

Si intentamos entender *CR-I* como un argumento que representa una explicación, entonces tenemos dos "sensaciones" epistémicas:

- i) En un sentido, parece extraño que A y C constituyan un *explanans* para B. Sin embargo,
- ii) en otro sentido parece intuitivo que A y C constituyan un *explanans* para B.

La extrañeza de i) se debe a la presunción de la inexistencia de cierta relación entre el evento capturado por el *explanandum* y los eventos capturados en el *explanans*. La no extrañeza de ii) se debe a la presunción de existencia de cierta relación entre el evento capturado por el *explanandum* y los eventos capturados en el *explanans*. La postulación de existencias o inexistencias de ciertas relaciones entre sus respectivos *relatas* tiene que ver con el tipo de *relatas* involucrados por la relación. Si los *relatas* son considerados como eventos no fácticos es probable que no estemos dispuestos a aceptar la existencia de una relación de causalidad entre ellos. En contraste, si los *relatas* son eventos fácticos, sí. En consecuencia, una forma de recuperar la intuición acerca de las asimetrías que acepta el lado no causal como explicativo, es entender la relación de pertinencia explicativa en el lado no causal como una que se da no entre eventos fácticos sino entre eventos no fácticos.

Pocos autores protagonistas en la discusión sobre la explicación científica se ocuparon de esta otra cara de la moneda de la asimetría, debido a que la mayoría de ellos concentró sus esfuerzos por elucidar la explicación en el marco de los eventos susceptibles en principio de contrastación empírica. A partir de la opinión de que la explicación de una proposición significa un cambio de tema, es decir, que pasar de lo

fáctico a lo no fáctico es un cambio de tema¹⁰², la búsqueda por elucidar la explicación científica omitió considerablemente la integración de este otro tipo de explicaciones.

Lo que está por debajo de la contraintuitividad hacia la tentativa de que un argumento como *CR-I* sea considerado como una explicación es, por supuesto, la noción de causalidad y muchos pensadores no tardaron en percatarse de ello. En el caso de los modelos de Hempel y Oppenheim, cuyos autores partieron por lo menos en principio de un análisis de la explicación causal, el problema no estaba en una falta de visión del aspecto causal de las explicaciones sino en que su modelo era incapaz de distinguir esta importante variable.

Algo muy importante de aclarar es que si algunos modelistas intentaran escapar al dilema presentado antes, mediante la aceptación de una relación de pertinencia explicativa que no enlace (en uno de los lados de la asimetría) eventos fácticos sino eventos no fácticos, todavía tendrían que adecuar sus modelos para capturar esta distinción. Si no hay modificaciones en el modelo, las asimetrías indicarían una ambigüedad en el modelo: proposiciones que son tomadas como representando eventos fácticos son también tomadas como si representaran eventos no fácticos. La integración formal de esta otra cara del problema de la asimetría sería un punto clave para la construcción de un modelo uniforme de explicación.

Una estrategia que puede resultar exitosa para realizar la integración de los dos lados de una asimetría es intentar concebir *el núcleo* de la relación de pertinencia explicativa de manera que pueda subsumir ambos lados de la asimetría. Para este propósito, desde mi perspectiva, son de gran ayuda dos cosas: por una parte, examinar los ejemplos de asimetría a la luz de la distinción que he hecho entre tipos de *relata* (fácticos y no fácticos) y, por otra, mantener la noción de justificación separada de la noción de explicación. Así, cuando hacemos una pregunta-porqué exigiendo una justificación de

¹⁰² Como Hempel y Oppenheim lo dejan claro. Scriven es una excepción, pero no profundiza mucho sobre ello.

algún evento fáctico, nos encontramos dispuestos a aceptar cualquier argumento que concluya una proposición que describe el evento en cuestión. En este caso la intuición de *bases para creer* de la que nos habla van Fraassen parece suficiente para *dar cuenta* del evento no fáctico. En cambio, cuando hacemos una pregunta-porqué exigiendo una explicación de algún evento fáctico, nos encontramos exigiendo algo más, las bases para creer inferenciales no parecen suficientes para dar cuenta del evento. Un argumento a favor de la ocurrencia del evento no nos será suficiente para explicar el evento fáctico: tendrá que decirnos cuáles son los eventos responsables de que haya tenido lugar el evento en cuestión. La explicación debe ir más allá del ofrecimiento de *bases para creer* en el *explanandum*, falta todavía algo para que pueda dar cuenta del *explanandum*¹⁰³. No obstante, en realidad, tendríamos que pensar también en la intuición de dar cuenta, para el caso de los eventos no fácticos. Para el caso de los eventos no fácticos, debe suceder algo similar. Dar cuenta de un evento no fáctico debe llevarnos también a la responsabilidad que tienen otros eventos para que el evento en cuestión haya tenido lugar. La intuición de dar cuenta puede ayudarnos a elucidar el núcleo que hace que un esquema argumental, que ofrece razones para creer, pueda funcionar como una explicación. La explicación podría verse como modelable mediante una relación inferencial y condiciones que señalan factores que pueden dar cuenta del evento *explanandum*.

Si lo pensamos de este modo, la intuición de *dar cuenta* nos ofrece una alternativa de integración para ambos lados de las asimetrías. Pero, como veremos más adelante, no es necesario que el alcance de la exigencia de la noción de dar cuenta (como guía para identificar una relación de pertinencia explicativa), sea igual para los dos tipos de *relata* que he identificado antes.

Anticiparé un poco la determinación de dichos alcances, dejando clara la forma en que difiere de mi punto de vista, una opinión como la de Hempel. Para Hempel puede

¹⁰³ Por supuesto, dar cuenta de un *explanandum* es, *ipso facto*, dar bases para creer en él.

existir una relación entre eventos fácticos que no es causal pero sí es explicativa. Mi opinión es que cualquier relación no causal entre eventos fácticos no *da cuenta* del *explanandum*. Sólo proporciona bases para creer en el *explanandum* pero no lo explica. Las intuiciones contradictorias de que el lado no causal de una asimetría es explicativo y de que no lo es, reflejan una confusión entre tipos de eventos y esto, creo, nos acerca a una concepción más correcta del tipo de cosa que debemos pensar que es una relación de pertinencia explicativa.

Si la explicación implica *dar cuenta* del *explanandum* parece exigirse el señalamiento de *factores responsables* de éste. Si la explicación implica dar cuenta del *explanandum* entonces la única relación que conozco que puede cumplir la función de ser explicativamente pertinente entre eventos fácticos es la de causalidad. Qué se entienda con exactitud acerca de esta relación (la causalidad) es un problema importante en la Filosofía de la Ciencia, no obstante, parece que la idea intuitiva de *producción* de fenómenos físicos es la que está en juego cuando, en Ciencia, se exige una explicación. Si una explicación debe comprometerse a dar cuenta del *explanandum*, en el sentido de señalar eventos plausiblemente responsables de él, una intuición adecuada de relación de pertinencia explicativa es justamente la relación de causalidad, sea en el sentido de producción de un evento a partir de otros eventos, causalidad en el sentido de causa eficiente aristotélica o sea en el sentido de fundamentación ontológica. Es esta la intuición que aparece en el debate de los contraejemplos a los modelos clásicos de explicación científica, y es esta relación la que puede cubrir el alcance exigido por la intuición de dar cuenta de un evento. Por tanto, es plausible que, en nuestra intuición de lo que es explicación, estemos suponiendo esta relación de causalidad.

Para el caso de eventos no fácticos es difícil proponer una sola relación de pertinencia explicativa. Sin embargo, mi intuición es que los eventos que se encuentran conectados no causalmente pero mediante una relación que es explicativamente pertinente, lo están sólo si se les toma como eventos no fácticos. Porque creo que en

realidad no hay nada que dé cuenta de un evento fáctico de no ser un evento que pueda funcionar como su causa tentativa. En este sentido, los contraejemplos de asimetría son auténticos contraejemplos a los modelos clásicos de explicación. Cualquier correlación matemática o estadística o alguna otra clase de relación no será suficiente para *dar cuenta* del evento, en el sentido de que no sería suficiente para señalar los eventos que plausiblemente son responsables de la ocurrencia del *explanandum*.

Volvamos a la relación de pertinencia explicativa entre eventos no fácticos. Una tentación es la de capturar la intuición que subyace al considerar el lado no fáctico de una asimetría como una explicación pensándolos como eventos psicológicos. Pero pensarlos de este modo puede llevarnos a postular precisamente la misma clase de relación de pertinencia: la relación de causalidad. Yo puedo pensar en dar cuenta del evento B *qua* mi creencia de que B. En este caso lo que está siendo explicado es por qué tengo la creencia de que B y esto puede ser explicado a través de la idea de que creo en A y en C. Un conjunto de creencias explican otro. Esto puede bien ser subsumido como casos de pertinencia explicativa causal, pues un evento psicológico podría ser tomado, desde cierto punto de vista, como una clase compleja de evento fáctico. Es de nuevo entonces un caso de relación entre eventos fácticos. Lo que me parece sin embargo, es que, en realidad, la postulación de asimetrías parece llamar la atención hacia conjuntos de eventos que no pueden ser tomados como manteniendo relaciones causales. Es por este motivo que se atribuye a un lado de la asimetría una relación no explicativa. De otro modo no podría tenerse este supuesto. En última instancia, la interpretación psicológica del lado no pertinente de la asimetría podrá explicarse con la misma clase de relación de causalidad que atribuimos a su lado pertinente, pero podemos seguir preguntándonos, ¿Hay sin embargo alguna forma en que podamos recuperar la intuición del rechazo del lado no pertinente de la asimetría? Quiero decir, ¿hay alguna forma de pensarlo como una explicación? O ¿Hay alguna relación de pertinencia que guarde este lado de la asimetría y que sea una pertinencia explicativa? Una forma de hacerlo, en el caso CR-I, es pensarlo como eventos semánticos y no como psicológicos, la clase de relaciones de pertinencia

explicativa en esta clase de eventos puede suponer los eventos que enlaza, como si fuesen simples proposiciones. Como he avanzado la idea antes, en el caso CR-I podríamos suponer que lo que se intenta explicar es el evento no fáctico de que la proposición "La galaxia NGC7619 se aleja de nosotros" es verdadera. En tal caso, los eventos de que la proposición A es verdadera y de que la proposición C también lo es nos llevarían a dar cuenta del evento no fáctico de que B es verdadera. Si esto fuese así, ¿Qué clase de relación de pertinencia explicativa podría dar cuenta de estos eventos? Una clase de "producción" de unas proposiciones a partir de otras, puede verse en la relación de derivación. Verificar la relación de derivación entre las proposiciones A y C, y la B, puede dar cuenta, ya que se trata de eventos no fácticos, de la proposición B. La relación de derivación puede verse, en el dominio de los eventos semánticos, como una relación análoga a la relación de causalidad que ocurre entre eventos fácticos. Si B es tomada como un evento semántico, entonces el *explanans* A y C señala eventos semánticos que pueden considerarse como plausiblemente responsables de B. Así, la relación de pertinencia explicativa entre eventos no fácticos estaría capturada como una relación de derivación, pues es ésta la que daría cuenta de los eventos no fácticos representados mediante proposiciones. Considero que esta opción de interpretación del lado no fáctico de la asimetría es más coherente con lo que uno esperaría que se explica en ese caso, es decir, parece, en primera instancia más plausible pensar que lo que se quería era dar cuenta de la verdad de una proposición y no dar cuenta de una creencia.

De este modo los casos de asimetría, en su lado aparentemente no pertinente, podrían tomarse como casos que responden a la exigencia de una explicación de un evento no fáctico, como la verdad de una proposición. Si hacemos esto, la relación de pertinencia supuesta entre las proposiciones del *explanans* y la proposición del *explanandum* es una derivación, en el sentido de que si los miembros del *explanans* son verdaderos, eso explica por qué es verdadero el *explanandum*, la verdad de las premisas explica la verdad de la conclusión. Tomar esta relación como base, pero entre eventos semánticos de este tipo, verdades de proposiciones, es coherente con la idea de que no

ocurre la relación de causalidad en un lado de la asimetría, pero coherente además con la intuición de que, en algún sentido diferente, el otro lado de la asimetría también puede verse como una explicación. Si tomamos ambos lados de estos argumentos (asimetrías) como refiriéndose a la misma clase de eventos (eventos fácticos por ejemplo), entonces se producen las asimetrías.

Dadas las anteriores reflexiones, podríamos clasificar entonces las relaciones de pertinencia explicativa en dos clases: las de pertinencia fáctica y las de pertinencia no fáctica. De modo que lo importante en una explicación es determinar, antes que nada, el tipo de eventos involucrados en ella. Después de haber determinado el tipo de eventos de que se trata, eliminando la ambigüedad que pueda estar en el lenguaje de su emisión, entonces podemos intentar checar si la relación de pertinencia que les corresponde puede establecerse o no. En este caso, la propuesta de las reflexiones anteriores, nos conduce a considerar entonces dos clases de relaciones de pertinencia: la pertinencia causal y la pertinencia lógica. Consecuentemente, nos lleva a considerar dos sentidos de explicación: explicaciones asimétricas y simétricas. En el primer caso, interpretamos la explicación como refiriéndose a eventos fácticos, que pueden ser causales en dos sentidos diferentes: por algo parecido a la causalidad eficiente, o por fundamentación ontológica. Cuando cambiamos el orden de los elementos en una explicación de este tipo puede ser que no tengamos más una explicación, en el supuesto de que mantenemos la misma clase de *relata*. En el segundo caso, interpretamos la explicación como refiriéndonos a eventos semánticos. Cuando cambiamos el orden de los elementos en una explicación de este tipo seguiremos teniendo una explicación, bajo el supuesto de que mantenemos la misma clase de *relata*. Otro modo de ver esto es pensar que cuando, en una estructura argumentativa explicativamente asimétrica, cambiamos los eventos involucrados en el lado no explicativo, de eventos fácticos a eventos semánticos, y corroboramos una relación inferencial, tenemos también una explicación, pero de eventos diferentes. Un modo más simple de ponerlo es: si la intención de la explicación es dar cuenta de eventos semánticos, no es una explicación asimétrica; si, en cambio, la intención es dar cuenta de

eventos fácticos, podría ser asimétrica, en el primer caso, la relación de pertinencia explicativa que uno debería verificar es la relación inferencial, en el segundo, es la relación causal (eficiente o por fundamentación).

2.1.5. Causalidad, Contexto y Explicación

2.1.5.1. Dos Problemas de la Causalidad en la Explicación.

Examinar otros dos problemas relacionados con la causalidad y la explicación nos dará mayor luz sobre la distinción entre relación de pertinencia explicativa y la explicación propiamente dicha. Creo que esta distinción es clave para entender el tipo de perspectiva que deberíamos construir para articular una solución a todos estos problemas.

Algunos filósofos, como Scriven, Achinstein, McCarthy y Lewis, han señalado otro problema que tiene que ver con la pertinencia explicativa causal. El problema es que no parece ser suficiente restringir el *explanans* de modo que describa algún factor causal que sea capaz de producir el *explanandum*. [Achinstein 1981] lo ilustra con el siguiente ejemplo¹⁰⁴:

(Arsénico)

Juan comió una cucharada de arsénico en el tiempo t

Cualquiera que tome una cucharada de arsénico muere en 24 horas

Juan murió en 24 horas a partir de t

El *explanans* de este ejemplo cumple con la relación de derivación y cumple con el requisito de describir factores causales capaces de producir el *explanandum*. Sin embargo,

¹⁰⁴ [Achinstein 1981] Achinstein, Peter, "Can there be a model of explanation?" en Ruben David-Hillel (ed.), *Explanation*, Oxford University Press, EE.UU. 1993, p. 146. Fuente Original: "Can there be a model of explanation?", en *Theory and Decision*, 13, pp. 201-27, 1981.

supongamos que Juan se ha dado cuenta de que lo que ingirió fue arsénico y entonces sube a su auto y corre a toda velocidad rumbo al hospital más cercano. Supongamos que Juan sufre un accidente automovilístico y muere en él de camino al hospital. Juan murió por el accidente y no por el arsénico.

Achinstein usa este tipo de ejemplos para demostrar que la relación de pertinencia explicativa, aun cuando supone que se cumple la condición empírica de los modelos clásicos (que las premisas del *explanandum* sean verdaderas), no puede determinarse *a priori*. Por supuesto podríamos aducir que el arsénico fue la causa remota de la muerte de Juan. Esto nos pondría ante otro problema relacionado con la causalidad, que mencionaré después. Pero podemos encontrar casos en los que no podríamos apelar a la idea de una causa remota. Por ejemplo, el caso *Puente* de Scriven. Lo importante en el caso *Arsénico* es que la relación de pertinencia explicativa no se cumple *a priori* a pesar de cumplir con la verdad de los elementos del *explanans*, con la relación de derivación y con la mención de factores causales adecuados. Estas características las comparte con el ejemplo *Puente*, de Scriven. Así, ambos ejemplos, parecen hacer necesaria la restricción del *explanans* de modo que describa no sólo *una* causa adecuada para producir el *explanandum* sino *la* causa que de hecho lo produjo. Si el modelo no es sensible a *la* causa del *explanandum*, no puede modelar bien la explicación, a esto lo llamaré "**el problema de la sobredeterminación epistémica causal**".

Como Achinstein lo anticipó, la relación que de hecho ocurre entre eventos fácticos involucrados en una explicación no puede determinarse *a priori*. Sin embargo, ¿Es necesario que una explicación describa los hechos que realmente ocurrieron? Si la respuesta a esta pregunta fuese afirmativa estaríamos condenados a calificar de explicación sólo a aquellos candidatos cuyo *explanans* fuera *la* verdadera causa que produjo el *explanandum*. Y esto llevaría a cumplir el objetivo de Achinstein, de mostrar que no es posible la construcción de un modelo formal de explicación científica, pues este factor causal es a menudo cognoscible sólo *a posteriori*. Sin embargo, una exigencia así

reduciría mucho el conjunto de cosas que normalmente pensamos que son explicaciones. En esta cuestión, las intuiciones del sentido común parecen contradecirse. Intuitivamente algo es una explicación independientemente de si su *explanans* es verdadero, pues de otro modo no veríamos como explicaciones las ocurridas en teorías que históricamente han fallado, pero también intuitivamente algo es una explicación sólo si su *explanans* es verdadero, de otro modo no nos parece que realmente explique al *explanandum*. Analicemos esto con calma. Cuando x es propuesto como una explicación y se conoce la falsedad del *explanans* de x, las preferencias normales acerca de x son algo parecido a las siguientes:

- a) "x es *una* mala explicación",
- b) "x es *una* explicación errónea",
- c) "x no es *la* explicación,
- d) "x no explica".

Mientras que a y b apoyan la intuición de que algo puede explicar sin ser verdadero, c y d apoyan lo contrario pues sugieren que hay una sola, que es la que corresponde a ciertos hechos reales.

El problema de la sobredeterminación epistémica causal descansa sobre la idea de que la explicación debe estar fundamentada en una relación de pertinencia explicativa que realmente ocurra entre los eventos fácticos involucrados. Una exigencia encaminada a una racionalidad más dura y no a una propuesta de explicación epistémica que distinga entre explicación y buena explicación *a la* Scriven. Las preferencias a y c ilustran muy bien dos posiciones que pueden tenerse respecto de este problema: o bien estamos dispuestos a aceptar que "x es una mala explicación" o bien estamos dispuestos a aceptar que "x no es la explicación". La diferencia entre ambas posiciones es que en la primera uno no se encuentra comprometiendo la noción de explicación con alguna correspondencia con los hechos, no la comprometemos, como hemos dicho antes, con su *adecuación*, en palabras

de Scriven. He dicho que para solucionar el problema de las salvedades (en la introducción mencionando mi tesis de maestría), el problema de la baja probabilidad y el problema de la ambigüedad epistémica, (en este mismo capítulo) es mejor partir de una posición epistémica de lo que es la explicación. Además, he defendido que una posición así, no comprometida con un ideal de buena explicación y tendiente más bien a formular una noción mínima de explicación, nos ayudará además a representar mejor ciertos aspectos de la explicación, como su relación con el contexto y el cambio a partir de nueva información. Una posición como ésta puede también servir de marco para abordar el problema de la sobredeterminación epistémica causal. El examen de otro problema de la causalidad nos dará mayores elementos para optar por esta perspectiva. Examinaré ahora este segundo problema.

Una misma trama causal puede tener muy variados factores. [Lewis 1986] pone el ejemplo de un accidente automovilístico en el que hay varios factores que son candidatos a constituir el *explanans*. En un accidente, ilustra Lewis: la disponibilidad de gasolina, el cumpleaños de la abuela paterna del conductor, la construcción de la carretera fatal, la posición y velocidad del carro una partícula de segundo antes del impacto, pueden ser cada uno factores explicativos distintos para distintas exigencias de explicación que dependen del contexto¹⁰⁵. Para Lewis, la diferencia importante consiste no en identificar la historia causal de un evento, sino en cuáles son los factores causales que se aducen como significativos para el contexto. El contexto en el que se exige una explicación sobre un evento determina los factores a los que el *explanans* apelará para *dar cuenta* del *explanandum*. Así, a la hora de construir una explicación, no hay forma precisa para discriminar algunos factores causales de otros en una misma trama causal. Se implica que entonces un modelo tendría que ser sensible a determinar la causa correcta en una dada trama causal, a esto lo llamaré "**el problema de la elección causal**" y tiene que ver con las causas remotas que mencioné antes.

¹⁰⁵ [Lewis 1986] Lewis, David, "Causal Explanation", en Ruben, David-Hillel (ed.), *Explanation*, Oxford, University Press, EE.UU., 1993, p. 183. Fuente Original: "Causal Explanation", *Philosophical Papers*, ii, Oxford, University, Press, Oxford, Nueva York, 1986, pp. 214-40. La paginación la haré siempre en relación a la compilación de Ruben.

El problema recién planteado consiste en cómo discriminar el factor causal explicativo para una determinada exigencia contextual. Y aún si podemos, *prima facie*, identificar dicho factor, queda el problema de cómo capturar el factor adecuado en un modelo de la explicación. Los filósofos que se han percatado de la necesidad de la relación causal en general no han implementado algún mecanismo que represente formalmente su presencia en su modelo y, cuando lo han hecho, no es un mecanismo que integre los lados diferentes de una asimetría (como las que examinamos en la sección anterior).

Para el problema de la elección causal puede adoptarse una posición análoga a la que acabo de sugerir en el caso del problema de la sobredeterminación epistémica causal. Lewis parece tender a una posición en esta dirección. Para él algo es una explicación independientemente de si es aceptable o no¹⁰⁶. Así, Lewis no debería tener ningún problema en, por una parte, suponer que el énfasis que puede hacerse sobre un *explanans* en discriminación de otros depende del contexto en que se planteó la explicación y, por otra, suponer también que cualquiera de esos *explanantes* constituye una explicación del mismo *explanandum*. No hay necesidad de elegir el factor causal: cualquier factor causal que pueda producir al *explanandum*, que pueda dar cuenta de él, explica. Esta solución, expresada más generalmente, puede extenderse también para el problema de la sobredeterminación epistémica causal: cualquier relación de pertinencia explicativa que plausiblemente dé cuenta del *explanandum*, explica por lo menos de un modo mínimo. Y esta minimalidad es clave en mi perspectiva de modelación pragmática y epistémica de la explicación.

Ahora bien, aceptar *explanans* que posiblemente no correspondan a los hechos es mucho más sensato que aceptar *explanans* que no correspondan al contexto. En un laboratorio científico en el que se intenta explicar la combustión de una vela es más fácil aceptar la intervención del flogisto como explicación que el acto, efectuado por un

¹⁰⁶ *Ibid.*, pp. 186-7.

ayudante, de prender fuego a la vela. Sin embargo, a pesar de que algunos filósofos puedan rechazar la primera explicación y sostener la segunda observemos que tanto la exigencia de correspondencia con los hechos (adecuación scriveana) como la exigencia de adecuación contextual dependen de una misma concepción general de la explicación: una concepción comprometida con el ideal de buena explicación. En un contexto determinado una relación de pertinencia explicativa mínimamente adecuada puede no parecer una explicación para el *explanandum*. Pero esto se debe más bien a que no parece una buena explicación para el *explanandum*. Definido desde la distinción de Scriven entre explicación y fundamentos de la explicación, el ideal de que algo es una explicación si y sólo si es una buena explicación es algo que puede esclarecer ambos problemas.

Para la sobredeterminación epistémica causal y para la elección causal, la cuestión está en comprometerse o no con un ideal de buena explicación. Comprometiéndose con el ideal de buena explicación que hemos caracterizado, es posible sostener que algo no es una explicación si no cumple con cierta adecuación contextual. Pero precisar la noción de explicación en este sentido implicaría tener claro cuándo una explicación es contextualmente adecuada. Una noción de adecuación contextual que distinga entre el *explanans* que refiere al acto de encender la vela y el *explanans* que refiere al flogisto, debe ser muy compleja. Por lo tanto, encontrar los elementos conceptuales que puedan ayudarnos a identificar este tipo de adecuación podría resultar ser una empresa no muy exitosa.

En [Van Fraassen 1980] se intentó integrar este tipo de adecuación contextual a al modelo erotético-contextual (E-C). Lo que van Fraassen llama "**el problema de los rechazos**" es una versión generalizada del problema de la elección causal. Para van Fraassen ni las asimetrías ni los rechazos pueden ser solucionados por la clase de contraste sola¹⁰⁷. Hace falta involucrar más elementos del contexto. Básicamente, la estrategia en el modelo E-C fue utilizar una clase de contraste, pero varios de los

¹⁰⁷ [Van Fraassen 1980], en su versión de 1996, Op. Cit., p. 161.

elementos contextuales quedaban sin representar. Como vimos, una clase de contraste es un conjunto de alternativas para el *explanandum*. Se supone que la clase de contraste tiene la función de ayudar a delimitar el contexto de la respuesta, esto es, delimitar la explicación. El contexto reuniría varios factores vagamente caracterizados: el tipo de relación de pertinencia, la información del tipo de suceso que es el *explanandum*, el interés de los involucrados en el acto de la explicación y la clase de contraste. No hay, así, para el modelo E-C, explicación independiente del contexto¹⁰⁸.

Van Fraassen parece confundir el problema de las asimetrías con el problema de los rechazos, específicamente, el problema de la elección causal. Cuando van Fraassen aduce ciertos elementos contextuales para solucionar las asimetrías, siempre pone ejemplos acerca de rechazos explicativos. Sólo en abstracto intenta mostrar cómo el contexto daría una solución a las asimetrías. En el mejor de los casos, van Fraassen muestra únicamente su confianza en que el contexto puede solucionar el problema de las asimetrías¹⁰⁹.

Respecto del intento por solucionar el problema particular de la elección causal, van Fraassen es más exitoso. Los elementos que según van Fraassen componen el contexto constituyen, al menos, claves que proporcionan información adecuada para delimitar una explicación. A pesar de ello los elementos adolecen de un mal común: son difícilmente precisables. Por tanto, capturarlos en un modelo formal de la explicación es sumamente difícil. Van Fraassen, en un intento por manejar más concretamente la función del contexto en su modelo, los simplifica en dos resultados: la determinación de la relación de pertinencia y la determinación de una clase de contraste. La idea que parece desprenderse de [van Fraassen 1980] es que el análisis del interés de las personas involucradas nos llevará a un tipo de relación de pertinencia y, a su vez, el análisis del tipo de suceso-*explanandum*, a una clase de contraste adecuada¹¹⁰.

¹⁰⁸ Cf., *Ibid.*, pp. 161-165.

¹⁰⁹ Cf., *Ibid.*, pp. 161-167.

¹¹⁰ Cf., *Ibid.*, pp. 161-163.

No obstante, ninguno de los dos elementos conceptuales resultantes del análisis de van Fraassen, caracterizados como él los presenta, son adecuados para capturar la adecuación contextual sugerida por el ejemplo del flogisto. En el contexto de una teoría como la de van Fraassen, en el marco de su noción de explicación contexto-dependiente, saber el interés causal de las personas que piden el *explanandum*, no me lleva directamente a saber el tipo de relación de pertinencia que estoy solicitando pues, pongamos por caso, hay contextos en los cuales no sabría si se trata de causalidad formal o eficiente. Saber que el *explanandum* es un evento fáctico no me ayuda en nada para saber la clase de contraste, justo porque en la teoría van fraasseana no me ayuda a saber el tipo de relación de pertinencia explicativa que se exige. Aún más, saber que se está pidiendo un tipo de relación causal, digamos, eficiente en el sentido de Aristóteles, junto con saber que el evento que constituye nuestro *explanandum* es un evento fáctico, no es suficiente para determinar una clase de contraste única y esto es casi obvio, básicamente por el hecho de que una lista como la clase de contraste no puede determinarse de manera unívoca, ¿Con qué criterio habremos de tener una única lista?

La relación de pertinencia en van Fraassen, si pudiera ser determinada de algún modo, no serviría de mucho pues la relación de pertinencia en el modelo E-C es completamente arbitraria. No hay ningún tipo de restricción para la relación de pertinencia. La crítica en **[Kitcher & Salmon 1987]** al modelo E-C está basada principalmente en este hecho. Ellos mostraron cómo la relación de pertinencia de van Fraassen, dado que no tiene ninguna restricción, produce la trivialización de su modelo: dado cualquier *explanandum*, cualquier factor arbitrariamente elegido y cualquier relación de pertinencia, hay un contexto que sitúa al factor mencionado como un factor explicativo de aquél *explanandum*¹¹¹.

¹¹¹ **[Kitcher & Salmon 1987]**, Kitcher, Phillip & Salmon Wesley C., “van Fraassen on Explanation”, en Ruben, David-Hillel (ed.), *Explanation* Oxford University Press, EE.UU., p. 314. Fuente Original: “van Fraassen on Explanation” en *Journal of Philosophy*, 84, 1987, pp. 315-30.

Dado que, como he dicho, van Fraassen mismo acepta la insuficiencia de la clase de contraste para solucionar el problema de las asimetrías, si se derriba su noción de relación de pertinencia (como hicieron Kitcher y Salmon) no hará falta además hacer una crítica sobre su clase de contraste. De cualquier forma veamos el siguiente caso para ilustrar que la clase de contraste tampoco ayuda mucho. Uno puede construir distintas preguntas con la misma clase de contraste. En la ilustración de van Fraassen un individuo hace construir una torre de 53m de altura porque dicha altura le recuerda un antiguo evento romántico. La pregunta inicial ¿Por qué la torre es de 53m. y no de 35m.? Puede significar ¿Por qué el individuo la construyó de 53m. y no de 35m.?, ¿Por qué sabes que la torre es de 53m. y no de 35m.?, ¿Por qué es verdad que mide 53m. y no 35m.? o ¿Por qué al construirla quedó de 53m. y no de 35m.? En todas estas preguntas la clase de contraste es la misma. Las preguntas no iniciales pueden verse, sin ningún problema, como intentos de especificación de la pregunta inicial. El punto interesante radica en que, modificando sólo la redacción de la clase de contraste en la pregunta inicial, es decir sin modificar la redacción del presupuesto, no parece posible modificar la pregunta inicial para que refleje el significado de alguna de las otras preguntas. Así, por lo menos entre este grupo de preguntas no hay clase de contraste que precise por sí sola la información que la pregunta inicial exige. La información adicional en cada pregunta no inicial, permanece sin capturarse a través de la clase de contraste. Por tanto, la diferencia entre los distintos *explananda* en cada pregunta no queda determinada por la clase de contraste. Como se implica por nuestro análisis de las asimetrías, no estando determinado el *explanandum* de que se trata, no podrá determinarse el tipo de relación de pertinencia que se exige para la explicación. Así, la clase de contraste no puede determinar la relación de pertinencia. Pero, como hemos visto antes, esto es aún peor, pues aún en el caso de tener determinado el *explanandum*, no podemos tener una relación de pertinencia determinada. Esto ocurre porque la relación de pertinencia explicativa no está determinada en el modelo de van Fraassen: elíjase cualquier evento *explanandum* y, de cualquier forma, no está clara la relación de pertinencia explicativa que uno debería

esperar. Así, incluso la precisión del *explanandum* no me permite desambiguar la explicación.

Lo útil de la pregunta de van Fraassen es más bien que puede traer consigo más datos acerca del *explanandum* precisando el presupuesto. Ninguna especificación contextual es útil si no es vertida sobre el *explanandum*, es decir, si no entendemos con precisión el tipo de evento *explanandum*. Y ninguna precisión del *explanandum* es útil si puede haber cualquier relación de pertinencia que dé cuenta de él. El ejemplo que el propio van Fraassen elabora no es inmune a esto. El ejemplo que él elabora sobre Adán, adolesce de esta insuficiencia de la clase de contraste. ¿Por qué Adán comió la manzana *en vez de devolverla a Eva?*, donde lo que va en cursivas es la clase de contraste, no soluciona el problema de los rechazos. Esto puede ser interpretado (con la misma clase de contraste) como tres preguntas distintas: ¿Por qué es verdad que Adán comió la manzana? O bien ¿Por qué crees realmente que Adán comió la manzana? O más interesantemente ¿Por qué Adán no pudo resistir la tentación de comerse la manzana? (esta última alternativa destaca porque se distingue de la primera solo por un aspecto intencional). Sin embargo, una vez aclarado el *explanandum*, y si podemos conectarlo no ambiguamente con alguna relación de pertinencia particular, la confusión contextual desaparece).

La precisión del *explanandum* es algo que depende del contexto, pero si nos comprometemos con una relación de pertinencia particular para cierta clase de eventos, el contexto no nos coloca en tanto problema. Cuánto debe precisarse el *explanandum* depende de la información del receptor, esto es inevitable aún para una visión semántica del *explanandum*, es decir, aún, para el *explanandum* visto como una proposición teórico-dependiente de una relación de derivación. Tal vez sea posible eliminar la imprecisión del *explanandum* si lo tomamos como conectado a una relación de pertinencia que ocurra entre eventos de cierta clase. Este primer paso parece necesario: restringir la relación de pertinencia que estamos buscando al clarificar una explicación, a una relación entre

eventos, como hicimos en nuestro análisis de las asimetrías. Y, como segundo paso, ligar no ambiguamente el tipo de evento con un tipo específico de relación de pertinencia. El evento *explanandum* siempre es inconfundible como entidad del mundo, su singularidad no es ambigua. Así, en estas nuevas condiciones, si el tipo de evento en el *explanandum* me determina la relación de pertinencia, una vez que sabemos el tipo de *explanandum*, sabemos la relación de pertinencia explicativa que cabe esperar. El problema aparecerá sólo cuando intentamos construir la explicación, no cuando queremos establecer una relación de pertinencia. Una relación de pertinencia así definida puede ser perfectamente objetiva e independiente del contexto, aunque no la conozcamos o no podamos describirla con claridad y precisión, es decir, aunque mantengamos una perspectiva epistémica de la explicación. La idea básica es que para determinar si un candidato a explicación es o no una explicación, hace falta, como hemos dicho, determinar el tipo de evento al que se refiere el *explanandum*, es decir, lo que se está inquiriendo con la pregunta-porqué. Esta determinación es claramente contextual. No parece haber escapatoria: hay que responder una explicación, antes de poder modelarla. Pero lo importante en este respecto es examinar si la relación de pertinencia es la mínimamente adecuada para el tipo de evento de que se trata en ese contexto. No se trata entonces de si es tal relación es la que realmente ocurrió o no, ni de si es o no la que se espera en el contexto para que cumpla con ser una buena explicación. Una propuesta así, clarifica la conexión entre una explicación y su relación de pertinencia. Además, nos coloca en una posición que no exige para la noción de explicación una adecuación como la mencionada por Scriven, la exigida por los contraejemplos de Achinstein o de Lewis, pero tampoco una adecuación contextual como la que asume van Fraassen.

Por supuesto, la pregunta puede seguir siendo ¿Cómo el modelo puede ser sensible al tipo de evento *explanandum* de que se trata? Pero la pregunta en este caso puede tener por lo menos dos sentidos bien diferentes: ¿Cómo el modelo puede desambiguar de qué tipo de evento se trata? o ¿Cómo el modelo puede capturar diferentes tipos de eventos? La desambiguación de la explicación, mediante la precisión

del *explanandum* no debe ser necesariamente parte del modelo de explicación. Una vez que sabemos lo que ha de ser el candidato a explicación, una vez que sabemos a qué pregunta debe responder, entonces usamos el modelo para clarificar por qué ha de considerarse o no una explicación en relación a cierta teoría o contexto. ¿Por qué entonces van Fraassen insiste en la determinación del *explanandum* vía modelo E-C? Quizá la respuesta es que intenta adecuar contextualmente las explicaciones. Dada su propuesta contextualista de la explicación, le parece que la adecuación debe venir del contexto, pero para ello le parece imprescindible que el modelo sea sensible a qué se refiere el *explanandum* con una precisión muy fina. Sin embargo, ¿Por qué ha de exigirse a un modelo de explicación que, por decirlo así, “perciba” el uso preciso de una preferencia lingüística?, ¿es necesaria esta característica para que un modelo de explicación sea útil? Yo creo que no. Un modelo puede tener otros fines, como la descripción de la relación entre los componentes de algo, una vez que lo hemos traducido a su estructura formal, es decir, una vez que sabemos sobre qué clase de evento se está pidiendo la explicación. Es así como uno dice en otros ámbitos, por ejemplo, que tal o cual argumento es bien capturado por el esquema de *Modus Ponens*. Una vez entendido como *Modus Ponens*, el argumento concreto capturado puede ser evaluado, y podemos ver las relaciones entre sus componentes de manera más clara. Podemos entender, por ejemplo, por qué creemos que ese argumento concreto es un argumento válido.

Una consecuencia lateral de este análisis que no debe pasarse por algo, y que es muy importante para continuar con los objetivos de esta última parte del capítulo, es que una relación de pertinencia completamente relativizada al contexto y una clase de contraste no son suficientes para precisar la noción de explicación, y que esto muestra de nuevo un subyacente ideal de buena explicación, ahora aplicado a una adecuación contextual. El caso de van Fraassen es un ejemplo de la dificultad para precisar elementos conceptuales apropiados, con los que pueda identificarse la adecuación contextual requerida para la solución del problema de la elección causal, pero también para el problema de la sobredeterminación causal. Considero que el intento de van Fraassen es

un buen esfuerzo por elucidar el problema. Y su fracaso es una razón adicional para abandonar lo más que sea posible ideales de buena explicación de esta clase. Así también, la adecuación contextual no debería ser puesta como requisito para la elucidación de la noción de explicación. Esto implica pensar la adecuación contextual como parte de los fundamentos de la explicación y, así, como parte del ideal de buena explicación que hemos caracterizado. En todo caso, lo que será de mi interés en esta investigación es intentar escudriñar una parte necesaria de toda explicación mínima.

Esto contribuye finalmente en favor de la posición que yo adopto tanto para el problema de la sobredeterminación epistémica causal como para el problema de la elección causal. Desde mi perspectiva, si se deja abierta la posibilidad de varios factores causales para un mismo *explanandum* entonces tendremos que estar dispuestos a aceptar, como explicaciones, construcciones que contengan *explanans* que aparezcan como extraños en el contexto en el que se pide la explicación. Análogamente, tampoco importará para el dictamen de que algo sea una explicación, la cuestión de si los eventos involucrados en la explicación realmente estuvieron relacionados por esa relación de pertinencia o no. Simplificadamente, algo puede ser una explicación, aún en el caso de tener un *explanans* que sea extraño al contexto o que no describa la relación real acaecida entre los eventos involucrados.

Obsérvese que esta forma de afrontar ambos problemas, tiene implicaciones importantes sobre la forma de concebir la explicación. Una explicación no necesita corresponder a alguna realidad extraepistémica. La explicación en este sentido es claramente concebida como una actividad intelectual falible, una actividad epistémica. Es conveniente, entonces, distinguir algunos conceptos: la **relación de pertinencia explicativa**, y la **relación de explicación**.

Entenderé por "relación de pertinencia explicativa" una relación que mantienen dos eventos. Cuando estos eventos son fácticos la relación tendrá que ser causal, cuando

son no fácticos, entendidos como la verdad de una proposición, la relación deberá ser inferencial. Podemos decir que alguien ha propuesto que existe una relación de pertinencia explicativa entre dos eventos independientemente de si dicha relación de hecho ocurre entre ambos eventos. La relación de explicación, en cambio, no es una relación entre eventos sino entre un conjunto de proposiciones llamado "*explanans*" y un evento, que llamamos *explanandum*. Así, el acto de explicar un evento es la postulación de una relación de explicación entre un conjunto de proposiciones y un evento a explicar, implica una relación entre un objeto abstracto y un objeto del mundo fáctico o del mundo semántico. Los *relata* de tal relación son un conjunto de proposiciones y un evento. Cuando alguien ofrece una relación de explicación se encuentra postulando además una relación de pertinencia explicativa *mínimamente adecuada* (es decir, causal o inferencial) entre, por un lado, los eventos referidos en el conjunto de proposiciones y, por otro, el evento a explicar (*explanandum*). La relación de explicación puede cumplirse aún a pesar de que la relación de pertinencia explicativa postulada no se cumpla realmente, pues puede proponer relaciones de pertinencia mínimamente adecuadas y esto puede no corresponder con la realidad o con las exigencias de cualquier contexto. Estas distinciones nos encaminarán a estructurar una noción epistémica y pragmática mínima para la explicación científica que pueda articular una nueva forma de abordar los problemas que hemos revisado.

2. 1. 5. 2 La Noción de Relación Causal

La distinción entre relación de explicación y relación de pertinencia explicativa facilita una caracterización epistémica de la explicación. La relación de pertinencia explicativa se concibe así como una relación real en el mundo, mientras que la relación de explicación consiste en la postulación de una relación de pertinencia explicativa mínimamente adecuada. En el caso de eventos fácticos la relación de pertinencia mínimamente adecuada es la relación causal; en el caso de eventos no fácticos, la relación inferencial. No obstante el carácter epistémico de la explicación, la posición que estoy

proponiendo intenta escapar lo más posible de dependencias contextuales. Aunque, sin duda, será imposible escapar del todo.

La situación puede describirse como sigue. Por un lado, como he apuntado antes, debemos intentar revisar los efectos que tiene el contexto en la explicación, como parte de la caracterización de una perspectiva epistémica de ésta; por otro, queremos evitar la relativización que puede producir una dependencia del contexto *a la* van Fraassen. La estrategia que seguiré para salir del problema será la siguiente: intentaré mantener una idea de explicación mínimamente adecuada manteniendo invariables del contexto las relaciones de pertinencia explicativa causal e inferencial; al mismo tiempo, intentaré mantener las consideraciones para construir una buena explicación, definidas como he descrito (*a la* Scriven) como dependientes contextualmente. La dificultad de este intento se traducirá en mantener la explicación en el nivel epistémico mientras que las relaciones de pertinencia explicativa mínimamente adecuadas se mantienen invariables de contexto a contexto. Esto tendrá algunas implicaciones importantes sobre la manera en que usaré la noción de relación causal en la explicación.

La forma en que suele hablarse de la causalidad en los textos sobre explicación presupone, en general, que no hay duda acerca de lo que se quiere decir con "relación causal" o con "factores causales". Salmon es uno de los pocos que intentaron elaborar una noción de causalidad especialmente para elucidar la explicación. Para Salmon no es posible elucidar la explicación si no se tiene una idea clara de lo que es una relación causal. Pero ésta no parece ser la opinión general. La mayoría de los filósofos clásicos que han escrito acerca de la explicación, han obviado el tema de la causalidad aunque prácticamente ninguno ha justificado esta posición metodológica.

Una forma de elucidar la explicación es proponer una metaexplicación-cómo de la explicación. En este sentido, elucidar la explicación es hacer una propuesta acerca de cómo funciona la explicación y principalmente saber cuál es su mecanismo interno. Esto

puede ofrecer, en mi caso, una justificación metodológica para no profundizar en el problema de precisar la noción de relación causal, y es parte de mi posición pragmática sobre el tema.

Tal como he construido esta perspectiva, la explicación es una relación entre proposiciones y eventos que postula una relación de pertinencia entre eventos de dos clases. La existencia real de dicha relación de pertinencia es independiente de la calidad de ser explicación del candidato a explicación que la postula. En el caso de la relación causal tanto como en el de la relación inferencial, lo que juega un papel en la estructura de la explicación es *la propuesta* de que exista una relación tal entre dos eventos determinados y no la existencia de la relación. De modo que uno estaría tentado a prescindir de una noción precisa de relación causal y dejar que su precisión dependa del contexto en el que se plantea cada explicación. Bastaría con precisar lo que entendemos por "postular una relación causal entre dos eventos". No obstante, una precisión subordinada al contexto trivializaría la noción de explicación en el mismo sentido en el que se trivializa la noción de explicación de van Fraassen: para cualesquiera dos eventos y una relación entre ellos, habría un contexto que los hace encajar en una explicación. Una relativización de este tipo entre eventos no fácticos, con base en la relación de pertinencia lógica, tendría menos problemas de aceptación. Si el contexto incluye la estipulación de ciertas formas de inferencia no parece haber ningún problema en aceptar como explicaciones de eventos no fácticos las que cumplan con alguna relación de inferencia estipulada en dicho contexto, esta es la forma en que podemos no hacer que la relación de inferencia sea *post-hoc* o *ad-hoc*. Mi posición para el caso de eventos no fácticos, será muy parecida a ésta. Para el caso de eventos fácticos, en cambio, puede haber mayores dificultades: podría postularse una relación causal entre, por ejemplo, la erupción de un volcán en México y el bostezo de un pingüino en la Antártida. Esto no parece nada intuitivo. Lamentablemente, hasta donde he presentado mi propuesta, las consecuencias podrían ser peores que las de la relativización contextual que he expuesto. Si la explicación de un evento fáctico no depende del ideal de buena explicación que hemos

caracterizado y, así, tampoco de si corresponde con los hechos reales o con contextos, entonces con cualquier relación causal postulada tendremos una relación de explicación. En el supuesto de que la explicación es epistémica, toda postulación de una relación causal entre cualquier par de eventos es válida como relación de pertinencia explicativa no sólo para un contexto determinado sino para cualquier contexto. La calidad de ser explicación depende sólo de algunas propiedades de *la postulación* de la relación de pertinencia. Dichas propiedades no tienen que ver con alguna correspondencia con relaciones realmente mantenidas entre los eventos involucrados. Por lo tanto, las relaciones de pertinencia postuladas en una explicación, aún desde esta perspectiva epistémica, deberían tener alguna restricción si no queremos que cualquier postulación de relaciones entre eventos sea catalogada como explicación. Como vemos, mi propuesta de que las únicas relaciones de pertinencia explicativa sean la relación causal y la relación inferencial, no soluciona del todo el problema.

En las siguientes líneas examinaré un poco más de cerca el problema de la precisión de las relaciones de pertinencia, luego intentaré precisar, mediante un tipo de adecuación teórica o contextual, el uso que haré de la noción de relación causal en la explicación y, finalmente, sugeriré una forma más estrecha de concebir la relación de causalidad.

Vayamos al extremo de la situación. Un ejemplo más general que muestra drásticamente las consecuencias de proponer relaciones de pertinencia mínimamente adecuadas pero completamente irrestrictas, se encuentra relacionado con el ejemplo clásico del brujo. El brujo hechiza a alguien para que siempre que ponga sal en el agua, ésta se disuelva. De manera que siempre que esa persona coloca la sal en agua, la sal se disuelve y entonces en algún sentido el encantamiento de la sal explica su disolución. El ejemplo del brujo aparece en [Kyburg 1965]¹¹² y generalmente se ha interpretado como un caso de contrapertinencia explicativa, esto es, se le interpreta en relación con el

¹¹² [Kyburg 1965] Kyburg, Henry E. Jr., "Comment" en *Philosophy of Science*, 32, pp. 147-51.

problema del supuesto argumental (SA). Sin embargo, lo importante del ejemplo respecto de la discusión de este apartado es que puede reformularse para que muestre lo siguiente: si se considera a la explicación como dependiente solamente de la postulación de ciertas relaciones de pertinencia, aunque sean mínimamente adecuadas, pero sin necesidad de ninguna restricción, entonces podríamos aceptar que los miembros de cualquier par de eventos estén vinculados por dichas relaciones de pertinencia. El ejemplo de Kyburg puede reformularse de la siguiente manera: supongamos que pensáramos equivocadamente que un brujo muy poderoso encantó toda la sal que existió, existe y existirá. El embrujo consistiría en que cada vez que la sal estuviera en el agua, ésta se disolvería. Supongamos también que no tuviéramos acceso a la *causa verdadera* de la disolución de la sal. Así, la causa verdadera quedaría oculta por una causa postulada, a saber, el encantamiento, y esto es independiente de que mantengamos nuestra idea de causa independiente del contexto. El ejemplo es importante pues sugiere lo gravemente oculta que puede quedar la relación de pertinencia real, por relaciones de pertinencia postuladas. Y a través de ello muestra también la ficción en que puede convertirse nuestra idea de explicación, a pesar de haber establecido relaciones de pertinencia mínimamente adecuadas, pues podría mantenerse una relación de causalidad para el caso del encantamiento.

El ejemplo reformulado exige mayores restricciones para la postulación de relaciones de pertinencia entre eventos. Sin embargo, si no deseamos comprometer la noción de explicación con el ideal de buena explicación, entonces no podemos suponer que la restricción consiste en exigir que las causas deben ser reales para ser explicativas. No tendremos entonces este recurso. Una salida es que las restricciones provengan del contexto. Como he dejado claro, la restricción usual a partir del contexto consiste en relativizar la relación de pertinencia a cada contexto particular. Lamentablemente ello significa que, por ejemplo, si en un contexto determinado la relación de contigüidad entre dos eventos es válida como relación de pertinencia explicativa entre ellos en ese contexto, la contigüidad explicaría. Esta es la clase de adecuación contextual por la que van Fraassen

opta. No obstante, no estamos seguros de que esto pueda efectivamente sernos de alguna utilidad. El problema del cual partimos es, precisamente, cómo lograr que no cualquier relación pueda jugar el rol de una relación de pertinencia explicativa. Contextualizar como van Fraassen la relación de pertinencia explicativa es asegurar la intuitividad de la relación de pertinencia explicativa en cada contexto determinado, a costa de desvanecer la noción intuitiva de explicación. Otra versión de la misma estrategia es intentar relativizar, más específicamente, la relación causal a cada contexto. Pero esto es equivalente a usar el término "relación causal" para designar todo tipo de relaciones (para cualquier relación siempre se puede encontrar algún contexto, por muy raro que éste sea, en el que la relación sería considerada como una relación causal), lo cual daría como resultado exactamente las mismas consecuencias indeseables que la relativización contextual de la relación de pertinencia. De cualquier forma, se hace necesario integrar algún tipo de restricción.

De manera que podemos plantear esta dificultad como un nuevo dilema: O bien restringimos las relaciones de pertinencia explicativa desde alguna correspondencia con relaciones reales o bien las restringimos mediante el contexto. Llamaré a cada opción "**restricción externa**" y "**restricción interna**", respectivamente. Una explicación que no satisficiera las condiciones impuestas por el contexto, las restricciones internas, sería llamada por Scriven, como vimos, una "explicación irrelevante", y también coloca esta categoría como un aspecto de los fundamentos de la explicación, no de la noción de explicación. Si optamos por una restricción externa entonces nos comprometemos con ello a que la explicación sea adecuada en el sentido de Scriven. Si optamos por la restricción interna, trivializaremos la noción de explicación. Una forma adecuada para restringir las relaciones de pertinencia tendrá que superar este dilema. Creo que la clave está en la forma de entender la relativización de la relación de pertinencia. Intentaré construir una solución sobre la idea del segundo cuerno del dilema.

El problema de una relativización *a la* van Fraassen es que produce el efecto de que la relación de pertinencia puede ser cualquier cosa. En ese caso la relativización transforma, de contexto a contexto, el tipo de relación de pertinencia explicativa exigida. Si pudiéramos concebir una relativización que no permitiera cambiar el tipo de relación de pertinencia explicativa, podríamos evitar la trivialización de la noción de explicación.

Antes de construir una propuesta encaminada sobre la sugerencia del párrafo anterior, haré algunas breves aclaraciones acerca de la distinción entre la adecuación contextual que quiero proponer y la adecuación contextual en el sentido de van Fraassen.

Intentaré precisar un poco más la idea de van Fraassen. La concepción de contexto en van Fraassen está pensada en términos no proposicionales. Al parecer, para él los elementos del contexto de una pregunta son los elementos pragmáticos de la situación en la que es exigida una explicación; y los elementos contextuales de una respuesta son los elementos pragmáticos de la situación en que es dada la respuesta. Un intento de precisar esta idea puede hacerse apelando a la distinción entre semántica y sintáxis de un lenguaje. Así, podemos interpretar "pragmático" respecto de la explicación como todo aquello no sintáctico que se encuentre modificando la semántica del enunciado que pregunta y todo aquello no sintáctico que se encuentre modificando la semántica del enunciado que contesta y que se considera la explicación. Como hemos visto, para van Fraassen toda respuesta debe ser una buena respuesta. La bondad de la respuesta consiste en una especie de adecuación entre ambos contextos, el de la pregunta y el de la respuesta, en el sentido de que exista cierta coherencia entre el significado resultante de cada una de las dos (la pregunta y la respuesta). Esto es la adecuación contextual vanfrasseana. Pero, como he dejado claro antes, resulta difícil encontrar elementos conceptuales con los que podamos identificar claramente esta adecuación contextual.

Mi concepción de teoría intenta capturar lo proposicionalmente capturable de la noción de contexto, como interpreto que manejó a ésta última van Fraassen. Es el aspecto

proposicional del contexto. Concibo una teoría, provisionalmente, como un conjunto de proposiciones que corresponden a un conjunto de afirmaciones explícitas o presupuestas. Parentéticamente podemos suponer que, entre los presupuestos de una teoría, se hallan algunos tipos de inferencia permitidos. Así, podemos relativizar la relación de pertinencia explicativa entre eventos no fácticos a cada teoría particular. Para el caso de eventos fácticos, los cuales son el motivo de la presente discusión, la solución no será tan simple.

Una teoría puede construirse a partir de una o varias teorías científicas. Pero también puede construirse a partir de un contexto concebido como lo hace van Fraassen. En uno y otro caso los elementos de una teoría pueden ser completamente falsos o significar cosas extrañas a las verdades supuestas por la Ciencia o la Filosofía. Pueden corresponder a afirmaciones míticas, místicas o mágicas tanto como a afirmaciones de sentido común, científicas o filosóficas. De este modo, una adecuación teórica en mi sentido no es lo mismo que una adecuación contextual en el sentido que le atribuyo a van Fraassen. Sin embargo, en estos mismos sentidos, la adecuación teórica puede mostrar parcialmente en términos proposicionales un tipo de adecuación contextual. La adecuación teórica que propongo mostrará un tipo de adecuación contextual que es distinto del de la adecuación contextual vanfrasseana, un tipo de relevancia de la explicación en palabras de Scriven. Notemos sin embargo, que, a fin de cuentas, se trata también de una adecuación contextual.

Veamos ahora algunos aspectos formales relacionados con la posibilidad de concebir una relativización que nos permitiera por lo menos parcialmente mantener contexto-independiente el tipo de relación de pertinencia explicativa. Lógicamente hablando, una relación puede definirse como un conjunto de pares ordenados. Cada par ordenado consta de dos objetos que cumplen con la relación. La idea es que si capturamos el conjunto de pares en los que se cumple la relación entonces hemos capturado la extensión de la relación. Al conjunto de pares ordenados puede calificársele como una definición extensional de la relación en cuestión. En este sentido, un cambio en

el significado de la relación puede implicar un cambio en el conjunto de pares ordenados que la definen extensionalmente. Así, típicamente se entiende que dos nociones pueden no ser las mismas intensionalmente, a pesar de que tengan la misma extensión. La inversa, en cambio, se entiende como válida: no se puede cambiar la extensión sin cambiar la intensión.

Van Fraassen sugiere, al comienzo de su estudio sobre la explicación, en **[Van Fraassen 1980]**, que para entenderla será necesario penetrar en su dimensión intensional¹¹³. Notablemente, en **[Orayen 1976]**, también se hace una sugerencia similar. Observemos desde un inicio que una explicación para van Fraassen es una respuesta pero en un sentido intensional. Con ello él intenta decir que se trata de una noción que es contextualmente dependiente. No obstante van Fraassen no da ninguna idea precisa de lo que intenta decir cuando habla de la intensionalidad de la explicación¹¹⁴. Aunque su noción de intensionalidad puede fácilmente interpretarse como lo he descrito en el párrafo anterior, van Fraassen parece querer referirse a algo adicional que tiene que ver con el contexto. En el punto donde uno espera que dé alguna caracterización de lo intensional en la explicación, van Fraassen procede a relativizar contextualmente la relación de pertinencia. Empero, van Fraassen pone algunos ejemplos que muestran que en algunos casos es contraintuitivo pensar que dos nociones son equivalentes, a pesar de ser extensionalmente equivalentes. A propósito de las asimetrías, van Fraassen intenta justificar con estas disquisiciones la relativización contextual de su idea de explicación. Pero no parece quedar muy clara la línea que lo lleva desde la intuición de lo intensional en la explicación hasta la relativización contextual de su noción de relación de pertinencia.

De cualquier forma, las ideas de van Fraassen pueden servir de preludeo a una noción de relación de pertinencia explicativa definida intensionalmente. Si logramos escapar del principio de que dos nociones son distintas si ocurre que los conjuntos de los

¹¹³ Cf., **[van Fraassen, 1996]** (1980), Op. Cit., p. 186-7.

¹¹⁴ Aunque el estudio de Orayen es mucho más analítico, tampoco avanza con precisión sobre esta idea. Pero deja muy claro, con una prueba deductiva, la dificultad de una modelación deductiva de cierta clase de explicaciones.

pares ordenados que las definen extensionalmente son distintos, estaremos en una mejor posición para enfrentar el dilema recientemente planteado.

No es posible abandonar el principio formulado en el párrafo anterior pero sí es posible, en algún sentido, escapar a sus limitaciones. Esto es así si pensamos en la relación de pertinencia como una relación que se aplica a un grupo de teorías. Tomemos el caso de la relación causal. Supongamos que definimos una relación causal sin mencionar ningún miembro del conjunto de pares que cumplen con la relación causal, es decir, definimos "relación causal" no extensionalmente. Postulemos que dicha definición de relación causal permanece constante en toda teoría o contexto. En cualquier teoría o contexto, para cualesquiera dos eventos que mantienen la relación causal, dichos eventos cumplen con lo estipulado en la definición intensional supuesta. Si la relación de pertinencia es mantenida inmutable en todo contexto, entendiendo por contexto un contexto teórico como hemos caracterizado, parece difícil hablar de una relativización de la relación causal a distintos contextos. Sin embargo podemos concebir que *la aplicación* de la misma relación causal sea distinta en cada teoría o contexto. La forma de hacerlo es permitiendo que difieran en algunas de las teorías-contextos, los conjuntos de pares ordenados que cumplen con la relación causal, según la teoría que se tome en cuenta. Esto supone de nuevo una concepción epistémica de la relación de explicación. Bajo dos teorías-contexto distintas T1 y T2, podemos tener la misma idea intensional de relación causal R pero proponer que dicha relación ocurre, en cada teoría, entre distintos conjuntos de pares de eventos C1 y C2. Así, podemos concebir una relativización contextual que no afecte esencialmente el significado de nuestra relación de pertinencia explicativa.

Permítaseme hacer conjeturas realistas para explicarlo. Podemos suponer que la interpretación típica de la relación entre intensionalidad y extensionalidad permanece válida en la realidad: si son distintas las extensiones de la relación causal en cada teoría debe ser distinto su sentido intensional. Ambos sentidos de causalidad R, en las diferentes teorías-contexto, no pueden ser verdaderos. Sus principios de causalidad son diferentes.

Algo debe estar equivocado en alguna de las dos teorías o en ambas pues, concedamos, realmente a R debe corresponderle un único conjunto C de pares ordenados que cumplen con ella. A pesar de esta intuición realista, las cosas no son necesariamente así desde su tratamiento a nivel epistémico y pragmático. El que tengamos esta idea del Principio de Causalidad no impide que cada teoría postule, sin acuerdo obligado entre las mismas, conjuntos de pares de eventos entre los que se presume que R se cumple. De cualquier forma, en general, lo que he aseverado antes de estas conjeturas es compatible también con una posición no realista. Podríamos decir que, en vista de no haber criterio para saber si T1 o T2 están realísticamente en lo cierto al postular sus respectivos C1 y C2, la intensionalidad de R puede mantenerse sin ningún problema igual en ambas teorías a pesar de que postulan, cada una de ellas, conjuntos distintos de pares de eventos entre los que se presume que R se cumple. Tales conjuntos siempre pueden diferir pues son *postulaciones* o *conjeturas* acerca de R. Por lo demás R, sin embargo, permanece intensionalmente invariable y esto mantiene un poste indicador para la identificación de explicaciones a pesar de mantener, al mismo tiempo, una posición más bien pragmática de lo que entendemos por explicación mínima.

Con este procedimiento nos encontramos adecuando la extensión de la noción de relación causal, relativizándola a una teoría-contexto particular. Es decir, haciéndola contexto-dependiente. Mientras tanto, mantenemos la intensión de la relación causal invariable en toda teoría-contexto. Tenemos entonces dos dimensiones de la relación causal: la epistémica y la real. Podemos incluso mantener que una relación de pertinencia explicativa se mantiene siempre en la dimensión de lo real. A pesar de ello, la idea que tengamos acerca de los eventos que cumplen con la relación causal puede diferir de una a otra perspectiva teórica. Es decir, pueden ser epistémicamente distintas. La relación causal, concebida de esta manera, no hace trivial nuestra noción de explicación. Por una parte, pueden aceptarse las consecuencias de una concepción epistémica de la explicación, como el hecho de que sean aceptables distintas postulaciones de relaciones de pertinencia explicativa, sin requerir de correspondencia con hechos reales ni

adecuación contextual. Por otra parte, no toda relación de pertinencia explicativa es aceptable. Específicamente, las explicaciones de eventos fácticos tendrían que cumplir con postular relaciones causales definidas intensionalmente de manera unívoca.

El tratamiento anterior para la relación causal entre dos eventos permite una clase de adecuación contextual a la par que evita la trivialización de la noción de explicación. Aunque la aplicación de la relación causal puede diferir de teoría a teoría, es la misma relación en toda teoría. El ejemplo del brujo de Kyburg puede entenderse mejor ahora: desde esta perspectiva, en un contexto científico no es aceptable la postulación de una relación causal entre el encantamiento y la disolución de la sal. Pero en un contexto diferente, digamos, en una cultura que contemple ritos como factores causales adecuados, puede mantenerse la idea de que existe una relación causal entre el encantamiento y la disolución de la sal. Podemos proyectar parcialmente cada contexto en una teoría. Así, tendríamos dos teorías: la científica y la mágica. Cada teoría-contexto acepta que la relación causal se da en ciertos pares ordenados y esto es lo que puede distinguirla de otras teorías-contextos.

El uso de la noción de relación causal en mi noción mínima de explicación sigue estando representado por la postulación de la relación en lugar de la relación misma. Y la postulación de una relación de pertinencia entre dos eventos se encuentra relativizada a una teoría-contexto particular. Sólo si la teoría-contexto *acepta* dicha relación entre ambos eventos entonces estamos ante una explicación para ese contexto. Toda explicación debe cumplir con este tipo de adecuación contextual. No obstante, la relación de pertinencia exigida para toda explicación en toda teoría o contexto es, en el caso de eventos fácticos, la relación causal, y esto nos permite pasar de la desambiguación del *explanandum* a la determinación del tipo de relación de pertinencia explicativa exigida. Y, de este modo, la relación de pertinencia no está determinada sólo por factores subjetivos ni sólo factores extraepistémicos. Consecuentemente, este procedimiento nos permite tener un criterio de evaluación relativo a un contexto, evaluar así un tipo de relevancia

(como diría Scriven) de la explicación, de modo que a partir de un argumento, pueda determinarse si es o no una explicación del *explanandum* precisado.

Como ya he dicho, entiendo la elucidación de la explicación en el sentido de clarificar su mecanismo interno (explicación-cómo). Así, lo importante en dicha elucidación es intentar identificar mejor los elementos que juegan un papel importante en una explicación. Y no sólo esto. También es necesario clarificar las relaciones entre dichos elementos internos y observar si hay otros elementos externos que se encuentran vinculados con ellos. Bajo este sentido de lo que es una elucidación, es posible dejar al margen un análisis profundo de la noción de relación causal. Si lográramos ponernos de acuerdo en una noción de causalidad adecuada para tener el rol de relación de pertinencia explicativa, es probable que esa noción satisficiera nuestra elucidación de la explicación. Sólo habría que cuidar que la resultante noción de relación causal no trivializara nuestra noción de explicación. Para hacer un análisis de la relación explicativa no es necesario involucrarse de lleno en la discusión sobre el problema de la elucidación de la noción de relación causal.

Uno puede, sin embargo, desear dar un sentido más que formal a su modelo de explicación científica, para avanzar hacia una noción mínima de explicación que aporte una idea no completamente formal sino ya por lo menos parcialmente interpretada de ésta. Y esto no puede hacerse prescindiendo de una definición, aun cuando fuera provisional, de dicha relación de pertinencia. En estas últimas líneas del capítulo propondré una noción de relación causal para que la idea de explicación tenga mayor sentido en la estructura de los elementos constituyentes del modelo formal al que tienden todas estas reflexiones.

Desde luego, la noción intensional de relación causal que queremos que se tome en cuenta para la explicación debería intentar comprender una gama amplia de concepciones acerca de la causalidad. Debería, por ejemplo, incluir la concepción

indeterminista de la causalidad tanto como su concepción determinista. Esto, en particular, no es difícil, pues es posible concebir sin inconsistencia que algunos eventos reales del mundo mantengan relaciones causales deterministas mientras que otros, relaciones causales indeterministas.

De cualquier modo, elegir una noción de relación causal para que funcione como relación de pertinencia en la explicación, es complicado. Hay muchos tipos de causalidad y algunos de ellos quedarán fuera del ámbito de la explicación. De acuerdo a la idea de explicación que he estado intentando caracterizar, la elección de uno de estos tipos de causalidad o de algún compuesto entre ellos marcará un sesgo que dividirá entre las relaciones que sustentan una relación de explicación y las que no. Así, podría ser que muchas relaciones causales no fueran consideradas como jugando un rol explicativo. Pero debemos tomar esto como un ejercicio reflexivo en el camino de hacer una investigación de mayores profundidades sobre el tema, un tímido avance hacia la continuación de una investigación que concluya tan solo con un modelo formal de la explicación.

Siguiendo con nuestro avance, otra opción es suponer que cualquier relación de tipo causal entre eventos constituye una relación de pertinencia explicativa. Es importante enfatizar la distinción entre esta opción y una relativización contextual de la noción de relación causal *à la* van Fraassen. Hacer que la relación de pertinencia explicativa entre eventos fácticos sea cualquier tipo de relación causal es algo predeterminado para su aplicación en toda teoría y, por ende, para todo contexto. Cosa muy distinta es relativizar la noción de relación causal a cada contexto y no es a esto a lo que quiero referirme. La opción consiste en definir "relación causal", de antemano y para toda teoría y contexto de forma que abarque todo tipo de relación causal entre eventos. Esta opción es suficientemente moderada para dejar contentos a algunos estudiosos de la explicación. No obstante, parece haber tantos tipos de causalidad que puede correrse el riesgo de trivializar de nuevo la noción de explicación. De cualquier forma en este caso la trivialización no es ya una consecuencia necesaria.

A pesar del riesgo de trivialización, dejaré abierta la posibilidad de interpretar la relación causal de mi noción de explicación en este sentido amplio y moderado que he descrito.

Sin embargo, como dije antes, yo me siento inclinado a pensar que la relación causal más adecuada a las explicaciones de eventos fácticos es la relación de causa eficiente de Aristóteles y, también, la relación de fundamentación ontológica. Aunque la causalidad eficiente aristotélica tendría que sufrir al menos tres modificaciones. Por un lado, me refiero a la causalidad eficiente considerada como una relación que vincule sucesos y no objetos. Por otro lado, una relación de causalidad eficiente tendría que concebirse en términos no temporales, es decir, privilegiando la cualidad productiva de la relación independientemente de si la causa precede temporalmente al efecto. Finalmente, una relación de causalidad no tiene por qué pensarse como una relación determinista.

Apostar por esta relación causal es aventurado, pero creo que esta relación de causalidad es una relación que coincide muy bien con el sentido en el que exigimos contemporáneamente una explicación de un evento fáctico. Un análisis del sentido en el que parece entenderse la noción de relación causal en los textos sobre explicación puede indicar razones inductivas para considerar que, al menos intuitivamente, esta noción de causalidad es el sentido por excelencia que se supone en las disquisiciones sobre el tema. Por otro lado, en el ámbito científico, me parece que muchas explicaciones causales son entendidas también en el sentido que he sugerido.

No obstante, no intentaré defender más profundamente en esta investigación la idea de que la única relación de causalidad en la explicación de eventos fácticos sea este tipo de relación causal. Me contentaré sólo con la idea, más general, de que entre eventos fácticos la relación de pertinencia es necesariamente causal. Esto ya constituye una

conclusión no muy fácil de defender. Mi argumentación a favor no irá más allá de los párrafos que he escrito en secciones anteriores del capítulo y de las ventajas y la adecuación del modelo, las cuales serán presentadas en el capítulo IV.

Por último, acerca de la noción de relación causal, aún a pesar de que para muchos filósofos la relación causal juega un papel importante en la explicación, en los modelos de explicación generalmente no queda muy clara la forma de representar esta relación. A menudo su representación es completamente omitida. Acudir a las propuestas de asimetría es un buen recurso para mostrar esto. Recordemos el ejemplo CR de asimetría. En la formulación del caso CR-P parece proponerse una conexión causal entre los eventos involucrados. Sin embargo, tal conexión no se encuentra ahí, en la formulación del ejemplo. El establecimiento de la relación causal es dejado al contexto. Si la relación causal juega un papel importante en el modelo, puede ser que pensemos que debería tener alguna representación en el modelo. De lo contrario no tendríamos cómo resolver el problema de las asimetrías desde el modelo. Si fuésemos estrictos, la formulación CR-P y la formulación CR-I no proponen relación causal alguna, son meras derivaciones deductivas. El conocimiento de que CR-P contenga la propuesta de una relación causal queda completamente vago. Lo mismo se aplica en el caso de argumentos explicativos de tipo estadístico. Lo que esto muestra es que normalmente se deja la caracterización de la relación de pertinencia explicativa como parte de la formulación de las condiciones adicionales al modelo formal de explicación. Se piensa que la base de un modelo de explicación puede ser una estructura formal básica con algunas condiciones adicionales para la inferencia.

Sin embargo, la forma en que he propuesto la interacción entre los candidatos a explicación y el contexto en las líneas anteriores, puede exigir por lo menos una representación mayor de la relación de pertinencia explicativa, en el nivel formal. Si pretendemos que el modelo represente una interacción entre explicaciones y contexto, a través de la relación de pertinencia explicativa, necesitaremos un modo en que el modelo

reaccione ante la confirmación de que los eventos propuestos por un conjunto de proposiciones dan efectivamente cuenta, relativamente a un contexto dado, del *explanandum*. Si aquellos eventos no dan cuenta de éste, entonces el argumento constituido por ese conjunto de proposiciones no es en realidad un *explanans* para ese *explanandum* en esa teoría-contexto. Esta forma de evaluar sería útil para notar las relaciones candidato-a-explicación/contexto. De modo que una forma de no monotonicidad podría también examinarse a partir de si un aumento de información modifica los pares que sean considerados *relatas* para la relación de pertinencia explicativa en un contexto dado. Si estos pares cambiaran, algunos pares conjuntos-de-proposiciones/*explanandum*, podrían dejar de ser considerados como explicaciones para ese contexto. Así, el aumento de información puede modificar pares de cosas que guardaban una relación de explicación, a partir de que las relaciones de pertinencia explicativa que se han propuesto por el candidato a *explanans*, ya no funcionan para el par de eventos asumido.

Así, en este capítulo hemos apuntado hacia una representación más dinámica de la explicación, como un modo de abordar los problemas presentados en Filosofía de la Ciencia a los modelos clásicos de explicación. Esta representación, hemos dicho, puede mantener el supuesto argumental (SA), pues resulta adecuado para representar intuiciones básicas de la explicación científica. En particular, hasta ahora, he justificado la importancia de que el mecanismo formal argumental que modele la explicación permita baja probabilidad y no monotonicidad (análisis del SA y de la ambigüedad epistémica), pueda representar dos clases de relaciones de pertinencia explicativa (examen de las asimetrías) e interactúe con contextos sin que se vuelva trivial la propia noción de explicación (reflexiones sobre la relación de pertinencia explicativa y sobre la ambigüedad epistémica). Esto nos ha llevado a considerar como una posible estructura básica del modelo, una relación de inferencia por *default*. A esta relación de inferencia sería importante enriquecerla con capacidades para la representación de diferentes *explananda* y diferentes relaciones de pertinencia explicativa y, también, con su interacción, mediante

una forma de entender estas relaciones, con un contexto dado. Todo ello bajo una perspectiva argumental, epistémica y pragmática de explicación mínima que nos lleve a incorporar las consecuencias de los contraejemplos a los modelos clásicos. En el siguiente capítulo exploraremos más las conexiones que sería importante hacer notar en el modelo, entre explicación y contexto.

CAPÍTULO 3

EL CONTEXTO Y LA MODELACIÓN INFERENCIAL DE LA EXPLICACIÓN CIENTÍFICA: PARACONSISTENCIA Y NO MONOTONICIDAD

Introducción

Una de las conclusiones del capítulo anterior fue que el contexto y su interacción con la explicación es un factor imprescindible para la modelación de ésta, dados los problemas clásicos presentados. En este capítulo nos centraremos en dos problemas centrales de esta interacción: los contextos inconsistentes y el cambio en la información del contexto en el que es propuesta la explicación. Cada problema nos llevará a dos propiedades importantes para un modelo argumental de la explicación: la paraconsistencia y la no monotonicidad.

3.1. Explicación en Contextos Inconsistentes

3.1.1. La Explicación de Frege.

La explicación es un elemento dinámico de la actividad científica en el sentido de que algunos de sus constituyentes son directamente dependientes de un contexto histórico. Las relaciones que guarda una explicación con el objeto que se intenta elucidar a

través de ella pueden cambiar vistas desde un nuevo marco conceptual acotado por una situación histórica particular.

Bajo ciertas condiciones algunas explicaciones exitosas han sido presentadas, inicialmente, con serios problemas. En ocasiones, dichos problemas se van atenuando por modificaciones posteriores. Otras veces, este tipo de dificultades se fortalecen con el tiempo hasta convertirse en factores importantes para el abandono total de la explicación propuesta.

Un caso particularmente controvertido respecto de la perspectiva clásica para la modelación de la explicación científica ocurre cuando el *explanans* está constituido por un cuerpo inconsistente de conocimientos. Tal característica fulminaría, aparentemente en forma inmediata, la seriedad de la explicación. Sin embargo, la historia nos enfrenta con algunos ejemplos de explicaciones de este tipo especial que se sostienen a pesar del descubrimiento de su peculiar estructura lógica. Me propongo en esta parte del capítulo enfatizar el primero de los dos problemas que directamente involucran el contexto en la explicación: el problema de la inconsistencia de algunos contextos teóricos, mostrando un caso concreto de este tipo de explicaciones y enfatizando la dificultad de modelarla formalmente y su relación con el tipo de racionalidad que nos revela.

En 1893, en su *Grundgesetze der Arithmetik* vol. 1, Frege intenta solucionar el problema de la definición del concepto de número¹¹⁵, como parte de su trabajo para la fundamentación de la aritmética. La definición de Frege constituyó un intento por explicar adecuadamente el concepto de número y, partiendo de ahí, proporcionar un fundamento a la aritmética. Para ello, hizo uso de la teoría ingenua de conjuntos de George Cantor la cual, a la postre, resultó ser una teoría inconsistente.

¹¹⁵Ya en [Frege 1884], Frege lo había hecho pero con menores detalles en su *Grundlagen der Arithmetik*, párrafos 69- 76. Véase la traducción de Hugo Padilla: [Frege 1972] Frege, Gottlob, *Conceptografía - Fundamentos de la Aritmética - Otros Estudios*, Colección Filosofía Contemporánea, Instituto de Investigaciones Filosóficas, UNAM, México, 1972, pp. 175-6. Para los *Grundgesetze* véanse párrafos 42 y 43 en su primer volumen.

La clave para la explicación de Frege es considerar a un número x como la clase A tal que contenga a todas las clases B cuya cardinalidad es x . El procedimiento no es circular pues está fundamentado en una relación que Cantor usó para demostrar algunas relaciones entre distintos números infinitos: la relación uno a uno entre los miembros de dos conjuntos. La relación uno a uno, toma un elemento del primer conjunto, digamos el conjunto B_1 y un elemento del segundo conjunto, B_2 ; luego toma un segundo elemento del conjunto B_1 y un segundo elemento de B_2 ; la relación es uno a uno si ocurre que una vez recorrido todo el dominio (todos los elementos de B_1) se ha terminado con los elementos de B_2 y no existe un elemento de cualquiera de los dos conjuntos al cual le corresponda más de un elemento. Así, es posible verificar que todas las subclases B poseen el mismo número de miembros, sin suponer la noción de número.

La explicación de Frege se encuentra montada sobre la teoría de conjuntos cantoriana. La posibilidad de tener una clase como la clase A mencionada en el párrafo anterior, está fundamentada en lo que Frege llamó "axioma de comprensión" [Véase Dieudonné, 1989]¹¹⁶, es decir, la Ley V de Frege. Lo anterior, debido a que a partir de esta ley se deriva, en el sistema de Frege, la posibilidad de construir una clase, en base a cualquier propiedad estipulada como requisito de membresía para sus elementos. La Ley V de Frege puede expresarse como sigue¹¹⁷:

$$\forall x f(x) = \forall x g(x) \leftrightarrow (x)(f(x) \leftrightarrow g(x))$$

Sin esta ley no podríamos obtener algo como esto¹¹⁸:

¹¹⁶Dieudonné, Jean, *En Honor del Espíritu Humano. Las Matemáticas Hoy*, Alianza Universidad, España 1989. p. 306.

¹¹⁷Véase la introducción de Montgomery Furth en [Frege 1964]: Frege, Gottlob, *The Basic Laws of Arithmetic*, University of California Press, Berkeley y Los Angeles, 1964, pp. xi-xiii. Traducción de alguna parte de los *Grundgesetze der Arithmetik*. Por supuesto, la notación no es la de Frege, y " $\forall x f(x)$ " puede leerse como "el conjunto formado por los individuos que satisfacen f ".

¹¹⁸Idem.

$$f(y) \leftrightarrow y \in \wedge x f(x)$$

El cual se ofrece como primera derivación en el primer volumen del *Grundgesetze der Arithmetik*¹¹⁹. Ello nos permite transitar de la predicación entre un concepto y la extensión de un conjunto correspondiente, como el llamado axioma de abstracción de Cantor.

3.1.2 El *explanans* inconsistente: la Paradoja de Russell.

Bertrand Russell descubrió una paradoja en relación con esta forma de construir a las clases. Alejandro Garciadiego, quien realizó una investigación histórica acerca de las paradojas descubiertas por Russell expone algunos datos interesantes en relación a esta paradoja. Al parecer, la primera versión de la paradoja puede considerarse una nota en el primer borrador de sus *"Principles of Mathematics"*, en 1901 [Garcia diego 1992]¹²⁰. Su formulación más precisa aparece por primera vez en la carta de Russell a Frege del 16 de junio de 1902¹²¹. En la carta Russell menciona que le ha escrito a Peano sin obtener respuesta y se refiere a una contradicción importante provocada desde los supuestos de Frege, particularmente, desde la transición irrestricta cantoriana entre una propiedad y un conjunto. La contradicción la expone Russell de la siguiente manera:

"Sea ω el predicado: ser un predicado que no puede predicarse de sí mismo. ¿Puede ω predicarse de sí mismo? De cada respuesta se sigue su opuesta." [Robles 1982]¹²²

¹¹⁹Ibid, p. xi-xii. La derivación corresponde a los párrafos 54 y 55 del primer volumen de *Grundgesetze der Arithmetik*. Véase apéndice I de [Frege 1964], op. cit., pp. 123-126.

¹²⁰[Garcia diego 1992] Garciadiego Dantan, Alejandro R., "Bertrand Russell y los Orígenes de las "paradojas" de la Teoría de Conjuntos", Alianza Universidad, España, 1992. p. 139.

¹²¹Ibíd, p. 152.

¹²²[Robles 1982] Robles, José Antonio, *"Bertrand Russell: Antología I"*, SEP - Setentas - Diana, México, 1982. p. 65. La traducción es hecha de la versión inglesa de Beverly Woodward, con la aprobación de Russell, que fue publicada en [Heijenoort 1967], Heijenoort, Jean van, *From Frege to Gödel*, Harvard University Press, Cambridge, Massachusetts, 1967.

La paradoja puede entenderse también de esta forma: Supongamos que existan clases que no se pertenecen a sí mismas y clases que sí se pertenecen a sí mismas. Ahora bien, pensemos en una clase mayor que contenga a todas las clases que no se pertenecen a sí mismas. Llamemos a esta clase "la clase Φ " y llamemos a las clases que no se pertenecen a sí mismas "las clases ψ ", esto es, todas las que están contenidas en Φ . A partir de lo anterior, si suponemos que Φ se pertenece a sí misma, entonces, dado que sólo pertenecen a Φ las clases ψ , tenemos que Φ es una ψ pero toda ψ , por definición, no se pertenece a sí misma; si, por el contrario, suponemos que Φ no se pertenece a sí misma entonces, por definición, Φ es una clase ψ y esto significa que Φ está dentro de Φ dado que toda ψ lo está. Por lo tanto, Φ se pertenece a sí misma si y sólo si Φ no se pertenece a sí misma.

La paradoja de Russell es uno de los ejemplos más sencillos de este tipo de contradicciones; sólo requiere de dos nociones: el primitivo " \in " y el concepto de clase. No obstante su simplicidad, pone en consideración la consistencia de la teoría de conjuntos de Cantor. Es importante decir, que Cantor había ya notado, aunque no expresada como en la paradoja de Russell, que su teoría generaba inconsistencias, lo que lo llevó a admitir que había algunos conjuntos inconsistentes y otros consistentes, asunto que no consideró especialmente problemático. Así, parecería que la suposición de una lógica subyacente en las teorías, es un punto clave para la consideración de la inconsistencia como un problema¹²³. Frege usa el postulado problemático de la teoría de Cantor para su explicación de la noción de número, a saber, lo que hemos mencionado como el axioma de abstracción, el *explanans* de su explicación es, pues, inconsistente, y él pareció estar bien consciente de que esto significaba un grave problema para su teoría.

Después de la paradoja de Russell, aparecieron muchas otras igualmente serias respecto de la teoría de conjuntos de Cantor¹²⁴. Todo ello afectaba no sólo al proyecto

¹²³ Debo esta observación a Max Fernández.

¹²⁴ Antes de la paradoja de Russell, Burali-Forti había encontrado una paradoja acerca de la existencia del número cardinal mayor. Russell comenta que esta paradoja no había provocado el efecto de la suya, tal vez

entero de Frege para la fundamentación de la aritmética. Por una parte, el uso irrestricto del axioma de comprensión se había efectuado antes para la definición de "lugar geométrico"¹²⁵; además, la teoría de Cantor de los conjuntos había penetrado ya al análisis matemático haciendo que muchos teoremas importantes dependieran de la salud de esta construcción intelectual. Algunos de estos teoremas eran el teorema de Heine-Borel y el teorema de Baire de las funciones [Bell 1985]¹²⁶. Así, después del golpe al proyecto fregeano, el análisis también se vio fuertemente afectado por las paradojas. E. T. Bell lo expresa como sigue:

"Al derrumbarse la teoría de clases del sistema de los números, el análisis quedó sin cimientos, suspendido en medio del aire como el ataúd de Mahoma, sustentado sólo por el milagro de la fe."¹²⁷

3.1.3 Implicaciones: T. de Conjuntos y Matemáticas.

Es interesante observar tres tipos diferentes de reacción ante este formidable acontecimiento intelectual. Primero, Cantor, no consideró problemático el asunto de la inconsistencia. Segundo, en el caso de Frege, al principio pensó, como Russell, que sería posible arreglar el problema mediante algún cambio no esencial. Después Frege decidió considerar que todo su programa de investigación dirigido a la fundamentación de la aritmética era, muy probablemente, insostenible. Consecuentemente, Frege decide añadir la siguiente nota al final de la impresión del segundo volumen de su *Grundgesetze der Arithmetik*, en 1903:

debido a su complejidad. [Russell 1964] Russell, Bertrand, *La Evolución de mi Pensamiento Filosófico*, Aguilar, Madrid, 1964, pp. 80-81.

¹²⁵[Dieudonné 1989] Dieudonné, Jean, *En Honor del Espíritu Humano. Las Matemáticas Hoy*, Alianza Universidad, España 1989, p. 306.

¹²⁶[Bell 1985] (1940) Bell, E. T., *Historia de las Matemáticas*, Fondo de Cultura Económica, México, 1985, p. 497. Traducción del original: *The Development of Mathematics*, McGraw-Hill Book Co., Nueva York, 1940.

¹²⁷Ibid, p. 293.

"Difícilmente algo más desafortunado puede sucederle a un hombre de ciencia que el derrumbe de uno de los pilares de su ciencia una vez que ha concluido la tarea.

Esta fue la posición en la que me situó una carta del señor Bertrand Russell justamente cuando se concluía la impresión del presente volumen.

Lo que aquí está en juego no es tan sólo mi forma particular de fundar la aritmética, sino el si se le podrá dar a la aritmética algún tipo de fundamentación lógica."¹²⁸

Tercero, en contraste con esta reacción, Russell parece no haber creído inicialmente, al igual que Frege, la veracidad de esta contradicción, pero cuando aceptó que el problema era legítimo decidió reformular los supuestos empleados en la investigación de Frege:

"Al principio pensé que debía de haber algún error trivial en mi razonamiento. Examiné cada paso bajo un microscopio lógico, pero no pude descubrir nada incorrecto. Escribí a Frege acerca de ello, y me replicó que la aritmética se tambaleaba y que ahora veía que su ley V era falsa. Frege quedó tan desasosegado por esta contradicción que dio de lado el intento de deducir la aritmética de la lógica, y, al cual, hasta entonces, había dedicado principalmente su vida. Como los pitagóricos cuando tropezaron con los inconmensurables, buscó refugio en la geometría y al parecer consideró que el trabajo de su vida hasta aquel momento había estado mal orientado. Por mi parte, me di cuenta de que la dificultad residía en la lógica más bien que en las matemáticas, y era la lógica lo que había de reformarse. Me confirmé en esta opinión al descubrir una fórmula por medio de la cual se podía formar un número estrictamente infinito de contradicciones."**[Russell 1964]**¹²⁹

¹²⁸[Robles 1982], *op. cit.*, p. 62.

¹²⁹[Russell 1964] Russell, Bertrand, *La Evolución de mi Pensamiento Filosófico*, Aguilar, Madrid, 1964, p. 79. La obra aludida es [Russell 1944] Russell, Bertrand, *My Philosophical Development*, Londres, George Allen y Unwin 1944. Véase también [Robles 1982], *op. cit.*, p. 47, tomado de "Mi desarrollo Mental" (1944): *Escritos Básicos* (1903-1959), Aguilar, México, 1969, pp. 49-68.

Para 1902, Russell se encontraba escribiendo su libro *Principles of Mathematics* y al concluirlo, en 1903, no había llegado todavía a una solución adecuada para ésta y otras dos paradojas de la teoría de conjuntos¹³⁰.

No obstante, la diferencia entre estos dos tipos de reacción, el tipo de reacción de Frege y el de Russell, radica en el grado de certeza que le conceden a la teoría de conjuntos cantoriana. Por supuesto, la lamentación de Frege ante la paradoja se encuentra involucrada con la negación de todo un programa de investigación, mientras que para Russell puede no tener esas megalíticas consecuencias. Sin embargo, lo que lleva a Frege a esa conclusión es su confianza en la veracidad de la teoría de Cantor. Russell, por su parte, y sin soslayar un ápice la importancia de la paradoja una vez entendida, toma el asunto con cautela y, a pesar de considerar que la contradicción es legítima ---y por tanto que no se debe a un error de razonamiento ni en Cantor ni en Frege ni en él---, decide buscar la solución revisando algunos de los supuestos iniciales de la teoría de Cantor. Mientras Frege decide renunciar a la estrategia de explicación por así decirlo, Russell opta por modificar los supuestos más remotos del *explanans* o añadir a éste algún tipo de supuesto restrictivo, de tal forma que dichos cambios, mientras permitieran la construcción de las clases construidas por Frege para su explicación, prohibieran la construcción de clases tan problemáticas como las que Russell mismo usa en su paradoja.

Como es natural en la historia de la ciencia, las explicaciones plausibles no son abandonadas en forma inmediata, no antes de una revisión de sus posibilidades de fertilidad y otras ventajas respecto de alternativas contemporáneas. En este caso, abandonar la explicación de la noción de número significaba, si somos consecuentes, abandonar cualquier explicación basada en esta teoría de los conjuntos; no obstante, hacer esto significaría renunciar a un poderoso instrumento de análisis para la matemática

¹³⁰La paradoja de Cantor y la de Burali-Forti.

y, con ello, a un poderoso instrumento de explicación. Hermann Weyl expresa así la importancia de la teoría de conjuntos:

"El método de la teoría de conjuntos ha impregnado no sólo al análisis sino también a la matemática primaria, la teoría de los números naturales." [Weyl 1965]¹³¹

Es importante aclarar que los resultados mencionados por Weyl eran ya conocidos por Bertrand Russell y sus contemporáneos. Otra nota al respecto es la de Dieudonné quien se refiere a la reconstrucción de los resultados sobre teoría de conjuntos que hace Dedekind en 1872:

"La importancia de este lenguaje [el lenguaje de la teoría de conjuntos] reside en que permite a los matemáticos, a partir de lo últimos años del siglo XIX, considerar relaciones entre objetos de naturaleza *completamente indeterminada*: son simplemente elementos de conjuntos considerados como objetos primitivos de una teoría axiomática." [Dieudonné 1989, p. 188]¹³²

De esta manera, algunos supuestos del *explanans* de Frege eran inconsistentes pero la estrategia de explicación, esto es, la definición de número mediante el uso de clases, fue tan valiosa que Russell la conserva todavía en 1919, año en que publica su *Introduction to Mathematical Philosophy* en donde propone una definición de la noción de número dándole todo el crédito a Frege [Véase Russell 1975]¹³³. Dado que son los supuestos los que producen la contradicción, en el fondo, el problema no estaba directamente en lo que usualmente llamamos *explanans* sino, más bien, en el contexto el

¹³¹[Weyl 1965] Weyl, Hermann, *Filosofía de las Matemáticas y de la Ciencia Natural*, Centro de Estudios Filosóficos, UNAM, México, 1965, p. 52. Traducción de la versión original de 1949. Lo que está en corchetes es mío.

¹³²[Dieudonné 1989], *Op. cit.*, p. 188. Estructuras matemáticas como los grupos, los anillos y algunas nociones topológicas, se proponen como ejemplo en el texto. Lo que está entre corchetes es mío.

¹³³[Russell 1975] Russell, Bertrand, *Introduction to Mathematical Philosophy*, London George Allen & Unwin Ltd, Great Britain, 1975. p. 11.

que el *explanans* había sido propuesto. Este es un problema que surge cuando consideramos al contexto como parte esencial de las explicaciones, en cierto sentido, por lo menos algunas partes del contexto resultan imprescindibles en la forma en que un *explanans* da cuenta del evento en consideración. Podemos considerar así dos clases de *explanans*, uno directo, y uno indirecto o contextual. El primero de ellos es el conjunto de proposiciones que uno propone para dar cuenta del evento *explanandum*. El segundo, el contexto teórico en el que se propone la explicación y que puede ser el factor directamente responsable de que efectivamente el *explanans* directo dé cuenta del *explanandum*, como en el caso de Frege.

| Ya en 1919 Russell disponía de algunas soluciones alternativas para evitar la paradoja, como la teoría de los tipos y nuevas versiones de la teoría de conjuntos axiomatizada. En cierto sentido, la estrategia de la explicación de la noción de número en Russell preserva la solución que había propuesto Frege desde su *Grundlagen der Arithmetik*. La diferencia fundamental está en que los presupuestos de su *explanans* se encuentran, al parecer¹³⁴, libres de contradicciones.

Sin embargo, algo sumamente interesante es que los matemáticos no esperaron hasta la aparición de la primera modificación de la teoría de conjuntos, la teoría axiomática de conjuntos de Zermelo en 1908, para seguir realizando importantes desarrollos en varias áreas de la matemática. A pesar del descubrimiento del carácter inconsistente de la teoría de Cantor, muchos matemáticos siguieron usándola como instrumento para impulsar soluciones a viejos problemas matemáticos y, lo cual es de

¹³⁴Libres de contradicción sólo en el sentido de que han desaparecido las paradojas de finales del s. XIX y principios del s. XX; lo mismo ocurre, aunque con una estrategia de solución diferente a la de Russell, con las formulaciones axiomáticas de Zermelo-Fraenkel y von Neumann. Poincaré decía a este respecto: "Hemos puesto una valla en torno al rebaño para protegerlo de los lobos, pero lo que no sabemos es si habrá quedado algún lobo dentro ya de la valla". [Kline 1992] (1972) Kline, Morris, *El pensamiento Matemático de la Antigüedad a nuestros Días*, tomo III, Alianza Universidad, Madrid, España, 1992, pp. 1566-9. Del original: *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, 1972.

mayor interés para este trabajo, como un poderoso *cuerpo explicativo*. La topología¹³⁵ por ejemplo sufrió un impulso formidable.

El tratamiento que había efectuado Cantor de los espacios geométricos como conjuntos de puntos, hizo posible pensar que algunos resultados aislados sobre características abstractas del espacio matemático podrían unificarse en una teoría más general. Maurice Fréchet lo intentó por primera vez en su tesis doctoral de 1906, sus resultados fueron la estructura ósea sobre la cual se construyera la nueva topología del siglo XX. Un comentario ilustrativo de la significación del trabajo de Fréchet en los desarrollos de este siglo en Topología es el siguiente:

"Con esta obra de Frechet [sic] y la [...] de Hausdorff la topología pasó en el segundo decenio del siglo XX a la fase postulacional o completamente abstracta."**[Bell 1985, p. 476]**¹³⁶

Uno de los factores medulares en el renacimiento de esta disciplina matemática fue, sin duda, la teoría de los conjuntos de Cantor¹³⁷. Morris Kline lo dice de la siguiente manera:

"El primer esfuerzo importante por construir una teoría abstracta de espacios de funciones y de funcionales fue llevado a cabo por Maurice Fréchet (1878), eminente matemático francés, en su tesis doctoral de 1906. En lo que Fréchet llamó cálculo funcional, intentó unificar en términos abstractos las ideas contenidas en los trabajos de Cantor, Volterra, Arzela, Hadamard y otros.

Con objeto de conseguir el más alto grado de generalidad para sus espacios de funciones, Fréchet echó mano de todas las ideas básicas sobre conjuntos

¹³⁵En términos generales, la topología es entendida como el estudio de las propiedades invariantes de los espacios geométricos bajo transformaciones de cierta clase especial.

¹³⁶**[Bell 1985]**, op. cit., p. 476.

¹³⁷Ibid, p. 475.

desarrolladas por Cantor, aunque para él los puntos de los conjuntos eran ahora funciones." [Kline 1992, p. 1422]¹³⁸

Y, más sintéticamente:

"Fréchet, el año de 1906, impulsado por el deseo de unificar la teoría de conjuntos de puntos de Cantor y el tratamiento de las funciones como puntos de espacio, que ya se había hecho habitual en el cálculo de variaciones, inició el estudio de los espacios abstractos." [Kline 1992, p. 1531]¹³⁹

3.1.4 Racionalidad e Inconsistencia en la modelación de la explicación.

Lo sorprendente en el caso que nos ocupa es el hecho de que se utilice, confiando en una posible solución posterior, un cuerpo explicativo del cual ha sido demostrada su inconsistencia. Tenemos, pues, que existen explicaciones cuya estrategia esencial, construida en el *explanans*, puede ser preservada a pesar de poseer un estatus claramente inconsistente. En este caso la explicación propuesta por Russell conserva la estrategia de la explicación de Frege; y también, como hemos visto, es posible que cuerpos explicativos suficientemente fructíferos sean preservados y usados en otras empresas científicas, aún a pesar de ser explícitamente inconsistentes, particularmente los desarrollos en topología a principios de siglo.

Por tanto, en ocasiones los científicos continúan trabajando con sistemas que contienen contradicciones, por lo menos transitoriamente. Podemos llamar a esto "inconsistencia de transición", pero la consecuencia inmediata de una inconsistencia en un cuerpo teórico que consideramos como suponiendo una racionalidad deductiva clásica, es un problema que es claramente grave: la trivialización o explosión del sistema, esto es,

¹³⁸[Kline 1992] Kline, Morris, *El pensamiento Matemático de la Antigüedad a nuestros Días*, tomo III, Alianza Universidad, Madrid, España, 1992, p. 1422. Del original: *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, 1972.

¹³⁹Ibid, 1531.

la consecuencia de que cualquier tesis es inferible a partir de ese cuerpo teórico: Una vez que aceptamos una contradicción en el sistema, por un argumento muy sencillo, el sistema puede inferir cualquier tesis, independientemente de que ésta sea verdadera o no; así, el sistema prácticamente “explota” y se vuelve “trivial”, no puede explicar relevantemente nada a partir de entonces. Esto lo explicaré enseguida.

Podría pensarse que una situación como la que he descrito es desalentadora respecto de algunas perspectivas de la racionalidad. En particular, parecería plausible, después de un análisis como el que he expuesto, pensar que cualquier tentativa por elucidar, objetivamente, el mecanismo racional de algunas explicaciones, sería una quimera. Como comentaremos más adelante, este no es el caso.

El punto que me interesa enfatizar en esta parte del capítulo, es la dificultad que representa la modelación de esta clase de explicaciones bajo una perspectiva clásica de la Lógica. Si intentáramos modelar la explicación de Frege en los términos a los que acabo de aludir asumiríamos que en el *explanans* estaban contenidas todas las proposiciones, ya que es natural desde la lógica clásica que de una contradicción se siga cualquier cosa. Existen otros ejemplos como éste en la historia de las matemáticas, como el cálculo infinitesimal en sus primeros tiempos; también podemos encontrar ejemplos muy famosos en la física, como el caso de la teoría de Bohr acerca del átomo o el de la mecánica cuántica [Priest & Routley 1989]¹⁴⁰. En estas circunstancias, los pensadores involucrados con la teoría piensan en cómo reestructurar el cuerpo explicativo si es que hay alguna forma de hacerlo, pero no consideran que la teoría misma que está en juego posee toda afirmación posible, lo cual, dicho sea de paso, no obsta para la búsqueda seria de una solución para la estructura lógica de la teoría.

¹⁴⁰[Priest & Routley 1989c] Priest, G. y Routley, R.; "Applications of Paraconsistent Logic" en Priest, Graham; Routley, Richard; Norman, Jean; *Paraconsistent Logic. Essays on the inconsistent*. Philosophia Verlag, Alemania, 1989, pp. 368-380.

Con todo, es posible que el problema de la modelación de inconsistencias al que apuntan las aseveraciones del párrafo anterior pueda ser solucionado si partimos de una perspectiva lógica no clásica. Si pudiésemos construir un ambiente formal que fuera capaz de soportar inconsistencias y preservar el objetivo analítico para el que se pretende utilizar, tendría lugar un estudio formal de teorías como las que nos ocupan. La dificultad estriba en el hecho de que introducir formalmente proposiciones contradictorias, implicaría la posibilidad de inferir cualquier fórmula del sistema. Esto es llamado "la trivialización del sistema". Generalmente, dicha trivialización es provocada en los sistemas clásicos de lógica mediante el principio conocido como Principio del Pseudo Scoto¹⁴¹:

$$(S \wedge \neg S) \supset Z$$

El principio dicta que una vez que hemos obtenido una contradicción en el sistema entonces podemos obtener cualquier proposición del lenguaje. Una elegante justificación del principio puede ser realizada en base a las reglas de silogismo disyuntivo, simplificación y adición; esto es llamado "el argumento de Lewis":

$S \wedge \neg S$	Supuesto.
S	Simplificación.
$S \vee Z$	Adición.
$\neg S$	Simplificación.
Z	Silogismo disyuntivo.

Otra forma de este principio es la siguiente:

$$\{S, \neg S\} \vdash Z$$

¹⁴¹ En relación con este principio es importante mencionar la tesis doctoral de **[Gomes 2013]** Evandro Luis Gomes, "Sobre a História da Paraconsistência e a obra de da Costa: a instauração da Lógica paraconsistente", defendida el 12 de diciembre de 2013, Universidade Estadual de Campinas, Brasil. En este trabajo se hizo un estudio cuidadoso sobre este principio.

El principio puesto de esta manera expresa que a partir de dos proposiciones contradictorias, podemos obtener cualquier proposición del sistema. Una justificación puede ser la que sigue:

S	Supuesto.
$\neg S$	Supuesto.
$S \vee Z$	Adición.
Z	Silogismo disyuntivo.

La diferencia entre las dos presentaciones anteriores del Principio del Pseudo Scoto, es que en el primero establecemos que de una contradicción, esto es, algo de la forma $P \wedge \neg P$, se sigue cualquier fórmula del sistema, en el segundo establecemos que basta tener dos proposiciones contradictorias para tener el mismo efecto; de aquí, por conjunción, podemos obtener la primera formulación del principio.

Así, desde la lógica clásica, la contradicción luce como algo inseparable de la trivialidad. Admitir contradicciones bloqueando la trivialidad no parece entonces una tarea fácil. El argumento de Lewis ilustra perfectamente algunos puntos de ataque para evitar la explosión de fórmulas en el sistema; no obstante, las reglas usadas por el argumento de Lewis son profundamente intuitivas y esto constituye el carácter medular de la investigación sobre estos tópicos en relación a la naturaleza misma de la lógica.

Algunas investigaciones formales francamente alejadas de principios fundamentales de la lógica clásica han logrado crear sistemas de lógica con características inconsistentes que evitan el escollo de la trivialización, ellos han sido llamados por Francisco Miró Quesada: "sistemas de lógica paraconsistente".

Las motivaciones que se tienen para realizar trabajos sobre lógica paraconsistente son muy diversas, los primeros sintieron que sería muy interesante transpolar el gran acontecimiento de las geometrías no-euclidianas a la lógica, modificando alguno de sus principios fundamentales (no contradicción, tercio excluso, identidad) [D'Ottaviano 1990]¹⁴², algunos otros investigadores produjeron trabajos por razones involucradas con aspectos de la dialéctica, otros han visto en este tipo de lógicas una nueva perspectiva para abordar lógicas borrosas, también se ha pasado por la parcela de la lógica paraconsistente incluso por necesidad como es el caso de las lógicas relevantes, cuyas propiedades requeridas coinciden con algunos sistemas paraconsistentes; el gusto por representar los procesos del cambio científico también ha sido un motivo interesante para la producción de lógicas paraconsistentes. En general, los motivos a los que he estado haciendo referencia, convergen en la búsqueda de la representación formal de teorías inconsistentes que pueden ser de interés o, directamente, en el análisis de las propiedades de las teorías inconsistentes y de sus objetos¹⁴³. Pensar los cuerpos teóricos como suponiendo una lógica subyacente que no fuera clásica, sino más bien paraconsistente, evitaría la consecuencia explosiva en la modelación de contextos inconsistentes. Así, con una lógica de este tipo se puede capturar lo que se puede llamar “paraconsistencia de transición”, un estado inconsistente del cuerpo teórico de la explicación

Dentro de estos múltiples enfoques de la paraconsistencia, especialmente interesante para los fines de modelación de la Ciencia, es el proseguido por Newton C.A.

¹⁴² Podemos considerar los trabajos de Jaśkowski, en la tradición polaca, y los de Newton da Costa, en la tradición brasileña, como dos de los trabajos fundacionales de la Lógica Paraconsistente. Véase para el primero, [D'Ottaviano & da Costa 1970] D'Ottaviano, I.M. L., & N.C.A. da Costa, “Sur un probleme de Jaśkowski”, en *Comptes Rendus Acad. Sci.*, 270A (1970), 1349–1343. Para el segundo consúltese [da Costa 1974] da Costa, Newton C. A., “On the theory of Inconsistent Formal Systems”, *Notre Dame Journal of Formal logic NDJFL*, Vol. 15, Num .4, 1974; o bien, para la historia de la Lógica Paraconsistente, [D'Ottaviano 1990] D' Ottaviano, Ítala M. Loffredo, “On the Development of Paraconsistet Logic and da Costa's Work”, *The Journal of Non-Classical Logic (TJNCL)*, Vol.7, No. 1-2, 1990.

¹⁴³ Como el análisis de los objetos de la teoría de Meinong. [Priest & Routley 1989d] Priest, Graham, y Routley, Richard, “The Philosophical Significance and Inevitability of Paraconsistency”, en Priest, Graham; Routley, Richard; Norman, Jean; *Paraconsistent Logic. Essays on the inconsistent*. Philosophia Verlag Alemania, 1989, pp. 517 et s.

da Costa, uno de los fundadores de la Lógica Paraconsistente. Su propuesta asume de modo explícito que los cuerpos teóricos pueden considerarse como teniendo lógicas subyacentes y propone una jerarquía de cálculos paraconsistentes con los que pueden modelarse algunas contradicciones permitidas en el sistema, sin que esto signifique la aceptación irrestricta de toda contradicción.

Así, este aspecto peculiar de la racionalidad de la explicación científica, que surge al incluir el contexto en su modelación, la inconsistencia que presentan algunos sistemas de conocimiento y sus consecuencias en relación a explicaciones objetivamente formuladas, no nos invita, necesariamente, a un abandono de la búsqueda por sus representaciones formales. Sabemos que el posible descubrimiento de contradicciones no obsta para seguir usando un cuerpo explicativo fructífero pero, además, sabemos que existen herramientas lógicas que pueden ayudarnos a modelar este tipo de cuerpos explicativos, lo cual significa un punto a favor de la lógica como herramienta de análisis en algunos aspectos concretos de la racionalidad.

3.1.5 Un sistema Paraconsistente: Los Cálculos C_n .

Introducción

Para proporcionar un acercamiento a las herramientas formales que nos interesa destacar de los sistemas paraconsistentes es conveniente caracterizarlos. Me ocuparé de ello tanto respecto de sus definiciones generales como de las diversas formas de construcción, en ésta y la siguiente sección. Usaré para esta reconstrucción histórica un artículo que es una referencia obligada en la historia de la paraconsistencia, a saber, **[D'Ottaviano 1990]**, que ya he citado antes¹⁴⁴.

La tarea de caracterizar a las lógicas paraconsistentes puede resultar ardua en razón de la amplia diversidad de sistemas existentes. En esta sección revisaré algunas definiciones de los sistemas paraconsistentes y presentaré la jerarquía de cálculos paraconsistentes C_n , $1 \leq n \leq w$, de da Costa. Esto servirá de antecedente para mostrar cómo el sistema de da Costa puede usarse para la construcción de un modelo formal de la explicación científica.

¹⁴⁴ En la sección de Explicación e Inconsistencia.

El nombre de "Lógica Paraconsistente" fue propuesto por el filósofo Peruano Francisco Miró Quesada en 1976. En general, una lógica paraconsistente se entiende como una lógica en la cual el principio de la no contradicción no es válido en general pero de una contradicción no se trivializa el sistema.

Stanislaw Jaśkowski y Newton C. A. da Costa poseen el mérito de haber formulado por vez primera una definición de este tipo de sistemas. Para da Costa una lógica paraconsistente debe cumplir con 4 requisitos **[da Costa 1974]**¹⁴⁵:

- 1) El Principio de No Contradicción, a saber, $\neg(A \wedge \neg A)$, no es válido en forma general.
- 2) No es válido deducir una proposición arbitraria a partir de A y de $\neg A$.
- 3) Debe ser capaz de subsumir las aplicaciones de las reglas de inferencia de la lógica clásica que sean compatibles con el requisito 2).
- 4) Debe ser susceptible de extenderse al cálculo de primer orden.

El requisito 2) de da Costa es fundamental, pues constituye el carácter interesante de la lógica paraconsistente: la ausencia de trivialidad aún a pesar de la violación del Principio de No Contradicción. Los requisitos 1) y 4) no se consideran necesarios para la construcción de una lógica paraconsistente y obedecen a la tendencia de da Costa de preservar lo más que se pueda el carácter clásico en este nuevo tipo de lógicas. Da Costa afirma que el requisito 3) es vago **[da Costa 1974, p. 498]**, y se ha encontrado que algunas lógicas que intuitivamente parecen ser paraconsistentes no lo cumplen.

Por lo anterior, me parece conveniente intentar otra definición. En ello nos puede ser útil una clasificación más amplia, combinando una elaborada en [Peña 1993]¹⁴⁶ y otra que es usual en la escuela brasileña de paraconsistencia, para ello tomaré la clasificación supuesta en **[Carnielli&Coniglio&Marcos 2000]**¹⁴⁷. Esto nos permitirá diferenciar cinco tipos de teorías: inconsistentes, consistentes, contradictorias, explosivas y, finalmente, paraconsistentes. Podemos distinguir cuatro tipos de lógicas relacionadas con la inconsistencia, a saber:

- a) Lógica Inconsistente.- Es una lógica en la que por lo menos una fórmula y su negación son derivables en ella (es inconsistente con respecto a algún functor de negación) o viola al menos algún principio semántico (por ejemplo sobre valores de verdad) de no-contradicción.

¹⁴⁵ **[da Costa 1974]** da Costa N.C.A., "On the theory of Inconsistent Formal Systems", *Notre Dame Journal of Formal Logic NDJFL*, Vol. 15, Num .4, 1974, p. 498. Para esta formulación usé también: **[Arruda 1988]** Arruda, Ayda, "Panorama de la Lógica Paraconsistente", en CELIJS, Antología de la lógica en America Latina, Universidad de Carabobo, Valencia, Venezuela, 1988, p. 171; y Urbas, Igor, "Paracaonsistency and the C-Systems of da Costa", *NDJFL*, Vol. 30, No. 4, 1989. p. 583.

¹⁴⁶ **[Peña 1993]** *Introducción a las Lógicas No Clásicas*, IIF-UNAM, 1993.

¹⁴⁷ **[Carnielli & Coniglio & Marcos 2000]**, Carnielli, Walter A. & Marcos, Joao, "A taxonomy of C-Systems", en **[Carnielli & D'Ottaviano & Coniglio 2000]** Walter A. Carnielli & Ítala M. L. D'Ottaviano & Marcelo E. Coniglio, *The Logical Way to the Inconsistent. Proceedings of the World Congress of Paraconsistency*. State University of Campinas, Edit. Marcel Dekker Inc., Sao Paulo, Brasil, 2000.

b) Lógica Contradictorial.- Es una lógica inconsistente en la que puede derivarse la conjunción de una fórmula con su negación.

c) Lógica Explosiva.- Es una lógica que, dada una extensión de ella que sea inconsistente, se trivializa.

d) Lógica Paraconsistente.- Es una lógica que no es explosiva.

Esta clasificación proporciona una relación clara entre la paraconsistencia y la lógica clásica. En efecto, es posible entender la lógica paraconsistente como una lógica, ya sea inconsistente o consistente, cuyas extensiones inconsistentes no son triviales. El énfasis está puesto en esta última propiedad. Así, la clasificación nos permite ver claramente la distinción entre tener una lógica inconsistente, la cual puede ser clásica, y tener una lógica paraconsistente. Además, con esta clasificación, una caracterización como la de da Costa incluye el caso de una lógica (inconsistente o consistente) cuyas extensiones contradictoriales¹⁴⁸ no son triviales, pero también el caso de una lógica cuyas extensiones simplemente inconsistentes no son triviales, lo cual es mucho más general.

En [Arruda 1988]¹⁴⁹ se propone una caracterización más directa, basada en el concepto de trivialidad. Sintéticamente, su definición es como sigue:

Si
L es un lenguaje,
F el conjunto de fórmulas de L,
S es un subconjunto de F,
C es la lógica subyacente de S,
y por tanto S es una teoría,
S es inconsistente con respecto a un functor de C,
S no es igual a F;
Entonces,
C es una lógica paraconsistente con respecto a un functor de C.

La ventaja de estas caracterizaciones es que no se comprometen a una forma de expresión del Principio de No Contradicción ni, más específicamente, al requisito 1) de da Costa. Es interesante observar que Arruda sólo caracteriza "ser inconsistente respecto de un functor" y no, como hemos visto con Peña, "ser inconsistente respecto de un functor de negación". Señalar esta amplitud de la caracterización de Arruda es importante si consideramos los operadores M y L de las lógicas modales como parte de esta clase de funtores. En una lógica de este tipo podemos tanto obtener LA (siendo A una fórmula)

¹⁴⁸ Extensiones que contengan contradicciones.

¹⁴⁹ [Arruda 1988] Arruda, Ayda, "Panorama de la Lógica Paraconsistente", en [CELLIJS, 1988] C.E.L.L.I.J.S., *Antología de la lógica en América Latina*, Universidad de Carabobo, Valencia, Venezuela, 1988, p. 615.

como, de una forma sencilla, obtener también A. Esto significaría que la lógica modal es una lógica paraconsistente respecto del functor L, lo que parece muy extraño. Podemos precisar un poco la caracterización de Arruda si entendemos el tipo de funtores a los que se refiere, como funtores que contrarían el significado de la proposición, esto es, no añaden información a la proposición sino que cambian la información en un sentido contrario al contenido por el argumento del functor.

3.1.5.1 El Sistema de da Costa

El señalamiento ¿reiterado en todas las definiciones anteriores es la capacidad de capturar la inconsistencia evitando que el sistema se vuelva trivial. Esta capacidad ha sido capturada mediante varios tipos de propuestas. Los primeros intentos se basaron en modalidad y polivalencia, los segundos y más concretos se dirigieron más directamente sobre las dificultades en la propia sintáxis del sistema¹⁵⁰. El problema básico es separar inconsistencia de trivialidad. Puede haber distintas formas de ataque directo al problema. Una de ellas consiste en eliminar la adjunción (regla de conjunción), para que, a su vez, esto impida la aplicación del principio conocido como el Principio del Pseudo Scoto o de su regla; otra, sin eliminar la adjunción, es construir una serie de esquemas de axiomas que sea incapaz de producir el Principio del Pseudo Scoto. Sistemas como este último son los más clásicos de la lógica paraconsistente: las jerarquías C_n , $1 \leq n \leq \omega$, de Newton da Costa. Veamos una breve introducción formal a estos cálculos¹⁵¹.

Los cálculos C_n , $1 \leq n \leq \omega$, tienen como símbolos primitivos, los símbolos de variables proposicionales, los conectivos $\neg, \wedge, \vee, \rightarrow, \supset$, y los paréntesis.

Definamos α° de la siguiente forma:

$$\alpha^\circ =_{\text{def.}} \neg(\alpha \wedge \neg\alpha),$$

Llamamos α° a "fórmula bien comportada".

Podemos asumir los siguientes axiomas-esquema para C_1 :

$$(dC1) \alpha \supset (\beta \supset \alpha)$$

$$(dC2) (\alpha \supset \beta) \supset ((\alpha \supset (\beta \supset \gamma)) \supset (\alpha \supset \gamma))$$

$$(dC3) \alpha \supset (\beta \supset \alpha \wedge \beta)$$

$$(dC4) \alpha \wedge \beta \supset \alpha$$

$$(dC5) \alpha \wedge \beta \supset \beta$$

¹⁵⁰ Cfr. [D'Ottaviano 1990]. D'Ottaviano, Loffredo "On the Development of Paraconsistent Logic and da Costa's Work", *The Journal of Non-Classical Logic (JNCL)*, Vol. 7, Num. 1-2, 1990 pp. 89-152.

¹⁵¹ Para su exposición seguiré [D'Ottaviano 1990].

$$(dC6) (\alpha \supset \gamma) \supset ((\beta \supset \gamma) \supset (\alpha \vee \beta \supset \gamma))$$

$$(dC7) \alpha \supset \alpha \vee \beta$$

$$(dC8) \beta \supset \alpha \vee \beta$$

$$(dC9) \neg\neg\alpha \supset \alpha$$

$$(dC10) \alpha \vee \neg\alpha$$

$$(dC11) \beta^\circ \supset ((\alpha \supset \beta) \supset ((\alpha \supset \neg\beta) \supset \neg\alpha))$$

$$(dC12) \alpha^\circ \wedge \beta^\circ \supset (\alpha \supset \beta)^\circ \wedge (\alpha \wedge \beta)^\circ \wedge (\alpha \vee \beta)^\circ$$

Y podemos asumir a la regla de *Modus Ponens* como la única regla de C_1 :

$$(MP) \alpha, \alpha \supset \beta / \beta$$

Ahora, introduciremos otros conectivos para construir los cálculos C_n , $0 < n < \omega$, basados en la idea de iterar la construcción de α° :

$$A^{\circ\circ} = (A^\circ)^{\circ} =_{\text{def}} \neg(\neg(A \wedge \neg A) \wedge \neg\neg(A \wedge \neg A)).$$

Por tanto, definimos:

$$A^n =_{\text{def}} A^{\circ\circ\dots\circ} \text{ (n veces),}$$

y, una fórmula estable hasta n , será:

$$A^{(n)} =_{\text{def}} A^\circ \wedge A^{\circ\circ} \wedge \dots \wedge A^n$$

Así, para construir cualquier cálculo C_n $1 < n < \omega$, generalizamos (dC11) y (dC12):

$$(dC11) \beta^{(n)} \supset ((\alpha \supset \beta) \supset ((\alpha \supset \neg\beta) \supset \neg\alpha))$$

$$(dC12) \alpha^{(n)} \wedge \beta^{(n)} \supset (\alpha \supset \beta)^{(n)} \wedge (\alpha \wedge \beta)^{(n)} \wedge (\alpha \vee \beta)^{(n)}$$

Al cálculo que sólo tiene los axiomas-esquema (dC1)-(dC10) y la regla de *Modus Ponens*, le llamamos C_ω .

En cada uno de los cálculos C_n $1 \leq n < \omega$, podemos definir una negación fuerte, \neg^* , como sigue:

$$\neg^*A =_{\text{def}} \neg A \wedge A^{(n)}$$

Podemos, finalmente, denotar por C_0 , al cálculo proposicional clásico.

Las características primordiales de estos cálculos son¹⁵²:

- 0) El Principio del Pseudo Scoto no es válido en los cálculos C_n , $1 \leq n \leq \omega$.
- 1) C_n , $1 \leq n \leq \omega$, son paraconsistentes.
- 2) C_n , $1 \leq n < \omega$, son finitamente trivializables, es decir, que hay una fórmula (no un esquema) F , tal que el sistema obtenido por medio de añadir F como un nuevo postulado, es trivial.
- 3) C_ω no es finitamente trivializable.
- 4) Si añadimos a C_1 , el Principio del *Reductio* no restringido, entonces C_1 colapsa en C_0 .

En vista de lo anterior, es posible tener todas las consecuencias lógicas de C_0 como válidas en C_n , $1 \leq n < \omega$, para las proposiciones bien comportadas de grado correspondiente y utilizándose la negación fuerte \neg^* . Y todos los teoremas del cálculo proposicional positivo intuicionista son también teoremas de C_1 . Todo cálculo en la jerarquía $C_0, C_1, \dots, C_n, \dots, C_\omega$, es estrictamente más fuerte que el cálculo que le sigue en la secuencia. Esto es así en el sentido de que para todo C_n y su correspondiente C_{n+1} , todo lo derivable en C_{n+1} es también derivable en C_n pero no viceversa, es decir, que C_{n+1} es subconjunto propio de C_n .

Presentaré ahora algunos teoremas importantes de los cálculos C_n , $1 \leq n \leq \omega$, seguiré la formulación y la numeración (cuando ésta aparezca) que tienen en **[D'Ottaviano 1990]**.

Teorema 3.1.1. Todas las reglas y esquemas válidos del cálculo positivo proposicional son válidos en C_n , $1 \leq n \leq \omega$.

Teorema sobre la "clasicidad" de las fórmulas bien comportadas: Todos los teoremas de C_0 son válidos en C_n , $1 \leq n \leq \omega$, para proposiciones bien comportadas.

Y tenemos también para los cálculo C_n , $1 \leq n \leq \omega$, el **Teorema de Deducción**:

Teorema 3.1.4. Si Γ es un conjunto de fórmulas, $\Gamma, A \vdash_{C_n} B$, si y solamente si, $\Gamma, \vdash_{C_n} A \supset B$.

Finalmente,

Teorema 3.1.9. Sea A una fórmula de C_0 , y A^* la fórmula obtenida a partir de A sustituyendo por \neg^* cada \neg . Si $\vdash_{C_0} A$, entonces $\vdash_{C_n} A^*$, $1 \leq n \leq \omega$.

¹⁵² Cfr. **[Arruda 1988]**.

3.1.6 Paraconsistencia y Racionalidad

Introducción

Una de las ventajas principales que ofrece la paraconsistencia es su capacidad para la representación de contradicciones y su consecuente poder para el análisis formal de teorías o procesos de una clase especial. Tales teorías y procesos tienen dos propiedades básicas: 1) pueden contener contradicciones de algún tipo, pero sin trivialización del sistema y, sin embargo, 2) es deseable una exploración de éstos que parta de perspectivas que no los califican de antemano como resultados de alguna clase de irracionalidad.

Existen muchas alternativas para evitar la clase de presupuestos a los que me he referido. Pueden sin embargo distinguirse 2 grandes paradigmas respecto de la paraconsistencia que vale la pena anticipar desde ahora. El paradigma generado por la escuela de paraconsistencia desarrollada por Newton da Costa en Brasil, y el generado por Richard Sylvan y Graham Priest en Australia. La característica básica del primero es que se ve a la paraconsistencia como una alternativa de modelación para teorías inconsistentes, sin comprometerse necesariamente con la justificación normativista o bien la promoción de las contradicciones. Lo cual se ve reflejado en la clase de formalismos que generan, que pueden caracterizarse como “paraconsistencia débil”, siguiendo una clasificación en **[D’Ottaviano 1990]**. Es importante también notar que estos formalismos tienen una enorme influencia del intuicionismo en matemáticas. Esta es la perspectiva propuesta por los formalismos de Newton da Costa y, en general, la tendencia del paradigma generado por él y por varios estudiosos trabajando en conexión con él en Brasil. Deben destacarse en este respecto los trabajos del *Centro de Lógica, Epistemología e Historia de la Ciencia* de la Universidad Estatal de Campinas, de San Pablo. Tal vez los lógicos más destacados en este ámbito son Ítala D’Ottaviano, Walter Carnielli y Marcelo Coniglio. Por el contrario, la característica del segundo paradigma, el de la escuela australiana se compromete con una justificación normativista de la contradicción y con su promoción en diversos ámbitos. Estos desarrollos son herederos de una posición relevantista de la lógica y pueden colocarse, en la clasificación bipartita de **[D’Ottaviano 1990]**, como “paraconsistencia fuerte”. A menudo conectan sus propuestas formales con teorías filosóficas originales sobre la inconsistencia, las contradicciones o la dialéctica.

Sin soslayar lo anterior, una exposición histórica del desarrollo de la paraconsistencia, puede hacernos notar diversas alternativas para la construcción de los formalismos paraconsistentes. En este camino de caracterización podemos distinguir claramente tres vías de solución: la vía semántica, la vía modal y la vía sintáctica de desarrollo. Y, en esta última vía, notar también la diferencia entre los dos paradigmas que se mencionaron en el párrafo anterior.

Así, fundamentalmente, la construcción de lógicas paraconsistentes ha sido efectuada desde dos perspectivas distintas. La primera, cuyo nacimiento se debe a

Jaśkowski en 1948 y 1949¹⁵³, utiliza las lógicas modales desarrolladas en los años veintes del siglo pasado. La segunda perspectiva realiza un cambio en la estructura sintáctica de los sistemas lógicos clásicos, y su consecuente cambio semántico, y fue fundada por da Costa.

En este apartado haré una breve exposición de las últimas dos vías mencionadas para la construcción de lógicas paraconsistentes y después haré algunas reflexiones acerca del sentido de la racionalidad implícita en una perspectiva paraconsistente particular.

3.1.6.1 La vía modal

El sistema creado por Jaśkowski fue construido sobre el sistema S5 de modalidad. Como he comentado antes, su desarrollo obedece al objetivo de representar discusiones entre sujetos con la posibilidad de que algunos de ellos sostengan tesis mutuamente inconsistentes. El sistema D2 de Jaśkowski fue desarrollado en notación polaca. Una reconstrucción de sus preliminares es presentada en el texto de D' Ottaviano de 1990¹⁵⁴. En D2 se pueden encontrar dos tipos de implicación y dos de equivalencia, la implicación y equivalencia materiales y la implicación y equivalencia discursivas (las últimas denotadas respectivamente por " \rightarrow_d " y " \leftrightarrow_d ") definidas como sigue [D'Ottaviano 1990, p. 107]:

$$p \rightarrow_d q = df \Diamond p \rightarrow q$$

$$p \leftrightarrow_d q = df (\Diamond p \rightarrow_d q) \wedge (\Diamond q \rightarrow_d p)$$

En el sistema de Jaśkowski una afirmación aceptada por todo sujeto se entiende como una afirmación que ocurre en todo mundo posible. Por el contrario, si se trata de una afirmación en la que hay divergencias entre los participantes en la discusión¹⁵⁵ las afirmaciones serán entendidas como afirmaciones que ocurren en al menos un mundo posible. Así, colocar el diamante modal a la izquierda de una fórmula significa que algún agente (o sujeto) la propuso como tesis en una discusión.

A raíz de que D2 se encuentra construido con base en el sistema modal S5 es fácil suponer que es derivable dentro del sistema la Regla del Pseudo Scoto, a saber,

¹⁵³ [Jaśkowski 1948] Jaśkowski, S., Rachunek zda dla systemów dedukcyjnych sprzecznych, publicado bajo el título "Un Calcul des Propositions pour les Systemes Déductifs Contradictaires" en *Studia Societatis Scientiarum Torunensis*, vol. 1 (5), sectio A, pp. 55-77. Y [Jaśkowski 1949] Jaśkowski, S., O konjuncji dyskusyjnej w rachunku zdán dla systemów dedukcyjnych sprzecznych, publicado bajo el título "Sur la conjonction Discursive das le Calculu des propositions pour les Systemes Contradictaires" en *Studia Societatis Scientiarum Torunensis*, vol. 1 (5), sectio A, pp. 171-2. Para una reconstrucción histórica más completa sobre este punto véase [Gomes 2013].

¹⁵⁴ [D'Ottaviano 1990] También es revisada, en este texto, la reconstrucción axiomática realizada por da Costa y Dubikajtis en 1968.

¹⁵⁵ Los conjuntos de creencias de dichos sujetos son entendidos como consistentes. Priest, G. y Routley (ahora Sylvan), R.; "Systems of Paraconsistent Logic", [Priest & Routley 1989a]; en Priest, Graham; Routley, Richard; Norman, Jean (eds.), *Paraconsistent Logic. Essays on the inconsistent*. Philosophia Verlag, Alemania, 1989, p. 158 [Priest & Routley & Norman, 1989].

$(p \wedge \neg p) \rightarrow q$

En muchos otros sistemas la derivación de este principio pone en peligro de trivialización al sistema de Jaśkowski, pues hemos visto que dos agentes podrían proponer tesis mutuamente contradictorias.

Recordemos que el bloqueo de la trivialización del sistema tiene que ver con las reglas de conjunción, simplificación, adición y silogismo disyuntivo o algunas de sus equivalentes¹⁵⁶. La formalización de la discusión como un todo, incluyendo tesis contradictorias, requerirá entonces de algún mecanismo para bloquear la trivialización del sistema. Como vimos antes, la forma natural de efectuar dicho bloqueo es evitar la derivación del Principio del Pseudo Scoto. En este caso, Jaśkowski prefiere preservar el principio pero inutilizarlo en cierta medida, omitiendo la regla de conjunción:

$p, q / p \wedge q$

La justificación para la omisión de esta regla tiene una motivación principal: la representación de un proceso de discusión. En el proceso de una discusión no deberíamos confundir las distintas posiciones de los participantes asumiendo que se deben sostener ambas al mismo tiempo, como si se tratase de una sola posición respecto del tema que se encuentran discutiendo.

Algunos lógicos han opinado¹⁵⁷ que prescindir de una regla tan intuitivamente correcta como la conjunción es el problema más grave del sistema D2. No obstante, tomando en cuenta la justificación a la que he aludido antes, parece que el argumento más fuerte que puede rescatarse de la discusión es que, en realidad, el sistema de Jaśkowski no es (dado que omite la conjunción) una herramienta suficiente para el análisis de una verdadera situación inconsistente. Para decirlo más claro, en el sistema D2 no podemos representar inconsistencias internas del sujeto o de algún sistema de conocimiento, sin trivialidad, sino únicamente divergencias entre distintos sujetos.

Ahora bien, esta opinión no es muy justa. Por una parte, no debería minimizarse la utilidad del sistema de Jaśkowski con base en la suposición de que tiene que ser capaz de representar antinomias¹⁵⁸ sin trivialidad. Si hubiera de atacarse este punto deberá ser con base en la perspectiva de Jaśkowski, de otra manera la discusión pierde mucho sentido. La intención de Jaśkowski no es modelar inconsistencias al interior de un mismo conjunto de conocimientos, esto produciría trivialidad en su sistema. Lo que él quiere es tan sólo modelar posiciones mutuamente contradictorias en distintos conjuntos de conocimiento

¹⁵⁶ De acuerdo al argumento de Lewis expuesto cuando tratamos el caso de la explicación de Frege.

¹⁵⁷ Por ejemplo como sugieren Priest y Sylvan (antes Routley) en su "First Historical Introduction, A Preliminary History of Paraconsistent or Dialectic Approaches" [Priest & Routley 1989], en [Priest & Routley & Norman, 1989]. P. 48.

¹⁵⁸ El término "antinomia" es usado para denotar contradicciones del tipo $p \wedge \neg p$.

en una discusión. La intención de Jaśkowski no es la de comprometerse con la existencia de afirmaciones que sean, a la vez, contradictorias en sí mismas y racionales¹⁵⁹. Desde este punto de vista, la omisión de la regla de conjunción es algo sumamente intuitivo.

En realidad lo que es más interesante en la discusión acerca de los alcances del sistema de Jaśkowski y, en general, de cualquier sistema de lógica paraconsistente son los supuestos filosóficos que se encuentran en el fondo. Con ello la discusión toca también el problema de cuáles podrían ser los principios mínimos (si es que éstos existen) de toda racionalidad o, al menos, de lo que entendemos por "racionalidad occidental". En este sentido sería iluminador realizar un análisis histórico de la suposición del Principio de No Contradicción en el paradigma epistémico subyacente a diferentes teorías, pero es algo que no haré en éste trabajo.

3.1.6.2 La vía sintáctica

3.1.6.2.1. La versión moderada

La construcción de la paraconsistencia en la jerarquía de sistemas de da Costa ocurre por dos flancos. Por una parte, desde la sintaxis, la restricción para aplicar el Principio de Reducción al Absurdo (esquema de axioma 11 y similares) y la imposibilidad de inferir el Principio de No Contradicción. La inferencia de una fórmula arbitraria (trivialización) sólo es posible en los casos de fórmulas bien comportadas. Por otra parte, también desde la semántica se producen condiciones permisivas para la introducción de inconsistencias: las valuaciones proporcionadas para las fórmulas no prohíben el caso de que una fórmula y su negación sean verdaderas.

La posición que puede entreverse de los formalismos propuestos por da Costa es decididamente más comprometida con el supuesto de la aceptación de inconsistencias en ciertos sistemas de conocimiento en comparación con el sistema de Jaśkowski, quien no se propone modelar contradicciones en el sistema de creencias del sujeto sino sólo como resultado de la interacción entre un grupo de sujetos en el planteamiento de sus tesis. Así, la jerarquía de cálculos de da Costa acepta también claramente la derivación de antinomias, por lo que se puede considerar cabalmente como una lógica paraconsistente desde la perspectiva particular de Lorenzo Peña.

Debe observarse, sin embargo, que las condiciones de verdad de la valuación para los cálculos de da Costa impiden considerar al functor de negación usado en los sistemas C_n paraconsistentes en el mismo sentido en que es considerada la negación en los sistemas clásicos. Explicaré enseguida esta última afirmación.

¹⁵⁹ Todo lo cual, desde la clasificación propuesta al principio implica que el sistema de Jaśkowski no puede ser considerado como una lógica contradictorial, aunque sí inconsistente.

La valuación para los cálculos C_n paraconsistentes no permite determinar, a raíz de que la negación de una fórmula tenga valor 1, el valor de la fórmula sin negación. Lo anterior significa que es posible el caso en que una fórmula tenga valor 1 mientras que su negación posee el mismo valor, es decir, que una antinomia puede resultar verdadera. Pero también se deduce de ello que no podemos tener el caso en que ambas, una fórmula y su negación, tengan valor 0. Dicho de otra forma, para toda fórmula de los cálculos C_n , o bien ella o su negación es verdadera, pero no tenemos una restricción en contra de que ambas lo sean. Por supuesto, una posible interpretación de esto es que los cálculos de da Costa pueden modelar las contradicciones, considerándolas verdaderas. Y esto puede entenderse extra-lógicamente, más bien desde una posición ya epistemológica, como el hecho de que los cálculos de da Costa pueden modelar la inconsistencia de transición que a veces sucede en el cambio científico. Una inconsistencia que quizá puede después ser borrada de la teoría en cuestión, teniendo nueva información disponible.

Debe notarse que la valuación permisiva de inconsistencias está dada para el functor de negación débil de los cálculos C_n , a saber, para " \neg ", y que el functor de negación fuerte (" \neg^* ") se encuentra definido en términos de aquel. Por ello, cualquier antinomia derivada dentro de los cálculos C_n podrá verse como una contradicción que tiene como functor correspondiente la negación débil de da Costa, lo cual puede producir un efecto desalentador. Si alguien quisiese interpretar la jerarquía de cálculos de da Costa como sistemas que son capaces de formalizar lo podría atinar en llamar "verdaderas contradicciones", entonces esperaríamos que el sistema pudiera sostener contradicciones como ésta:

$$A \wedge \neg^* A$$

Lo anterior quiere decir que alguien podría tener la expectativa de que los cálculos de da Costa fueran capaces de expresar la verdad de una fórmula y, al mismo tiempo, la verdad de la negación fuerte de ésta. Es decir que, como consecuencia de una derivación de este tipo, y bajo una valuación clásica, tuviéramos que una fórmula es verdadera y falsa a la vez.

Sin embargo parece que existe una dificultad de fondo. Si quisiéramos modelar "verdaderas contradicciones", entonces basta un poco de análisis para notar lo que pasaría con el Principio de Pseudo-Scoto, a saber,

$$(A \wedge \neg A) \supset B$$

Lo que ocurriría es que el antecedente de este condicional sería falso y, dadas las condiciones de verdad de un condicional, el principio sería, en automático, verdadero. Cosa que afectaría la virtud paraconsistente de estos sistemas lógicos, pues haría que fueran más bien sistemas explosivos.

Esta es una de las razones por las cuales se requiere que la contradicción sea verdadera. ¿Qué solución podría darse a esto? Una consideración importante, desde mi punto de vista, es que la intuición puede inclinarnos a pensar que un científico no estaría dispuesto a admitir las contradicciones como algo verdadero, por lo menos en muchas de las ocasiones en las que se encuentra con ellas. Esta reflexión puede llegar a ser importante para la construcción de sistemas paraconsistentes que puedan contener esta característica: sus contradicciones son siempre falsas pero de ahí no se sigue cualquier fórmula. Debe tomarse en cuenta también que esta característica no sólo coincide intuitivamente con la práctica científica en el caso de considerar que ninguna contradicción es verdadera (observemos que esto no necesariamente implica que no pudiéramos usar como verdaderos los conjuntos de esa contradicción), sino también, coincide (y es igualmente importante) en que no se acepta explosividad: es también antiintuitivo pensar que un científico, al encontrarse con una contradicción legítima, considere que su teoría ha explotado en el sentido lógico que hemos definido¹⁶⁰.

Estas reflexiones deberían conducirnos a nuevos retos formales para construir sistemas paraconsistentes que modelen adecuadamente el cambio científico. Pero en este trabajo no abordaré estos retos, los dejaremos para una futura investigación.

Sin soslayar las reflexiones anteriores, pensándolo analíticamente no debería haber tanto problema al aceptar que lo que modela las contradicciones en la Ciencia, desde las jerarquías de da Costa, es un functor de negación débil. Independientemente de cuál sea el functor que le corresponda más fielmente a las negaciones supuestas en Ciencia en el caso de contradicciones, hay, por lo menos transitoriamente, contradicciones en los cuerpos teóricos científicos, y una forma adecuada de modelarlas es considerándolas verdaderas en el sistema (a través de la conjunción de la afirmación y la negación débil de una misma fórmula), en el sentido extra-epistémico de considerarlas, quizá, como casos de paraconsistencia de transición.

Así, en las investigaciones sobre paraconsistencia efectuadas por el grupo brasileño las antinomias vienen acompañadas de estatutos no sólo sintácticos sino también semánticos que permiten de antemano su aparición. El Principio de No Contradicción no es válido aunque sí rige en algunas fórmulas, las fórmulas bien comportadas.

Una posición radical aspira a fórmulas "verdaderamente rebeldes", que fueran falsas y verdaderas al mismo tiempo.

3.1.6.2.2. La versión radical

La aspiración a la que me he referido en el párrafo anterior cristalizó en propuestas del grupo australiano de investigación sobre paraconsistencia, particularmente las de

¹⁶⁰ Esta reflexión surgió, conjuntamente, en una interesante discusión sobre el tema, que tuve con Alicia Pazos.

Graham Priest y Richard Sylvan. Estas propuestas pueden pensarse como una alternativa dentro de la vía sintáctica para la paraconsistencia.

Los sistemas construidos por Sylvan y Priest se encuentran fundados en nociones relevantistas de implicación. Al menos la mayoría de los sistemas de lógica relevante pueden verse como clases especiales de sistemas paraconsistentes. Esto se debe a que el principal ataque de los lógicos de la implicación relevante está vinculado con el rechazo del silogismo disyuntivo como regla aplicable en nociones de consecuencia consideradas más fieles al tipo de inferencia que usamos realmente en nuestros sistemas de conocimiento, más fieles, por así, decirlo, a la noción de implicación relevante. Presentaré¹⁶¹ enseguida algunos de los axiomas del sistema DM¹⁶², un ejemplo de lógica paraconsistente basada en relevancia¹⁶³:

Dados los conectivos " \rightarrow ", " $\&$ ", y " \neg " como primitivos, y la disyunción de A y B, a saber, " $A \vee B$ ", definida como $\neg(\neg A \& \neg B)$, tenemos:

- D1. $A \rightarrow A$
- D2. $[(A \rightarrow B) \& (B \rightarrow C)] \rightarrow (A \rightarrow C)$
- D3. $A \& B \rightarrow A$
- D4. $A \& B \rightarrow B$
- D5. $[(A \rightarrow B) \& (A \rightarrow C)] \rightarrow (A \rightarrow (B \& C))$
- D6. $(A \& (B \vee C)) \rightarrow [(A \& B) \vee (A \& C)]$
- D7. $\neg\neg A \rightarrow A$
- D8. $(A \rightarrow \neg B) \rightarrow (B \rightarrow \neg A)$
- D9. $(A \rightarrow B) \rightarrow \neg(A \& \neg B)$

Para el caso que nos ocupa, el cambio de los conectivos primitivos en esta lista de axiomas, respecto de la lógica clásica, no tendría sentido sin la interpretación del condicional como algo distinto de la implicación material. Como consecuencia de ello, observemos, el signo " \rightarrow " no es definido en términos de disyunción.

El sistema posee además una tesis contradictoria postulada como axioma:

- D10. $p_0 \& \neg p_0$

¹⁶¹ Con base en la presentación que aparece en [D'Ottaviano 1990].

¹⁶² Presentado por Routley (ahora Sylvan) y Meyer en 1976. [Routley & Meyer 1976] Routley, R. & Meyer, R. K. "Dialectical Logic, classical logic and the consistency of the world", *Studies in Soviet Thought*, Núm. 16, 1976, pp. 11-25.

¹⁶³ Arruda, Ayda, "Aspects of the Historical Development of Paraconsistent Logic" [Arruda 1989], en [Priest & Routley & Norman, 1989]. DM es una abreviación de "Minimal Dialectical Logic" como parece sugerir el comentario citado por Arruda en la página 109.

El axioma D10 es un caso que viola el Principio de No Contradicción. Sin embargo, el principio es válido en el sistema DM. Para explicitar correctamente que es válido el Principio de No Contradicción enunciaré a continuación las reglas del sistema:

RD1. $A, A \rightarrow B \Rightarrow B$

RD2. $A, B \Rightarrow A \& B$

RD3. $A \rightarrow B, C \rightarrow D \Rightarrow (B \rightarrow C) \rightarrow (A \rightarrow D)$

Así, la derivación del Principio de No Contradicción quedaría como sigue:

- | | |
|---|----------------------------|
| (1) $(A \rightarrow A) \rightarrow \neg(A \& \neg A)$ | instancia de D9 |
| (2) $(A \rightarrow A)$ | instancia de D1 |
| (3) $\neg(A \& \neg A)$ | por RD1 mediante (1) y (2) |

Intuitivamente, después de la asunción del Principio de No Contradicción, una valuación adecuada para los conectivos de DM es la valuación clásica. Sin embargo ello desembocaría en una semántica contradictoria: p_0 es verdadera y falsa a la vez. Por tanto se añade un valor derivado para estos casos: el valor *verdadero-y-falso*.

En cierto sentido el functor de negación recoge el sentido de la noción clásica de negación a pesar de la adición de este nuevo valor híbrido en la semántica para DM. Por ello, este tipo de sistemas suponen una aceptación más franca del supuesto de que podemos tener sistemas de conocimiento que contienen contradicciones. En el sistema DM la validez del Principio de No Contradicción es indiscutible y sin embargo el sistema es capaz de contemplar al menos un caso de contradicción pero, también, su consecuente destrozamiento a nivel semántico: la adición del valor *verdadero-falso*. A pesar de lo anterior, DM no es lógicamente trivial.

La idea que parece subyacer, en los desarrollos de Lógica paraconsistente del grupo australiano, es una concepción que admite contradicciones a nivel ontológico. Esta idea básica parece intentar modelar contradicciones en el mundo, en un contexto epistémico que no admitiría contradicciones. Reflexionemos un poco al respecto. Un supuesto ontológico de esta naturaleza es adecuado para teorías de la realidad, como por ejemplo la que uno puede suponer en Hegel, en las que está implicado que los eventos tienen su respectivo evento contradictorio; donde, podría interpretarse, la contradicción tiene un papel en el devenir ontológico. Epistemológicamente por lo menos, suponer una teoría de esta naturaleza como, por ejemplo, un paradigma general de la Ciencia, sería desastroso. Aún a pesar de no suponer trivialización en la teoría que contuviera una contradicción, la identificación de una contradicción en una teoría científica tiene también un rol heurístico importante: nos invita a revisar los supuestos de los que partimos, nos invita, en general, a buscar un error en nuestros cálculos o en nuestros supuestos, o, incluso, en nuestra práctica de recolección de la evidencia. En esta línea de pensamiento, no creo que se trate de eliminar la restricción que nos impone el Principio de No Contradicción de la Lógica Clásica, sino tan sólo de construir sistemas que nos permitan

analizar los casos en los que la contradicción no ha podido evitarse, sea esto porque tenemos una dificultad a nivel ontológico o, tan solo, una a nivel epistémico.

Ahora bien, aunque la admisión del valor verdadero-falso, puede interpretarse como una forma de paraconsistencia de transición, resulta por lo menos artificial postular la contradicción desde los axiomas del sistema. Esto implica tener, de antemano, identificada la fórmula que produce la inconsistencia en la teoría que se intenta modelar. Los cálculos de la jerarquía de da Costa, en cambio, permiten modelar una teoría en la que no se sabe si tendrá contradicciones o no. Cuando la contradicción aparece, por ejemplo, añadiendo los axiomas propios de la teoría en cuestión, entonces por lo menos en algún nivel no podrá demostrarse el esquema de no contradicción para ella y, por tanto, no será una fórmula bien comportada en el sistema. Estos cálculos, pueden modelar más adecuadamente lo que he llamado inconsistencia de transición, pues si hay una modificación en los postulados propios de la teoría, tal inconsistencia no aparecerá más, y no tendremos para ello que hacer una modificación en los axiomas del sistema. Así, la jerarquía de cálculos de da Costa es por lo menos más adecuada en este respecto para modelar un cambio científico como el que implica una modificación en los postulados de la teoría en cuestión.

3.1.6.3 Algunas Consecuencias

Así, la investigación sobre la paraconsistencia, además de conformarse por estrategias de acuerdo a las diferentes perspectivas señaladas, posee también distintas formas de relación con compromisos epistémicos y ontológicos. Particularmente, sobre la línea histórica Jaśkowski-daCosta-Sylvan, puede observarse un gradual desapego al Principio de No Contradicción como principio fundamental de la racionalidad. En consecuencia, puede decirse que la paraconsistencia pone algo más en la mesa de discusiones que quizá tenga algunas repercusiones importantes. Miró Quesada, por ejemplo, observó que:

"el hecho de haber elaborado un sistema que funcione con eficacia, sin trivializarse, cuando se aplica a teorías inconsistentes revela algo fundamental: que la inconsistencia de una teoría no significa, necesariamente, su derrumbe como teoría. Y esto sólo puede significar una cosa: que la manera de funcionar de nuestra razón es muy diferente de lo que, tradicionalmente, se había supuesto. Es algo mucho más complejo y profundo, que no cabe dentro de los esquemas clásicos. La creación de sistemas como C_n no significa que nuestra razón, para funcionar adecuadamente, tenga que producir contradicciones mediante un dinamismo interno (tesis dialéctica tradicional), pero sí que, en caso de que así fuera, o de que el mundo funcionara dialécticamente, existe la posibilidad de

elaborar un sistema lógico riguroso que pueda aplicarse en relación a estas situaciones."¹⁶⁴

Sin duda, esta es la discusión de mayor interés que propone la paraconsistencia a la Filosofía de la Lógica, pero no sólo a ella, también a la epistemología y a la ontología en términos generales.

Sin embargo, existe un peligro en la consideración de sistemas de conocimiento paraconsistentes y se refiere a la fácil mutación de esta postura a un relativismo. Esto se puede apreciar cuando se toma conciencia de la posibilidad de transgredir, por ejemplo, el Principio de Reducción al Absurdo. Dejar pasar una contradicción sin la consecuente negación de alguno de los supuestos de los que se ha partido en una discusión significa perder, si no totalmente, al menos una gran parte de nuestra seguridad racional en el conocimiento.

El problema planteado en el párrafo anterior, en relación a la reducción al absurdo, ha sido analizado por Graham Priest¹⁶⁵. Él hace una distinción entre dos nociones de reducción al absurdo. Esta diferencia puede plantearse (aunque él usa otros términos) como la separación entre el uso de la reducción al absurdo en el contexto justificacional y su utilización en contextos creenciales. Su opinión puede expresarse mediante dos afirmaciones: 1) en el contexto creencial, el cambio de creencia con base en la demostración de contradicciones dentro de la propia postura del agente no constituye un cambio necesario ni suficiente, 2) en el contexto justificacional, existen secciones de los sistemas teóricos que no aceptan la paraconsistencia y otros que sí la aceptan. En los primeros es viable la prueba por reducción al absurdo, en los segundos, no.

Sin embargo aún no se ha ofrecido una respuesta adecuada a este problema, pues la pregunta inmediata ante soluciones del tipo que se ha tratado arriba es qué tipo de criterios nos ayudarían a discernir entre situaciones en las que debemos dar crédito a una prueba que supone el Principio de No Contradicción y situaciones que tienen que ver con supuestos paraconsistentes.

Con todo, si queremos realizar análisis formales de sistemas y procesos de conocimiento de modo que se aproximen lo mejor posible a la estructura de sus referentes, será casi imposible prescindir de ciertos supuestos paraconsistentes¹⁶⁶.

Podemos identificar a la explicación científica como un sistema del tipo que he estado mencionando, es decir un sistema que en ocasiones incluye inconsistencias. Como

¹⁶⁴ Miró Quesada, Franciso, "La lógica paraconsistente y el problema de la racionalidad de la lógica" [Miró 1988] p. 615, en [CELLIS, 1988]. Las itálicas son de Miró Quesada.

¹⁶⁵ Priest, Graham, "Reductio ad Absurdum et Modus Tollendo Ponens" [Priest 1989], en [Priest & Routley & Norman, 1989].

¹⁶⁶ Pueden tomarse como evidencia a favor de esta afirmación, los casos tratados en [Priest & Routley & Norman, 1989] y en [Carnielli & Coniglio & Marcos 2003].

expuse antes, la explicación científica puede entenderse como una estructura dinámica que se encuentra involucrada con información inconsistente; involucra por lo menos inconsistencias de transición relacionadas con su *explanans*.

Como se vio antes, la explicación de Frege de la noción de número contenía un *explanans* con supuestos inconsistentes. La teoría que servía de puntal para la explicación era contradictorial en el sentido de Peña y, sin embargo, la estrategia de explicación fue lo suficientemente prometedora como para que Russell se esforzara en reivindicarla (aunque con ciertos cambios importantes) más de una década después. Por otro lado, la base del *explanans*, en este mismo caso, constituyó una herramienta muy fructífera para ciertos avances en otras áreas de la matemática aún a pesar de que ya se tenía conciencia de que era una teoría inconsistente.

De esta manera, al menos transitoriamente, se conservó intacta la teoría de conjuntos de Cantor para la explicación Fregeana. Esto sucedió como si se hubiera dictado sentencia sobre el "mal comportamiento" (al estilo de da Costa) del supuesto de que algún conjunto puede pertenecerse a sí mismo, y como si se hubiera activado, simultáneamente, un bloqueo contra la explosión de fórmulas que ello provocaría con un sistema de lógica clásica subyacente a la teoría. La trivialidad de la teoría no ocurre en este tránsito a la modificación de los supuestos del *explanans*.

Modelar esta inconsistencia de transición mediante lógica clásica perdería matices interesantes del objeto de estudio. En primera instancia, en una situación de trivialidad no podrían modelarse claramente las relaciones de apoyo entre las proposiciones contenidas en la explicación. Será útil entonces intentarlo a través de herramientas lógicas paraconsistentes.

3.2 El Razonamiento del Sentido Común.

He mostrado que, si incluimos el contexto en la representación de la explicación, habrá una característica que resulta difícil modelar para la explicación científica: la paraconsistencia de transición. ¿Cómo articulamos esta representación en el contexto dinámico en el que hemos estado caracterizando la explicación? Para abordar esto es conveniente que revisemos desde su base pragmática la perspectiva dinámica desde la que se puede desarrollar esto. La perspectiva que intento construir, puede verse como un

intento de volver a la idea natural de explicación, que tiene mucha similitud con el funcionamiento del razonamiento del sentido común.

¿Cómo razonamos en el sentido común? ¿Cuáles son las diferencias entre nuestra forma de razonar en el sentido común y los razonamientos deductivos que modelamos en un curso de lógica clásica elemental? Algunos de los científicos de la computación que han estado enfocados en la producción de inteligencia artificial, en la segunda mitad del siglo XX, consideraron algunas similitudes pero también algunas sorprendentes y prometedoras diferencias. En esta sección ofreceré una introducción informal a una de las lógicas contemporáneas que se pretende que pueden modelar el razonamiento del sentido común: la lógica del razonamiento por *default*, de Raymond Reiter [Reiter 1980]¹⁶⁷. En la primera parte de esta sección, haré una caracterización intuitiva de las principales propiedades de los esquemas propuestos por Reiter para este propósito de modelación.

3.2.1 Algunas intuiciones sobre el razonamiento del sentido común.

A pesar de que por lo general coloquialmente omitimos indicaciones que refieran al tipo de inferencia usado en nuestros razonamientos, a veces indicamos la presencia de conclusiones obtenidas a través de un razonamiento de sentido común mediante frases como: "debe de suceder que", "normalmente", "es de sentido común que", "seguramente que", "es obvio que", "típicamente", "es posible que", "es de esperarse que", "podemos suponer que", "cabe pensar que", etc. ¿Es posible alguna elucidación formal de este tipo de matices de la relación de inferencia? Raymond Reiter pensó que sí.

Comencemos con algunos ejemplos. Supongamos que tocan el timbre de mi casa y que no espero a nadie. Llamemos a esto "Situación 1". Mi razonamiento inmediato puede

¹⁶⁷ Reiter, R., "A Logic for Default Reasoning", en *Artificial Intelligence*, Vol. 13, 1980, pp. 81-123. Reimpreso en Matthew L. Ginsberg, (comp.), *Nonmonotonic Reasoning*, Morgan Kaufmann Publishers, E.U.A. 1987.

ser: "No espero a nadie, debe de ser un vendedor". Este mismo razonamiento puede representarse sólo proposicionalmente como sigue:

(A) "Dado que no espero a nadie y han tocado el timbre de mi casa, debe de ser un vendedor."

O, más esquemáticamente,

(A')

No espero a nadie en casa.

Alguien ha tocado el timbre de mi casa.

Ese alguien es un vendedor.

Este argumento ilustra una de las características más peculiares que se suelen atribuir al razonamiento del sentido común: **el a la conclusión**. Las premisas de A' no son suficientes para inferir con certeza deductiva que la persona que se encuentra tocando el timbre es un vendedor. En un razonamiento de sentido común parece que partimos de premisas insuficientes para nuestra conclusión. Llamemos a este salto a la conclusión "**conjeturar**". Desde el punto de vista clásico de derivación, la conclusión de A' es aventurada. Llamemos a las proposiciones obtenidas de este modo "**conjeturas**".

Supongamos ahora que yo padezco de un odio especial contra los vendedores que llaman a mi puerta. A menudo el odio dota de perversidad y en esta suposición no es un caso de excepción. Así, supongamos que he diseñado un primitivo artefacto colocado encima de la puerta de mi casa tal que es capaz de vaciar un balde de agua fría exactamente en el lugar en donde estaría colocada una persona cuando toca el timbre. Para fantasear aún más, pensemos que el artefacto puede activarse con sólo mover una

palanca empotrada en la pared, muy cerca de mi computadora. Yo me encuentro intentando terminar mi tesis de maestría y trabajando en ella con mi computadora. No espero a nadie y de pronto escucho el timbre de mi casa. Bajo estas suposiciones, que llamaremos "Situación 2", y considerando A' como un razonamiento usual en mi cabeza, parece seguirse que activaré la palanca y que un vendedor saldrá perjudicado.

Ahora bien, si yo fuera más ingenioso y un poco precavido (o quizá menos perverso) hubiera diseñado otro artefacto para verificar, mediante un sistema de cámaras de video por ejemplo, si en efecto se trata de un vendedor o no. Sin embargo, el hecho de que yo intente verificar algo así podría implicar, además de mi paranoia, que el mío no es un razonamiento extraño al tipo de derivación supuesto en lógica clásica, es decir, podríamos modelar este nuevo razonamiento del siguiente modo deductivo:

(B)

No espero a nadie en casa.

Alguien ha tocado el timbre de mi casa.

Cuando no espero a nadie y tocan el timbre, entonces es posible que se trate de un vendedor.

Es posible que se trate de un vendedor.

En contraste, el razonamiento A' no lo es. La conclusión de A' no tiene fundamento suficiente en las premisas correspondientes, mientras que las premisas de B son suficientes para su conclusión.

Conjeturar no parece ser, tal como lo hemos caracterizado, llegar deductivamente a conclusiones posibles. En un sentido más general, para conjeturar no basta con llegar a

conclusiones que contengan operadores que atenúen la certeza de las mismas. Hace falta más bien que nuestra noción de consecuencia sea capaz de llevarnos, a partir de premisas insuficientes, a la franca afirmación¹⁶⁸ de la conclusión. Por ejemplo, veamos el “extraño caso” del ornitorrinco. Pensemos por ejemplo que estamos en la situación de considerar a un ornitorrinco, como un mamífero:

(C)

Los mamíferos no ponen huevos.

El ornitorrinco es mamífero.

El ornitorrinco no pone huevos.

La conclusión de C es errónea. Si C es tomado como un razonamiento deductivo, o bien tendríamos que apelar a su incorrección (negando la verdad de alguna de sus premisas) o bien nos veríamos *racionalmente* obligados a aceptar, contra nuestro conocimiento, su conclusión¹⁶⁹. Después de todo parece una regla adecuada inferir, cuando uno se encuentra con un mamífero, que dicho animal no pone huevos. No obstante, también *racionalmente*, uno no debería conformarse con una conclusión que va en contra de nuestro conocimiento, en este caso, nuestro conocimiento de los extraños ornitorrincos. ¿Cómo modelamos ambas intuiciones de racionalidad? Hagamos el siguiente ensayo. La regla podría expresarse con el siguiente esquema de razonamiento:

(D)

x es mamífero

x no pone huevos

¹⁶⁸Aunque esta afirmación sólo poseyera una certeza transitoria.

¹⁶⁹Dado que los ornitorrincos, a pesar de ser mamíferos, no ponen huevos.

El esquema D expresaría un tipo de razonamiento de la misma clase que el que ejemplifica A'. D y A' son **conjeturales**¹⁷⁰. A la luz de D, el razonamiento en C podría reinterpretarse entonces como:

(C')

El ornitorrinco es mamífero.

El ornitorrinco no pone huevos.

Sin embargo, la conclusión es fallida y de este modo D y C' ilustrarían cómo en un razonamiento de sentido común se puede llegar a conclusiones que pueden no ser ciertas y que esta incertidumbre no se debe a la corrección de los esquemas (no se debe necesariamente a que alguna de sus premisas sea falsa o incierta). Podemos decir, aludiendo a esta característica, que D y C' son **fallibles**. Para dejar esto más claro, comparemos a D con el siguiente esquema:

(E)

x es tigre

x es felino

Podemos ver a E como la expresión esquemática del hecho de que todos los tigres son felinos. E no parece tener ninguna excepción: siempre que se cumple el requisito antecedente del esquema E (x es tigre), debe cumplirse su consecuencia (x es felino). Así, E funciona entonces, por definición, como funcionan los esquemas deductivos,

¹⁷⁰En el sentido de que saltan a la conclusión.

independientemente de que su formulación no sea un esquema válido¹⁷¹. Funciona, como caso particular, de modo que siempre que su premisa es verdadera (siempre que el caso de sustitución de x, haga verdadera a su premisa), la conclusión lo será también. Este tipo de funcionamiento no lo tiene D, valga observar que ha sido interpretado como un esquema que tiene conclusiones falibles, es decir, que no siempre que se cumple su premisa, se cumplirá su conclusión.

Bien podríamos haber formulado el razonamiento por sentido común D, del mismo modo que E, nada nos impide ponerlo como un esquema en un sistema lógico; nada excepto el dictado de nuestra racionalidad. Si bien D puede ser formulado de modo que funcione como lo hacen los esquemas deductivos, tendríamos el inconveniente de concluir que un ornitorrinco no pone huevos. El inconveniente es de alguna importancia pues, formulando D deductivamente, no podríamos deshacernos nunca más de esta conclusión. Los razonamientos deductivos no tienen excepciones. Si se le quiere ver como una actividad racional, una forma de hacerlo es interpretando a D como un razonamiento a partir de esquemas generales que tienen **excepciones** de aplicación. D en este caso funcionaría, por ejemplo, cuando no se trata de un monotrema¹⁷².

Hasta aquí, A', D y C' muestran dos características principales del razonamiento del sentido común:

- i) Que es conjetural.
- ii) Que es falible.

¹⁷¹ Podríamos haber formulado E como un caso de *Modus Ponendo Ponens* (Si x es un tigre entonces es felino, x es un tigre, por lo tanto x es felino).

¹⁷²"Monotrema" es el nombre que se asigna en zoología al grupo de mamíferos que ponen huevos.

Pero hemos visto, además, que una interpretación racional posible de esquemas como estos es considerarlos como razonamientos con excepciones.

No todo razonamiento conjetural y falible es un buen razonamiento de sentido común. Como ejemplo volvamos al razonamiento A'. Supongamos ahora que existe otra persona en el mundo que odia tres veces más que yo a los vendedores cuando llaman a la puerta. En un arranque de locura, esta persona se dedica, por una terrible decisión, a asesinar vendedores en toda la ciudad. Supongamos también que, como era de esperarse por estos constantes y misteriosos homicidios, la profesión de vendedor desapareciera en la ciudad dado que se expandió el temor de ser asesinado por el psicópata susodicho. Llamemos a este estado de cosas "Situación 3". Asumamos que la Situación 3 comprende también la situación descrita antes para el razonamiento A', a saber, la Situación 1, donde no espero a nadie y alguien toca el timbre. Si alguien toca el timbre y no espero a nadie, concluyo que el que toca el timbre es un vendedor, según mi esquema general de razonamiento de sentido común. No obstante, parecería absurdo llegar a una conclusión como ésta dada la fama del psicópata. ¿Cómo hacemos que nuestro modelo de razonamiento distinga entre una situación y otra?

Si queremos que nuestra representación del sentido común sea más racional, se requerirá de algún mecanismo que justifique el salto a la conclusión. No toda conjetura puede ser aceptable. Así, siguiendo un uso racional de nuestros modelos de razonamiento, tenemos tres características:

- a) El razonamiento del sentido común es conjetural y falible.
- b) En los razonamientos conjeturales, no debería permitirse cualquier conclusión.

La última de estas consecuencias será el punto de partida para el resto de este trabajo, produce la tentativa de modificar nuestros razonamientos conjeturales de forma que representen formas racionalmente justificadas de razonamiento.

La **justificación** mínima que tenemos para permitir que un razonamiento conjeture una cierta conclusión es que al efectuar la inferencia no se transgreda el Principio de No Contradicción. Lo que está mínimamente justificado a inferirse es aquello que no contradiga el resto de nuestro conocimiento de fondo.

Volviendo a nuestros casos anteriores, si optamos por esta justificación mínima, el razonamiento A' no podría efectuarse en la Situación 3 ya que asume por sentido común que la persona que hizo funcionar el timbre es un vendedor y nuestro conocimiento de fondo nos indica, en un contexto trágico para el mercantilismo, que no hay vendedores en la ciudad. Efectuar la inferencia de A' produciría, junto con nuestro conocimiento de fondo, una contradicción, a saber, hay un vendedor en la ciudad y no hay un vendedor en la ciudad.

La representación de las justificaciones de una inferencia de sentido común es algo que debe examinarse por lo menos brevemente. Supongamos que quisiéramos representar la justificación de un argumento de sentido común, como aparece en el siguiente razonamiento:

(F)

No espero a nadie en casa.

Alguien ha tocado el timbre de mi casa.

Ese alguien es un vendedor.

Ese alguien es un vendedor.

La forma obvia de representación F es, por supuesto, un pésimo intento de asegurar la consistencia de la conclusión respecto de las premisas. La conclusión se seguiría deductiva y trivialmente de las premisas de F. La manera de introducir las justificaciones en un argumento de sentido común no debe hacer colapsar el tipo de inferencia que queremos representar, con el tipo de inferencia deductiva. En este punto no podemos dejar de referirnos a la propuesta de Reiter en su lógica del razonamiento por *default*. Trataré ahora de hacer una reconstrucción gradual e intuitiva que nos aproxime a su propuesta formal.

La manera en que Reiter estructuraría D (por buscar el ejemplo más simple) como un razonamiento de sentido común es la siguiente:

(D')

x es mamífero : **M** x no pone huevos

x no pone huevos

(donde **M** es un operador modal de posibilidad lógica)

D' se lee: "dado que x es mamífero, tal que es posible asumir que x no pone huevos, entonces se infiere por *default* que x no pone huevos". Preliminarmente se puede decir que el sentido que Reiter le da a "es posible asumir" es que es consistente asumir la proposición en cuestión¹⁷³ respecto del resto del conocimiento, o bien que no hay información en contra de la proposición en cuestión.

La reformulación correspondiente del razonamiento (F) quedaría:

¹⁷³En este caso la proposición en cuestión es "x no pone huevos".

(F')

No espero a nadie en casa \wedge Alguien ha tocado el timbre de mi casa : **M** Ese alguien es un vendedor

Ese alguien es un vendedor

Como puede verse, la parte que está antes de los dos puntos, que Reiter llama "**prerrequisito**", es algo que se toma como conocimiento seguro. Lo que se encuentra después de los dos puntos y el operador modal **M**, constituye para Reiter la "**justificación**" y, en cambio, depende de la revisión que se haga acerca de la información tomada como conocimiento seguro.

En D' y F' la proposición afectada por el operador modal es la misma que la conclusión obtenida. Puede darse el caso de que esto sea más complejo. Por ejemplo, puede ser que estemos justificados a conjeturar algo con base en otra proposición. Supongamos que no existiera el psicópata al que hemos aludido en nuestro ejemplo fantástico. Pero supongamos también que los domingos algunos religiosos osaran, como los vendedores, tocar el timbre de mi casa y así se unieran a estos últimos en lo que podríamos llamar "el asalto del tiempo ajeno" --caso menos fantástico por cierto. Esta situación (que llamaremos "Situación 4"), asumiendo que contiene la Situación 1 (escucho el timbre y no espero a nadie), podría orillarnos a razonar por sentido común conforme al siguiente esquema:

(G)

No espero a nadie en casa \wedge Alguien ha tocado el timbre de mi casa : **M** hoy no es domingo

Ese alguien es un vendedor.

Aunque a veces querríamos en domingo accionar indiscriminadamente la palanca del artefacto descrito en la Situación 2, este esquema de razonamiento no ocasionaría que

los domingos algún vendedor pescara un resfriado (bajo la Situación 2) porque su justificación no es consistente los domingos. Debe notarse que en G la justificación no es la misma que la conclusión. En este caso la justificación que tenemos para conjeturar la conclusión es la consistencia de otra proposición en relación con nuestro conocimiento.

En los esquemas D', F' y G, parece imprescindible que se cumpla el prerrequisito *antes de especular* si obtenemos o no la conclusión. La función del prerrequisito puede verse como la de **autorizar** la inferencia. Si no se da el prerrequisito, no tenemos derecho a una inferencia de este tipo; no tendríamos derecho a ello aún en el caso de que la justificación sí se cumpliera. La función de la justificación es la de **bloquear** o no la inferencia. Si se da el caso de que la justificación no sea consistente, es decir, que la proposición afectada por el operador modal es inconsistente con el resto de nuestro conocimiento de fondo, entonces la inferencia (aún en el caso de que estuviera autorizada) se bloquearía. A *grosso modo*, ésta es la manera en que interactúan los elementos de un razonamiento de sentido común desde la perspectiva de Reiter, de 1980.

Ahora bien, una noción de consecuencia que se defina en correspondencia con este tipo de esquemas de razonamiento (como D', F' y G) no cumplirá con la propiedad de monotonicidad, básica para la definición de nuestra idea de deducción. La monotonicidad puede expresarse del siguiente modo:

$$\text{Si } \Gamma \vdash n \text{ y } \Gamma \text{ está en } \Delta, \text{ entonces } \Delta \vdash n.$$

Que se lee “si de un conjunto Γ de proposiciones se puede inferir deductivamente la proposición n , y si Γ está en Δ , entonces también de Δ se infiere deductivamente n ”; “ \vdash ” es la forma usual de designar la noción de consecuencia deductiva. Una manera más general de formularlo es diciendo que Γ es un subconjunto de Δ :

Si $\Gamma \vdash n$ y $\Gamma \subseteq \Delta$, entonces $\Delta \vdash n$.

Para mostrar que, de acuerdo con nuestra reconstrucción del razonamiento del sentido común, el tipo de inferencia supuesto no cumpliría con el Principio de No Monotonidad, pensemos de nuevo en las Situaciones 1, 3 y 4.

Supongamos que nuestro conocimiento se encuentra determinado básicamente por la Situación 1 y que sólo poseemos el esquema F'. En la Situación 1 podíamos inferir por sentido común, sin mayor problema, que la persona que tocaba el timbre era un vendedor (a través del esquema F'). Supongamos ahora que ocurre un cambio en el estado de cosas y que este cambio resulta básicamente en la Situación 3, cuando aparece en escena el psicópata susodicho. La inferencia que habíamos efectuado para la Situación 1 contradice nuestra nueva información acerca de la cantidad de vendedores existentes en la ciudad. Así, la nueva información (en la Situación 3) bloquearía la inferencia de F'.

Designemos con "S1" al conjunto de proposiciones que definen la información proporcionada en la Situación 1 y, análogamente, designemos "S3" para el caso de la información proporcionada en la Situación 3. Designemos con " \sim " la noción de consecuencia del razonamiento de sentido común. Si la conclusión de F' (Que es un vendedor quien toca) se designa por "n", entonces, de acuerdo al párrafo anterior, tendremos un caso para \sim tal que:

$S1 \sim n$ y, a pesar de que $S1 \subseteq S3$, no es el caso que $S3 \sim n$.

Se dice entonces que una relación de inferencia de este tipo es **no monotónica** o **retractable**¹⁷⁴. Por supuesto, podemos también tener bajo esta misma noción de consecuencia casos en los que el aumento de información no afecte nuestras antiguas conclusiones, pero basta que no se dé el caso general de la propiedad de monotonicidad para saber que una noción de consecuencia es no monotónica. Asignemos, en la forma precedente, "S4" para el caso de la información proporcionada por la Situación 4 y designemos con "d" la proposición de que hoy es domingo. Si usamos el esquema F' bajo la información de S1, sucederá entonces que¹⁷⁵:

$$S1 \mid \sim n \quad \text{y} \quad S1, d \mid \sim n$$

En este caso *d* es irrelevante para la aplicación de F'. La monotonicidad parece cumplirse, pero un sólo caso no basta para determinar esta propiedad. Usando el esquema G bajo la información de S4 la no monotonicidad se manifiesta para $\mid \sim$:

$$S4 \mid \sim n \quad \text{y} \quad \text{no es el caso que } S4, d \mid \sim n$$

Lo anterior se debe al hecho de que mientras S1 está siendo considerado en relación al esquema F', S4 (que incluye a S1) lo está siendo en relación con el esquema G. Si S1 hubiera sido considerado en relación con el esquema G también mostraría, al añadirle la información *d*, la no monotonicidad de la noción de derivación $\mid \sim$.

¹⁷⁴ Nótese que la retractabilidad vino a partir de considerar excepciones a los esquemas, no a partir de su mera falibilidad. Así, falibilidad no implica retractabilidad. Falibilidad no implica revisabilidad, a menos que supongamos racionalidad. La retractabilidad puede verse como una forma de asumir racionalmente la falibilidad.

¹⁷⁵En el siguiente esquema debe tomarse en cuenta que lo designado por la expresión "S1, d" puede verse como designando, también, a un conjunto que incluye a S1.

Para terminar con esta idea intuitiva del razonamiento por *default*, notemos que en esta reconstrucción del razonamiento del sentido común, salta a la vista un detalle interesante. Obsérvese están vinculados a la no monotonicidad de este tipo de derivación no sólo la *información* de que disponemos en un momento dado sino también *los esquemas aceptados* de razonamiento. Recordemos la Situación 2 (la del artefacto malévolo). En ella se proporciona cierta información que nos lleva, aparentemente, a pensar que yo moveré la palanca y que un vendedor saldrá perjudicado (en el caso de que yo sea un hombre poco precavido). Demos por sentado que una persona, digamos X, llega a concluir que yo moveré la palanca, dada la descripción de la Situación 2. Sólo si X tuviera el mismo esquema que yo (a saber, el esquema F') o bien algún esquema que lo llevara a la misma conclusión que el esquema F', X podría llegar a la conclusión de que el individuo que ha resultado mojado es un vendedor. Una consecuencia de todo esto es que las inferencias mías y las de X, en el sistema de Reiter, pueden diferir. Dependerán cada una de su propio **contexto**, representado simplíficadamente por cierta información y ciertos esquemas de razonamiento. Así, también el contexto es un elemento importante en esta modelación del razonamiento del sentido común.

Recapitulando, he ilustrado en esta sección, los siguientes puntos:

- 1) Una forma de entender, racionalmente, a los esquemas de razonamiento falible es suponiéndolos como esquemas con excepciones.
- 2) Entender de esta manera al razonamiento del sentido común nos lleva a carecer de la propiedad de monotonicidad, básica en la definición del razonamiento deductivo.

3) La perspectiva de Reiter del razonamiento del sentido común relaciona a este tipo de razonamientos con la propiedad de no monotonicidad y con un tipo de dependencia contextual¹⁷⁶.

4) Dicho contexto puede representarse como un compuesto de dos clases de cosas: un conocimiento de fondo¹⁷⁷ y un conjunto de esquemas para razonar conjeturalmente¹⁷⁸.

En la segunda parte y última de esta sección, proseguiré con la exposición del sistema de Reiter en un tenor más formal y detallado.

3.2.2 Una introducción informal a la lógica del razonamiento por *default*.

La lógica del razonamiento por *default*, de Raymond Reiter¹⁷⁹, es una lógica propuesta para modelar el razonamiento del sentido común. En la primera parte de esta sección presenté una caracterización intuitiva de las principales propiedades de los esquemas propuestos por Reiter en uno de los artículos fundacionales de la lógica no monotónica. Mediante una serie de ejemplos, en la primera parte ilustré los siguientes puntos:

¹⁷⁶La dependencia contextual, será capturada sólo de una manera simplificada.

¹⁷⁷Este conocimiento será considerado por Reiter como un núcleo imposible de retractar y, en este sentido, inalterable.

¹⁷⁸En el sistema de Reiter este conjunto de esquemas de razonamiento serán meta-reglas. Aunque el carácter metalingüístico de las reglas de Reiter dificulta una interpretación intuitiva de las mismas, no me parece arriesgado decir que dichas reglas son la forma en que Reiter intenta capturar parte de nuestro conocimiento del mundo; dicho conocimiento se considera falible. En este sentido las meta-reglas de Reiter representarían nuestro conocimiento modificable.

¹⁷⁹ [Reiter 1980] Reiter, R., "A Logic for Default Reasoning", en *Artificial Intelligence*, Vol. 13, 1980, pp. 81-123. Reimpreso en Matthew L., Ginsberg, (comp.), *Nonmonotonic Reasoning*, Morgan Kaufmann Publishers, E.U.A. 1987. En las siguientes referencias, anotaré la paginación correspondiente a la compilación de Ginsberg, enseguida de "G:"

- 1) Una forma de entender racionalmente los esquemas de razonamiento falible es suponiéndolos como esquemas con excepciones.
- 2) Entender de esta manera al razonamiento del sentido común nos lleva a carecer de la propiedad de monotonicidad, básica en la definición del razonamiento deductivo.
- 3) La perspectiva de Reiter del razonamiento del sentido común relaciona a este tipo de razonamientos con la propiedad de no monotonicidad y con un tipo de dependencia contextual.
- 4) Dicho contexto puede representarse como un compuesto de dos clases de cosas: un conocimiento de fondo¹⁸⁰ y un conjunto de esquemas para razonar conjeturalmente¹⁸¹.

El propósito de esta segunda parte de la sección es ofrecer una introducción a la perspectiva de Reiter para modelar el razonamiento de sentido común. Lo haré con un objetivo muy específico: elaborar una aproximación paulatina a la comprensión de la noción de consecuencia involucrada en su sistema lógico.

En el marco de la perspectiva anterior, puntos 1-4, para Reiter el razonamiento del sentido común sigue algo parecido a una regla particular como ésta:

$AVE(x): M VUELA(x) / VUELA(x)$ ¹⁸²

Intuitivamente: “si algo es un ave y es posible que vuele, (por lo tanto) vuela”

¹⁸⁰Este conocimiento será considerado por Reiter como un núcleo imposible de retractar y, en este sentido, inalterable.

¹⁸¹En el sistema de Reiter este conjunto de esquemas de razonamiento serán meta-reglas. Aunque este carácter de las reglas de Reiter dificulta una interpretación intuitiva de las mismas, dichas reglas son la forma en que Reiter intenta capturar parte de nuestro conocimiento falible del mundo. En este sentido las meta-reglas de Reiter representarían nuestro conocimiento modificable.

¹⁸²Como los esquemas F', D' y G que anticipé en la primera parte de este artículo. En este caso la representación se encuentra "comprimida" en un sólo renglón.

Esto es lo que Reiter llama una "**regla *default***". Así, una regla *default*¹⁸³ tendrá la siguiente forma general [Reiter 1980, p. 88]¹⁸⁴:

$$\alpha : \mathbf{M}\beta_1 \dots \mathbf{M}\beta_n$$

ω

donde α es considerado el prerrequisito, en el sentido de que hace posible la aplicación de la regla; cada una de las β_i en la secuencia $\beta_1 \dots \beta_n$ es una justificación que es consistente asumir; ω es la conclusión que obtenemos por *default*; finalmente, **M** es un operador metalingüístico que aplicado a un fórmula indica que ésta se puede asumir consistentemente.

"Aplicar un *default* δ " significa obtener ω mediante δ , es decir, que una razón para inferir ω es una derivación directamente mediante δ . Se puede decir que la presencia de α "autoriza" la aplicación del *default*, pero esto no significa necesariamente que el *default* sea aplicado de hecho. Esto se debe a que faltaría la condición de que todas las justificaciones β_i del *default* se cumplieran. Sólo en el caso de que ambas condiciones estén dadas (prerrequisito y justificaciones) el *default* se aplica.

Obsérvese que he hablado de la "presencia" del prerrequisito α y del "cumplimiento" de las justificaciones β_i . He querido con ello anticipar la distinción entre ser derivable de un conjunto de proposiciones y ser consistente con un conjunto de proposiciones: que " α esté presente" intenta significar que sea derivable, mientras que " β_i

¹⁸³En ocasiones me referiré a estas reglas simplemente con "*defaults*".

¹⁸⁴[Reiter 1980] p. 88. G: p.71.

se cumpla" se refiere sólo a que β_i es consistente. Así, observemos, no es necesario derivar las β_i para cumplir con la segunda condición de un *default*, tan sólo basta que sean consistentes. ¿Derivar o checar consistencia, respecto de cuál conjunto de proposiciones? Esto es parte de lo que debemos saber para tener una idea más clara de la noción de consecuencia en la lógica de Reiter.

El mecanismo de activación de un *default* va como sigue: una vez que se cuenta con el prerrequisito, la aplicación o el bloqueo de la derivación dependen de la consistencia de las justificaciones. En realidad, el prerrequisito autoriza la aplicación del *default* pero también autoriza su bloqueo. Llamemos "**conjunto base**" al conjunto de proposiciones explícitas que sirve como punto de partida para determinar qué inferencias pueden efectuarse y llamaré "**conjunto de referencia**" al conjunto de proposiciones respecto del cual se checa la consistencia de una proposición. Veamos el siguiente ejemplo.

Supongamos el conjunto base de proposiciones $K = \{-q, r\}$ y el *default* siguiente:

(C1)

$r : Mq$

s

La aplicación del *default* anterior está autorizada por K, en virtud de r. No obstante, la justificación q no es consistente con K. De manera que, por lo menos respecto de K, no se cumple la segunda condición del *default* y, por tanto, se bloquea la aplicación del *default*. Ahora bien, en el ejemplo hice una concesión que es pertinente aclarar: dije antes que r tenía que ser derivable, pero no aclaré en qué sentido debe entenderse esta derivación. En el ejemplo anterior, r autoriza directamente al *default* sin necesidad de ninguna derivación complicada, pues r pertenece a K (en símbolos: $r \in K$). Sin embargo, r

también puede suponerse, en una formulación más general, como una consecuencia de K: en un sentido obvio sabemos que r es consecuencia de $\{-q, r\}$.

El conjunto de consecuencias resultante de un conjunto de proposiciones no vacío podría tener un número infinito de miembros¹⁸⁵. Por ello, intentemos una representación más compacta de lo anterior. Pensemos en una función de teoremicidad Th, la función de Tarski, una función recursiva sobre conjuntos de proposiciones que resulta en el conjunto de todas las consecuencias que pueden obtenerse con lógica clásica a partir de un conjunto inicial. En realidad, cuando hablamos de que un *default* está autorizado por un conjunto de proposiciones K no quisiéramos referirnos sólo a que algún miembro de K autoriza el *default* sino que también quisiéramos referirnos, más generalmente, a que alguna consecuencia¹⁸⁶ de K autoriza el *default*. Así, más explícitamente, el *default* anterior está autorizado por un miembro de $Th(\{-q, r\})$, el conjunto de consecuencias clásicas de K.

Pero llevemos aún más lejos esta idea de derivación. Para Reiter los prerequisites no necesariamente tienen que ser derivados sólo mediante lógica clásica, sería adecuado que también pudieran inferirse de otras reglas *default*. Una consecuencia de algún *default* puede servir, a su vez, de prerequisite para otro *default*; ello implica que tenemos que tomar en cuenta para la derivación no sólo las reglas clásicas sino también un conjunto de reglas *default*: para derivar algo en el razonamiento del sentido común, usamos no sólo conocimiento obtenido deductivamente, también usamos conocimiento justificado no deductivamente, conocimiento claramente no seguro. Pensemos en el siguiente ejemplo. Sean K y R, un conjunto de proposiciones y uno de reglas *default*, respectivamente:

¹⁸⁵ Pensemos en el caso de $K = \{p, q\}$ y en la regla clásica de conjunción: $p, q / p \& q$. El conjunto de consecuencias que resulta de aplicar esta regla a K contiene a $p \& q$, pero por lo menos también a $(p \& q) \& p$, $(p \& q) \& q$, $[(p \& q) \& p] \& p$, $[(p \& q) \& p] \& q$, etc.

¹⁸⁶ Si necesito $p \& q$ para obtener la consecuencia g, y tengo que mi conjunto base es $K = \{p, q\}$, es racional pensar, en el contexto clásico de reglas de inferencia, que puedo obtener g a partir de K. Esto es porque supongo que tengo no sólo a K, tengo a K y a todas sus consecuencias.

(C2)

$$K = \{\neg q, r\}$$

$$R = \{\neg q \wedge r : \mathbf{Ms} / t; t : \mathbf{Mf} / g\}^{187}$$

En este caso, el primer *default* de R está autorizado por $\text{Th}(\{\neg q, r\})$ y el segundo *default* de R estaría autorizado si pudiéramos derivar t del primero. Para aplicar el primer *default* será necesario que se cumpla su justificación. Obsérvese que s es consistente respecto de K, por lo que hemos de aplicar el primer *default*¹⁸⁸. Así, la consecuencia por *default* t, autoriza el segundo *default* y, dado que f también es consistente en el sentido en que lo es s, tenemos también a g. Así, añadiendo vocabulario de la primera parte de este artículo, la consecuencia $\neg q \wedge r$, autoriza al primer *default*, y la *conjetura* t autoriza al segundo.

En este punto tenemos que observar que si para realizar la derivación hemos estado añadiendo proposiciones al conjunto base¹⁸⁹. Usualmente consideramos todas las consecuencias de un conjunto base de proposiciones a la hora de revisar si es un conjunto consistente. Lo más racional es, entonces, que este conjunto incrementado sea el conjunto de referencia respecto del cual tendríamos que checar también la consistencia de las justificaciones. En realidad, la forma de verificar la consistencia de s y f en el párrafo anterior no es la forma en que Reiter propone que se efectúe. Reiter exigirá que en la verificación de consistencia se tome como conjunto de referencia todo el resultado de nuestras derivaciones, sean clásicas o no deductivas. Si esto no es así, esto hará posible que algunas de nuestras consecuencias parezcan irracionalmente justificadas. El ejemplo C2 no tendría ningún problema con lo acabamos de decir, pero veamos el siguiente caso:

¹⁸⁷ Usaré indistintamente “&” y “^”.

¹⁸⁸ Como se verá en seguida, en el sistema de Reiter esto no es exactamente lo que sucede: la consistencia no se checa sólo respecto del conjunto base. Pero concederé el caso (por el momento) sólo en función de una presentación didáctica del sistema.

¹⁸⁹ Primero tomamos en cuenta K, luego $\text{Th}(K)$, y por último se ha sugerido algo como la unión de $\text{Th}(K)$ con las consecuencias obtenidas por *default*.

(C3)

$$K = \{\neg q, r\}$$

$$R = \{\neg q \wedge r : \mathbf{M}s / t; t : \mathbf{M}\neg t / g\}$$

En C3, t está asegurada pues no hay nada que impida inferirla¹⁹⁰ del primer *default*. El segundo *default* está autorizado por la inferencia de t . Observemos además que la justificación del segundo *default* es consistente con K , el conjunto base. Si nuestro conjunto de referencia fuera K habríamos inferido t legítimamente y g con base en la suposición de que la negación de t es consistente con K . Pero Reiter no querría que una derivación así fuera legítima pues una consecuencia de $\text{Th}(K)$ junto con los *defaults* es precisamente t y sería muy poco intuitivo que algo fuera derivable a partir de K a través de la negación $\neg t$, es decir, mediante la negación de una proposición que es derivable a partir de K . En contraste, Reiter exige que la consistencia se cumpla respecto de *todas* las consecuencias del conjunto base y el conjunto de reglas *default*, es decir, respecto de $\text{Th}(K \cup \{t\})$. Si seguimos a Reiter entonces el segundo *default* de R se bloquearía y, por ende, g no sería una consecuencia para el ejemplo C3. Con esta perspectiva, Reiter soluciona casos que serían aún más graves:

(C4)

$$K = \{\neg q, r\}$$

$$R = \{\neg q \wedge r : \mathbf{M}s / t; r : \mathbf{M}\neg t / \neg t\}$$

En C4, tanto s como $\neg t$ son consistentes respecto de K y ambos *defaults* están autorizados. Así, si nos alejamos de la perspectiva de Reiter, tendríamos como resultado un conjunto de consecuencias inconsistente, pues estarían en él tanto t como $\neg t$. C3 y C4 ejemplifican bien la necesidad de tener en cuenta un punto de vista antiguo en la lógica: el conjunto de consecuencias de un conjunto base en relación con un tipo de derivación (en este caso un

¹⁹⁰Las dos condiciones se cumplen: prerrequisito y justificación.

conjunto de reglas clásicas y no clásicas) debe ser considerado como un conjunto que se ha producido aplicando simultáneamente todas las posibles derivaciones. Los *defaults*, tal como Reiter los presenta, no están ordenados por orden de aplicación. Para verlo más claro: R de C3 pudo haberse presentado como $\{t : \mathbf{M}\text{-}t / g; \neg q \wedge r : \mathbf{M}s / t\}$ y esta nueva presentación no cambia nada nuestro conjunto de consecuencias. ¿Cómo es esto posible? La razón es que el *default*:

$$r : \mathbf{M}s / t$$

no tiene en C3 ninguna restricción para aplicarse y por tanto, como se debe aplicar, t estará forzosamente en nuestro conjunto de consecuencias; en cambio, el *default* restante tendría problemas de inconsistencia, dada la inferencia irrestricta de t. Así, el *default* restante se bloqueará inevitablemente.

Mediante los casos C1-C4 podemos poner de relieve algunas notas pertinentes:

- (a) Para la autorización de las derivaciones mediante *defaults* se debe tomar en cuenta no sólo el conjunto base de proposiciones K sino también $\text{Th}(K)$ (como se ilustró con C1).
- (b) Pero además, para la autorización de las derivaciones mediante *defaults* se debe tomar en cuenta no sólo lo descrito en a sino también otras derivaciones por *default* (como se ilustró con C2).
- (c) Para el chequeo de consistencia en las derivaciones mediante *defaults* se debe tomar en cuenta no sólo el conjunto base de proposiciones K sino también $\text{Th}(K)$ (como se dejó implícito con C3 y C4).
- (d) Pero además, para el chequeo de consistencia en las derivaciones mediante *defaults* se debe tomar en cuenta no sólo lo descrito en c sino también las otras consecuencias por *default* (como se ilustró con C3 y C4).

(e) Así, para derivar mediante *defaults* es necesario tomar en cuenta todas las derivaciones (*default* y clásicas) como si hubiesen sido efectuadas simultáneamente.

He relacionado varias cosas: conjuntos base de proposiciones, conjuntos de referencia, conjuntos de reglas, la función de teoremicidad clásica y conjuntos de consecuencias con proposiciones derivadas por *default* y con proposiciones derivadas clásicamente. Afinaré ahora la noción de conjunto de consecuencias en el sistema de Reiter. Para decirlo con simpleza, podemos pensar que un conjunto de consecuencia es algo que resulta de completar conjuntos base en relación con conjuntos de reglas. El conjunto de reglas es una forma de caracterizar la capacidad de una función de teoremicidad. Desde esta perspectiva un **conjunto de consecuencias** es la forma de completar, con derivaciones, a otros conjuntos. Reiter llama "**extensiones**" a estos conjuntos de consecuencias. Para Reiter, las extensiones correspondientes a C2-C4, serían las siguientes:

Para C2, $E = \text{Th}(K \cup \{t, g\})$

Para C3, $E = \text{Th}(K \cup \{t\})$

Para C4, $E = \text{Th}(K \cup \{t\})$

Ordenar los pasos para efectuar derivaciones con las características descritas por *a-e* es una tarea compleja [Marek & Truszcynski 1993]¹⁹¹. Sin ordenar de este modo, Reiter sólo organizó estas características estipulando las siguientes tres propiedades necesarias para una extensión:

¹⁹¹[Marek & Truszcynski 1993] V. W. Marek y M. Truszcynski sí realizan esta tarea. Cfr. su *Nonmonotonic Logic. Context-Dependent Reasoning*, Springer-Verlag, Berlín, 1993, pp. 37-71.

I) La extensión debe contener al conjunto base que se ha tomado como punto de partida. Es decir, toda extensión preserva el conjunto base.

II) La extensión es un conjunto que contiene todas las consecuencias clásicas que pueden extraerse a partir del contexto (el resultado de las reglas *default* y el conjunto base). Más correctamente, es un conjunto cerrado bajo consecuencia clásica.

III) La extensión debe contener todas las consecuencias de los *default* que estén autorizados y que no sean bloqueables. En otras palabras, la extensión debe estar cerrada bajo consecuencia *default*.

I y II pueden sin mucho esfuerzo entenderse a través de la observación de las extensiones propuestas para C2-C4. III, en cambio, puede parecer menos obvia.

Para la construcción formal de las teorías *default*, Reiter parte de un lenguaje L como el que normalmente se usa para la lógica de primer orden. L contiene variables proposicionales y de individuo, letras de función, letras predicativas, constantes lógicas y signos de puntuación. Aunque en L se pueden expresar fórmulas cuantificacionales, simplificaré las formulaciones y ejemplos empleando un lenguaje para cálculo proposicional.

El sistema de Reiter cuenta con un conjunto de fórmulas bien formadas $H \subseteq L$ y también con un conjunto $I \subseteq H$ de fórmulas cerradas. Reiter define entonces la función Th_L de teoremicidad clásica sobre conjuntos de fórmulas $S \subseteq I$ tal que:

$$Th_L(S) = \{\omega \mid \omega \in I \text{ y } S \vdash \omega\}$$

Es decir, Th_{\perp} aplicada al conjunto de fórmulas cerradas S resulta en el conjunto de todas las fórmulas cerradas ω , para las que tenemos una prueba a partir de S .

A partir de ello, define una **teoría default** Δ como un par ordenado $\langle D, W \rangle$, tal que W (**presupuestos sobre el mundo**) es un conjunto de fórmulas cerradas y D es el **conjunto de reglas default** de que se dispone. Esta es la forma de caracterizar al **contexto** del que hablé más arriba, conocimiento seguro y conocimiento no seguro. W debe entenderse como el conjunto base K , y D como el conjunto de reglas R . Podemos definir un *default* cerrado como uno en el que toda fórmula que ocurre en él es cerrada. Así, si una teoría Δ posee un conjunto D tal que todo *default* en él es cerrado, se dice que Δ es cerrada.

Designemos con " Γ " una función sobre conjuntos de fórmulas $S \subseteq I$ y relativa a una teoría *default* cerrada $\Delta = \langle D, W \rangle$, que resulta en el más pequeño conjunto¹⁹² $\Gamma(S)$ que satisface las siguientes tres características:

$$D1. W \subseteq \Gamma(S)$$

$$D2. Th_{\perp}(\Gamma(S)) = \Gamma(S)$$

$$D3. \text{ Si } (\alpha: \mathbf{M}\beta_1 \dots \mathbf{M}\beta_n / \omega) \in D \text{ y } \alpha \in \Gamma(S), \text{ y } \neg\beta_1, \dots, \neg\beta_n \notin S, \text{ entonces } \omega \in \Gamma(S)^{193}$$

La función Γ caracteriza la noción de consecuencia por *default* relativa a una teoría, en la lógica de Reiter. Como puede verse, Γ incorpora las características I-III que anoté antes.

¹⁹²La exigencia de ser el más pequeño conjunto se refiere a lo siguiente. Supongamos que tenemos un conjunto $\Gamma(G)$ tal que es el más pequeño conjunto que cumple con D1-D3 sustituyendo $\Gamma(S)$ con $\Gamma(G)$ y S con G . Si tuviéramos una fórmula m completamente ajena a G , es decir, una fórmula tal que $\Gamma(G) \subset \Gamma(G \cup \{m\})$, entonces $\Gamma(G) \cup \{m\}$ cumpliría, de la misma forma en que lo hace $\Gamma(G)$, con D1-D3.

¹⁹³[Reiter 1980], p. 89. G: p.72.

Un conjunto de fórmulas $E \subseteq I$ es una extensión para una teoría Δ cerrada si y sólo si ocurre que $\Gamma(E) = E$ respecto de Δ . Se dice en este caso que la extensión E es un punto fijo de la función Γ . Obsérvese que inversamente, desde la definición de Γ relativa a Δ , si $\Gamma(S) = S$ entonces S es una extensión para Δ .

D1 nos asegura que nuestras suposiciones iniciales acerca del mundo no se pierdan: toda extensión las contendrá. Esta formulación responde con mayor formalidad a la condición I que mencioné antes. D2 supone que la extensión será un punto fijo de la función de teoremicidad. Un punto fijo de una función f es un valor r para la función, tal que si la función se aplica a r dará el mismo resultado¹⁹⁴. Así, como dicta la condición II, la extensión contendrá todas las consecuencias clásicas que se puedan obtener de ella. D3 obliga a que la extensión posea las conclusiones de los *defaults* aplicables y no bloqueables, esto es, los *defaults* de la teoría que cumplan con dos características: que su prerequisite pertenezca a la extensión y que sus justificaciones β sean consistentes con ella. D3 se encuentra descrita antes en la condición III.

En los casos C3 y C4 puede verse claramente la importancia de tomar en cuenta todas las consecuencias de las derivaciones (de acuerdo a D1-D3) a la hora de checar la consistencia de un conjunto de justificaciones. Examinaré ahora un detalle relacionado con esto que será importante para completar el entendimiento de lo que significa ser consecuencia en el sistema lógico de Reiter. Obsérvese que las inferencias dependerán del conjunto de reglas que se tome en cuenta como punto de partida (dependerán por tanto de la teoría *default* de que se trate). Esto puede ocurrir también en lógica clásica. Trivialmente, un conjunto de consecuencias a partir de un mismo conjunto base K puede variar según el conjunto de reglas que uno le aplique al conjunto base, sin importar el tipo de reglas que se use, pueden ser todas ellas deductivas. Así también, en la lógica de

¹⁹⁴Por ejemplo, en la función de identidad sobre x , $I(x) = x$. Si aplicáramos de nuevo la función a este resultado, el resultado sería x misma. Decimos así que x es un punto fijo de $I(x)$.

Reiter, una extensión de una teoría $\Delta = \langle D, W \rangle$ puede ser distinta de la extensión $\Delta' = \langle D', W \rangle$, donde $D \neq D'$. En cambio, que el conjunto de consecuencias varíe para un mismo conjunto base K respecto de un mismo conjunto de reglas es algo que no ocurre en cualquier tipo de lógica. En la lógica del razonamiento por *default* ocurre, pero no en la lógica clásica.

Supongamos la teoría $\Delta = \langle D, W \rangle$ tal que:

(C5)

$$D = \{-q : Ms / f; r : M-f / \neg s\} \text{ y}$$

$$W = \{-q, r\}.$$

En este caso $\neg q$ y r autorizan respectivamente el primer y el segundo *defaults*. No hay forma de que estas dos autorizaciones se impidan. Por el contrario, s es consistente sólo si $\neg s$ no se deriva (mediante el segundo *default*) y, recíprocamente, $\neg f$ es consistente sólo si no se deriva f (mediante el primer *default*). C5 nos lleva a dos extensiones distintas para la misma teoría:

$$E_1 = \text{Th}(W \cup \{f\})$$

$$E_2 = \text{Th}(W \cup \{\neg s\})$$

E_1 y E_2 son extensiones completamente legítimas para Δ según D1-D3. Este hecho marca una diferencia muy interesante de la lógica *default*, respecto de otro tipo de lógicas: en ella es posible tener varias alternativas de derivación que nos llevan a resultados distintos partiendo del mismo conjunto base de datos y del mismo conjunto de reglas. Puede haber

dos formas de completar el mismo contexto. Podemos llamar a esto "**multiplicidad de extensiones**".

Dado que podrá ocurrir que dos *defaults* arrojen cada uno conclusiones que no se pueden sostener juntas, tendremos que decidir entre un conjunto de creencias y otro, esto es, uno que contenga una de las conclusiones y otro que contenga la otra. Así, "parecería que los *defaults* pueden sancionar diferentes conjuntos de creencias acerca del mundo" [Reiter 1980, p. 86]¹⁹⁵.

Una teoría *default* puede entonces ser considerada como una teoría incompleta que puede extenderse mediante la aplicación de algunos de sus *defaults*, pero esto nos puede conducir, además, a resultados alterantivos. Esta parece ser una de las razones por las que Reiter propone que los *defaults* deben entenderse como meta-reglas que nos guían en la manera de construir las extensiones que completarían la teoría [Idem]¹⁹⁶.

Un efecto de la multiplicidad de extensiones sobre la idea de derivación a partir de una teoría en el sistema de Reiter es el siguiente: una fórmula ϕ es consecuencia de una teoría Δ si y sólo si ϕ pertenece por lo menos a una extensión E_i de Δ . Obsérvese que esta no es la única manera de tomar el asunto, pueden concebirse, de hecho las hay, lógicas similares que consideren a ϕ como consecuencia, sólo si aparece en toda extensión de la teoría¹⁹⁷.

Antes de terminar, ilustraré una de las propiedades más importantes de la lógica del razonamiento por *default*, la **no monotonicidad**, que es una característica común a las

¹⁹⁵"It would seem that defaults can sanction different sets of beliefs about a world". [Reiter 1980], p. 86. G: p. 71.

¹⁹⁶"Defaults therefore function somewhat like meta-rules; they are instructions about how to create an *extension* of this incomplete theory." *Idem*.

¹⁹⁷ Revítese, por ejemplo, [McDermott & Doyle 1980].

lógicas que, justamente, llamamos “no monotónicas”. Además, observamos antes que, al parecer, el razonamiento de sentido común involucra esta propiedad.

Está claro que respecto de una teoría Δ es concebible que algunos *defaults* sean bloqueables y otros no. El que un *default* se bloquee depende de lo que el conjunto de fórmulas W pueda producir en relación con el conjunto de reglas D , en el contexto de una función de teoremicidad. En esta situación, es fácil imaginar que al aumentar nuestra información sobre el mundo haya la posibilidad de que perdamos la capacidad para inferir algunas cosas que antes podíamos inferir, esto violaría el principio clásico de monotonicidad.

Las teorías *default* son no monotónicas en el sentido de que si tenemos dos teorías default cerradas $\Delta = \langle D, W \rangle$ y $\Delta' = \langle D \subseteq D', W \subseteq W' \rangle$ (situación que simplificaremos como $\Delta' \subseteq \Delta$), entonces podrá ocurrir que no haya un par de extensiones E' y E de Δ' y Δ respectivamente, tales que $E \subseteq E'$ [Reiter 1980, p. 91]¹⁹⁸. Veamos el siguiente ejemplo:

(C6)

Supongamos la teoría cerrada $\Delta = \langle D, W \rangle$ tal que

$$D = \{ :Mf / -s; :Ms / -t \}$$

$$W = \emptyset$$

Si tenemos además los conjuntos,

$$R = \{ r :Mg / -f \}$$

$$K = \{ -f \}$$

¹⁹⁸[Reiter 1980], p. 91. G: 73.

Supongamos la teoría $\Delta' = \langle D \cup R, W \cup K \rangle$.

Las extensiones resultantes serán,

Para Δ :

$$E = \text{Th}(W \cup \{\neg s\})$$

Para Δ' :

$$E' = \text{Th}(W \cup K \cup \{\neg t\})$$

Como puede corroborarse, en C6 no se cumple que $E \subseteq E'$ a pesar de que $\Delta \subseteq \Delta'$. Esto ilustra claramente la no monotonicidad que poseen las teorías *default* en la lógica de Reiter. Puede comprenderse fácilmente que también puede ocurrir esto en un caso más simple, cuando la información que aumenta es sólo respecto de W y no respecto de D . La propiedad de una consecuencia de poder dejar de inferirse es comúnmente llamada "**retractabilidad de las consecuencias**". **Retractar una consecuencia en un sistema** es entendido como perder la capacidad inferir dicha consecuencia en el sistema. Es interesante notar cómo, al ganar información en el conjunto base, no sólo podemos perder algunas consecuencias que podíamos inferir antes (no monotonicidad) sino que también podríamos ganar consecuencias que no podíamos inferir antes.

Finalmente, mencionaré tan sólo una característica más de la lógica de Reiter, que me parece importante resaltar. Dado que la noción de consecuencia estipulada por estas condiciones depende esencialmente de la revisión de consistencia en relación con toda la extensión, es difícil concebir que los *defaults* de una teoría nos lleven a la construcción de una extensión inconsistente para esa teoría. Sin embargo, cabe la posibilidad de que una teoría *default* contenga alguna extensión inconsistente.

Así, en el sistema de Reiter una teoría Δ cerrada tiene una extensión inconsistente si y solamente si una de sus extensiones es el conjunto de todas las fórmulas cerradas de L . Y este único caso de inconsistencia de una extensión de una teoría Δ sólo puede darse en virtud de la inconsistencia del componente W de Δ . De hecho, una extensión de Δ es inconsistente si y sólo si W de Δ es inconsistente [Reiter 1980, p. 91]¹⁹⁹.

He expuesto, a lo largo de las dos partes de este artículo, algunos elementos que nos aproximaron paulatinamente a la comprensión de la noción de consecuencia de la lógica del razonamiento por *default*, de Reiter. Con ello intenté ofrecer una introducción a la perspectiva de Reiter para modelar el razonamiento de sentido común. Es importante mencionar que el sistema de Reiter es apenas uno de los sistemas de lógica no monotónica que se han desarrollado desde principios de la segunda mitad del siglo XX. Algunos de estos sistemas son propuestas para el análisis del sentido común que difieren importantemente de la racionalidad subyacente en el sistema de Reiter. Tal vez esta introducción sirva al lector como un incentivo para el estudio de otros sistemas no monotónicos.

3.3. Explicación e Ignorancia

En el capítulo anterior enfatiqué algunas características lógicas de las explicaciones I-S, que mencioné bajo el rubro de "representación racional de la ignorancia":

- 1) La posibilidad de tener premisas falibles. Característica que consiste, básicamente, en que los miembros del *explanans* pueden no ser todos verdaderos.

¹⁹⁹[Reiter 1980], *op. cit.*, p. 91. G: 73.

2) El debilitamiento de la noción de derivabilidad (entre *explanans* y *explanandum*), que hace que el *explanandum* pueda dejar de derivarse por incremento de información.

3) La idea dinámica de la explicación. Característica que incluye la idea de que la explicación puede sufrir modificaciones cuando interviene un cambio en el contexto.

Como dije también en esa sección, Hempel y en general los críticos del modelo I-S de la explicación, como Coffa y Railton, interpretan la concepción epistémica de la explicación o bien como un mal necesario o bien como algo que debería evitarse al modelar esta actividad científica. Sin embargo, como sugerí desde entonces, la concepción epistémica de la explicación es una forma de abordar la elucidación de la explicación y no una consecuencia que se limite a cierto tipo de explicaciones probabilísticas. Una explicación D-N también puede ser entendida en términos de una relativización epistémica. Esto nos deja en posición para un análisis más directo del comportamiento de una explicación D-N cuando posee premisas falsas. Con dicho comportamiento me refiero a las relaciones de la explicación con los cambios ocurridos en su contexto. Por lo tanto, visto desde un plano epistémico, la explicación D-N comparte las características 1 y 3 con las explicaciones I-S.

No obstante, algunas características de las explicaciones I-S seguirán siendo diferentes respecto de las explicaciones D-N. Fundamentalmente, sus diferencias se mantendrán debido a que suponen tipos distintos de derivación. El tipo de derivación de las explicaciones I-S es más débil que el de las explicaciones D-N. A partir de lo dicho por Salmon y del análisis de van Fraassen, hay una característica distintiva entre una derivación deductiva y una derivación más débil (como es el caso de la derivación inductiva): que la última puede perder la capacidad para inferir algunas de sus conclusiones cuando aumenta la evidencia.

Hemos llamado a la propiedad referida en el párrafo anterior "no monotonicidad". En relación con la explicación, la no monotonicidad puede presentarse al menos de dos formas:

NM1) En una derivación con la que se intenta explicar un evento (derivación del *explanandum*).

NM2) En una derivación con la que se intenta proponer un *explanans* para un evento (derivación del *explanans*).

Como he dejado claro en la introducción general, mi interés en esta investigación es sentar las bases para la construcción de un modelo de explicación argumental en el sentido de construir un tipo de esquema de argumento que caracterice el tipo de relación de apoyo que ofrece el *explanans* al *explanandum* y no el tipo de inferencia que nos llevaría a proponer un *explanans* para un *explanandum* dado. En el primer caso, podemos decir que estamos tratando de construir un modelo de la explicación tomada como producto, en el segundo, un modelo de la explicación tomada como un proceso de postulación de hipótesis explicativas. Así, aunque considero esta segunda opción una línea de investigación muy interesante, mi perspectiva en esta tesis tiene que ver con la explicación tomada como un producto y no como un proceso. Consecuentemente, no abordaré el tipo NM2 de no monotonicidad en la explicación.

El tipo NM1 ocurre en argumentos cuya conclusión representa un evento a explicar. Podemos identificar al menos dos casos: cuando la conclusión deja de inferirse pero esto no implica que el argumento deje de ser una explicación y cuando la conclusión deja de inferirse y el argumento deja de ser una explicación. Llamaré a los primeros casos "casos NM1a" y a los segundos, "casos NM1b". Mostrar cualquiera de los dos tipos de caso es suficiente para reforzar mi tesis de que un modelo adecuado de explicación tendrá que ser no montónico.

Las explicaciones I-S proporcionan casos claros del tipo NM1a. El caso *Rec* expuesto en la sección 1.2 es un ejemplo. En ese caso, cuando aumentamos la evidencia de que el individuo en cuestión tiene 92 años, la conclusión de que se recuperará no parece seguirse inifiriendo. En su lugar tenemos entonces otro argumento, el que nos conduce a la no recuperación del individuo, a saber, $\neg Rec$. En estas circunstancias, la imposibilidad de sostener algunas conclusiones puede no deberse a la imposibilidad para que un conjunto de proposiciones siga guardando una relación *explanans-explanandum* con una proposición singular. Veremos más adelante cómo, desde mi perspectiva, *Rec* sigue siendo una explicación aunque no se derive más su conclusión.

Salmon dice que el problema está en que en este tipo de derivaciones la ley del refuerzo del antecedente, no se cumple. Sin embargo, es importante notar, que este comportamiento (el no cumplimiento de la ley antedicha) de algunas explicaciones no sucede si no son concebidas epistémicamente. La idea de que una explicación puede dejar de serlo si la información aumenta supone una concepción epistémica de la explicación. La razón es que si la explicación es concebida, en cambio, como algo que dispone de toda la información relevante acerca de la ocurrencia del *explanandum*, entonces no es posible pensar en que tal explicación deje de serlo debido a un aumento de información.

Railton pensaba que la explicación debía tomar en cuenta sólo relaciones completamente verdaderas, de las cuales no podía suponerse ninguna información desconocida. Por supuesto, esto no obsta para concebir que haya eventos objetivamente azarosos. Es decir eventos de los cuales, independientemente de nuestra información, no hay factores que los produzcan necesariamente. No obsta tampoco para que haya explicaciones para dichos eventos. Railton se propuso modelar este tipo de explicaciones, lo cual es un intento de llevar las explicaciones I-S a un plano no epistémico.

Así, es necesario distinguir dos cosas más: primero, el comportamiento de algunas explicaciones que consiste en perder, debido a un aumento de información, la capacidad

de derivar algunas conclusiones y, segundo, la característica de que algunas explicaciones, vistas como argumentos²⁰⁰, son falibles en el sentido de que su fenómeno *explanandum* puede no ocurrir. Ambas cosas son dependientes de una concepción epistémica de la explicación. La primera depende de una concepción epistémica en el sentido de que supone que la inferencia es dependiente de un cierto conjunto de conocimientos que no es inmutable; la segunda, en el sentido de que supone una distinción entre los hechos que se postulan existentes en base a un conjunto de conocimientos dado, y los hechos que son considerados como realmente existentes.

En el capítulo 1 traté de construir la forma de entender la explicación desde un punto de vista epistémico. Dije que una concepción epistémica de la explicación sería útil para reformular los problemas que nos sugiere la construcción del modelo I-S de explicación²⁰¹. Después establecí una conexión entre la no monotonicidad y la asunción argumental y dije que la forma plausible de asumir AA para modelar la explicación estaba relacionada con la idea de una concepción epistémica de la explicación²⁰². Una concepción epistémica de la explicación tendría que contemplar las distinciones que hice en los anteriores párrafos.

Una idea suficientemente clara de una concepción de la explicación como ésta no puede construirse sin un análisis del problema de la pertinencia explicativa²⁰³. Sin embargo, como hacer un análisis de este problema excedería los límites de esta tesis, trataré de bosquejar brevemente la concepción epistémica de la explicación a la que me refiero prescindiendo de él.

En las secciones que acabo de mencionar he sentado ya algunas bases para proponer un esquema general de una concepción epistémica de la explicación. Desde esta perspectiva, la explicación es un tipo de entidad abstracta que forma parte de nuestro

²⁰⁰La cual no es la posición de Railton.

²⁰¹Véase, *supra*, apartado *Ambigüedad Epistémica*.

²⁰²Véase, *supra*, apartado *Un Modelo No Monotónico*.

²⁰³Véase, *supra*, Introducción General.

conocimiento del mundo, sea lo que sea que entendamos por "mundo". Como en general lo es nuestro conocimiento, la explicación es una postulación, no algo establecido permanentemente. Es algo que puede corresponder con ciertos hechos del mundo, pero que no tiene que hacerlo necesariamente. Así, la representación racional de la ignorancia que mencioné recientemente constituye una parte importante de las características de una explicación vista desde esta concepción epistémica. Además de la forma en que propuse que se abordara la ambigüedad epistémica (representación racional de la ignorancia) y la no monotonicidad que la ocasiona, hay otras dos características importantes de mi concepción epistémica de la explicación: el rechazo del supuesto de alta expectabilidad y del ideal de buena explicación.

Para una concepción epistémica como ésta una explicación es esencialmente una manera de entender al mundo. La explicación se encuentra así en un plano epistémico. En el capítulo anterior usé como punto de referencia para la crítica de los modelos clásicos dos ejes: la tensión Hempel-Scriven y la tensión Hempel-Coffa. El desarrollo de la idea de explicación en esta sección me coloca en una posición que coincide ampliamente (aunque no completamente) con la posición general de Scriven. Por ejemplo, la idea de abandonar el ideal de buena explicación puede compararse con la sugerencia de Scriven acerca de separar explicación y fundamentos de la explicación.

Desde mi perspectiva una explicación no necesita corresponder a alguna realidad extraepistémica. La explicación en este sentido es una actividad intelectual falible. Es conveniente, entonces, distinguir algunos conceptos: la relación de pertinencia explicativa, y la relación de explicación. Entenderé por "**relación de pertinencia explicativa**" una relación que mantienen dos eventos. Una investigación completa sobre la explicación tendría que proponer un tipo de relación que debería ser considerada una relación de pertinencia explicativa, es decir, una relación que creemos que constituye algún criterio para determinar si algo es una explicación. Podemos decir que alguien ha propuesto que existe una relación de pertinencia explicativa entre dos eventos

independientemente de si dicha relación de hecho ocurre entre ambos eventos. Propondré, en cambio, el nombre de "**relación de explicación**" para designar no una relación entre eventos sino entre un conjunto de proposiciones llamado *explanans* y un evento, respecto de un contexto. Así, el acto de explicar un evento es el establecimiento de una relación de explicación entre un conjunto de proposiciones y un evento a explicar. Cuando alguien ofrece una relación de explicación se encuentra *postulando* una relación de pertinencia explicativa adecuada entre, por un lado, los eventos involucrados en el conjunto de proposiciones y, por otro, el evento a explicar. La relación de explicación puede cumplirse aún a pesar de que la relación de pertinencia explicativa postulada no se cumpla realmente. Finalmente, entenderé por "**explicación**" un argumento de cierto tipo cuya conclusión es la descripción del evento que se pretende explicar. Estas distinciones nos ayudarán a entender mejor la concepción epistémica que estoy proponiendo.

En esta sección ilustraré brevemente, con un ejemplo científico famoso, la dimensión epistémica en la que asumo que se desenvuelven generalmente las explicaciones. Mi objetivo será ilustrar este comportamiento de algunas explicaciones concebidas epistémicamente, que está relacionado con el aumento de información: la no monotonicidad. Para ello, expondré algunos casos del tipo NM1a y un caso de tipo NM1b. No me extenderé en variantes de los casos NM1b pues una exposición interesante de ellas exigiría ya una idea más concreta del modelo formal de explicación que no presentaré en esta investigación.

El comportamiento no monotónico no es peculiar del ámbito científico. También puede observarse en razonamientos que nos son más familiares: los razonamientos del sentido común.

En el razonamiento científico como en el del sentido común, a menudo hacemos inferencias que son retractables. Una consecuencia retractable es una consecuencia que puede dejar de inferirse aunque no perdamos premisas.

Formalmente hablando, tener consecuencias retractables implica estar usando una noción de consecuencia que viola el Principio de Monotonidad. Si una noción de consecuencia cualquiera \vdash cumple con la siguiente condición, decimos que la noción de consecuencia tiene la propiedad de ser **monotónica**:

Si a partir del conjunto D se infiere p , pero D es un subconjunto de F , entonces también de F se infiere p . Puesto en fórmulas:

Si $D \vdash p$, y tenemos que $D \subseteq F$, entonces $F \vdash p$.

En una noción de consecuencia no monotónica cabe la posibilidad de que, una vez que hemos aumentado información en nuestro conjunto de premisas, perdamos algunas de nuestras antiguas conclusiones. En símbolos:

Aunque $D \subseteq F$, y $D \vdash p$, no necesariamente $F \vdash p$.

A pesar de que en el razonamiento del sentido común hallamos con mayor facilidad ejemplos de no monotonicidad, no expondré estos casos hasta la sección 3.2. La razón es que en dicha sección explicaré un sistema concreto de lógica no monotónica a partir del razonamiento del sentido común; y, además, que para motivar el estudio de las lógicas no monotónicas en relación a esta investigación es suficiente y más adecuado mostrar ejemplos de no monotonicidad en la explicación científica. Sin embargo, terminaré esta sección con una introducción de la relación entre el sentido común y algunos conceptos importantes relacionados con la no monotonicidad. Identificaré dichos conceptos mediante algunos ejemplos de razonamiento del sentido común que serán desarrollados con mayor precisión a lo largo del capítulo.

El objetivo central del capítulo será proporcionar una introducción a los razonamientos que poseen la propiedad de no monotonicidad y, también, contribuir a la

familiarización con los principales elementos formales que pueden servir para construir un modelo formal de explicación.

3.3.1 No Monotonicidad en la Ciencia

Veamos un ejemplo paradigmático de explicación científica. En su *Filosofía de la Ciencia Natural* [Hempel 1977]²⁰⁴ Hempel presenta el caso histórico de las investigaciones de Ignaz Semmelweis acerca de la fiebre puerperal. Aunque Hempel usa este caso histórico principalmente para ilustrar la estructura lógica mediante la cual se someten a confirmación las hipótesis científicas, el caso ilustra también algunos aspectos concernientes a la explicación científica. Describiré muy brevemente el caso de Semmelweis basándome en la exposición de Hempel [Hempel 1977, pp 16-25].²⁰⁵

A mediados del siglo XIX se había detectado un problema en el Hospital General de Viena. Mientras que en la Primera División de Maternidad del hospital moría por fiebre puerperal (o fiebre de sobreparto) aproximadamente entre el 7 y el 11 por ciento de las mujeres que ingresaban, en la División Segunda sólo se registraba una mortandad de entre el 2 y el 3 por ciento. La situación era aún más enigmática si se consideraba que en la División Primera los encargados de atender a las mujeres parturientas eran los estudiantes de medicina mientras que en la División Segunda atendían comadronas.

Semmelweis se propuso investigar la razón de esta diferencia de mortandad. Su investigación se abocó entonces a localizar supuestos factores que pudieran estar influyendo en el incremento de mortandad por fiebre puerperal en la División Primera respecto de la División Segunda. Así, Semmelweis enfrentó varias hipótesis y las sometió a contrastación. Finalmente, de forma accidental, se encontró con cierta evidencia que lo condujo a la invención de una nueva hipótesis. Esta hipótesis resultó confirmada en algún grado y, consecuentemente, Semmelweis decidió que se trataba de la hipótesis correcta.

²⁰⁴Hempel, Carl G., *Filosofía de la Ciencia Natural*, Alianza Universidad, Madrid, 1977.

²⁰⁵*Ibid.*, pp. 16-25.

Cada una de estas hipótesis era, para Semmelweis, una forma de explicar la cantidad de muertes en la Primera División. La mayoría de ellas fracasaron. En el esquema hempeliano D-N el *explanandum* del caso de Semmelweis puede tomarse como un enunciado que habla sobre cierto porcentaje de mortandad por fiebre puerperal. El enunciado *explanandum* puede expresarse, por ejemplo, como "El porcentaje de muerte por fiebre puerperal en la División Primera está entre el 4 y el 9 más que en la División Segunda.

Una de las hipótesis fallidas fue que la posición en que yacían las pacientes era la causa del incremento de mortandad por fiebre puerperal. Mientras que las pacientes de la División Segunda yacían de costado, las pacientes de la División Primera yacían de espaldas. La explicación bajo esta hipótesis puede formularse simplificada en el siguiente razonamiento D-N:

(Posición)

A) Los miembros de un grupo de pacientes parturientas yacen de espaldas si y sólo si el porcentaje de muerte por fiebre puerperal en dicho grupo está entre el 4 y el 9 más que en grupos en los que no yacen de espaldas.

B) La División Primera es un grupo cuyos miembros son pacientes parturientas que yacen de espaldas.

C) El porcentaje de muerte por fiebre puerperal en la División Primera está entre el 4 y el 9 más que en grupos en los que no yacen de espaldas.

Donde A representa la hipótesis de que yacer de espaldas es la causa del incremento mencionado.²⁰⁶

El argumento es deductivamente válido. Una vez que hemos supuesto que las premisas de un argumento como éste son verdaderas, cualquier información que añadamos al respecto es irrelevante para la verdad de la conclusión. Ninguna información puede incapacitar la relación derivacional deductiva para que la conclusión siga derivándose. Esta propiedad de los argumentos deductivos es la llamada "monotonidad".

No obstante, Semmelweis concluyó que la hipótesis A era falsa. El experimento que llevó a Semmelweis a pensar que A era falsa fue hacer que las pacientes de la División Primera yacieran de costado. A pesar de esto, corroboró que la muerte por fiebre puerperal se mantuvo entre el 4 y el 9 por ciento. Hempel interpreta esto como un *modus tollens* a partir, simplificadaamente, del condicional $A \supset (\neg B \supset \neg C)$ y la negación del condicional que aparece en el consecuente, en palabras quizá más familiares, a partir de que 1) si A fuera cierta entonces, si las pacientes en cuestión yacieran de costado entonces el porcentaje de mortandad por fiebre puerperal disminuiría en la División Primera hasta aproximarse a su porcentaje en la División Segunda y 2) el resultado de que cuando las pacientes de la División Primera yacieron de costado los resultados de la mortandad no disminuyeron. Ante esta evidencia, Semmelweis concluyó que la hipótesis era falsa.

Hempel entiende el resultado del experimento como la negación de $\neg B \supset \neg C$. Esta interpretación hace posible que se siga la refutación de A²⁰⁷. El efecto que tuvo el

²⁰⁶Omitiré aquí el análisis de dos cuestiones: i) si en el caso estudiado, B representa el único factor causal que produce C o existen algunos otros que también pueden producirlo y ii) si dado B se sigue deductivamente C o se sigue tan sólo con cierta probabilidad menor que 1. Al parecer, Semmelweis mantiene i dado que, como veremos enseguida, asume que $C \supset B$, esto es, que es necesario el factor causal B para que se produzca C. Esto apoya mi interpretación bicondicional en A del ejemplo *Posición*.

experimento en la investigación de Semmelweis fue que éste abandonó la hipótesis A y asumió la información de que $\neg A$. Semmelweis asume que la hipótesis A era falsa.

Desde el punto de vista del modelo D-N, podemos interpretar todo esto como el hecho de que una de las premisas de *Posición* es falsa. Por lo tanto, *Posición* no cumple con la condición 6 del modelo según [Hempel & Oppenheim 1948]²⁰⁸. Se sigue que no es una explicación.

Posteriormente, un colega de Semmelweis sufrió una herida con un escalpelo que había sido usado para realizar una autopsia. Este individuo murió después mostrando los mismos síntomas que los que mostraban las pacientes con fiebre puerperal. Semmelweis cayó en la cuenta de que, en general, él y su equipo revisaban a las pacientes parturientas después de realizar prácticas en la sala de autopsias. Por ello, pensó que el factor que contribuía al incremento de la mortandad por fiebre puerperal era la "materia cadavérica" introducida en el torrente sanguíneo de las pacientes. Una explicación D-N para C bajo esta nueva hipótesis quedaría como sigue:

(Contaminación)

E) Los miembros de un grupo de pacientes parturientas, al someterlos a revisión, se les introduce materia cadavérica en el torrente sanguíneo, si y sólo si el porcentaje de muerte estará entre el 4 y el 9 más que en grupos en los que no hay introducción de materia cadavérica.

F) La División Primera es un grupo cuyos miembros son pacientes parturientas a las que se les ha introducido materia cadavérica en el torrente sanguíneo.

²⁰⁷Para evitar problemas con la tesis Duhem-Quine, supondremos que las hipótesis auxiliares son verdaderas. Así, la negación de la implicación $\neg B \supset \neg C$ implicaría la refutación de A. Por otro lado, esta interpretación puede que implique la refutación de A pero no implica la refutación de uno de sus componentes condicionales, a saber, $B \supset C$.

²⁰⁸Véase, *supra*, sección 1.2.

C) El porcentaje de muerte por fiebre puerperal en la División Primera está entre el 4 y el 9 más que en grupos en los que no hay introducción de materia cadavérica.

Análogamente al caso anterior, la suposición de que se trata de una causa única, se representa mediante el bicondicional E.

Este argumento tiene, como el anterior, la propiedad de ser monotónico, si ambos son considerados argumentos deductivos. Cualquier información que añadamos a E y a F es irrelevante para la derivación de su conclusión, C. La forma en que Semmelweis razonó para construir una contrastación de esta nueva hipótesis E fue como sigue: Si E fuera cierta entonces, si hacemos que todo el equipo que revisa a las pacientes de la División Primera se lave las manos con una solución de cal clorurada entonces el porcentaje de mortandad por fiebre puerperal disminuirá en la División Primera hasta aproximarse a su porcentaje en la División Segunda. El experimento se realizó. Los médicos y estudiantes que revisaban a las pacientes parturientas se lavaron a partir de entonces con esta solución. El resultado fue una disminución de la mortandad en la División Primera hasta quedar incluso por debajo del porcentaje de la División Segunda. Ante esta evidencia, Semmelweis consideró que la hipótesis E estaba confirmada y, así, que la materia cadavérica era el factor que incrementaba la mortandad por fiebre puerperal en la División Primera.

De nuevo, la "implicación contrastadora", como la llama Hempel, está presente en el experimento de Semmelweis. Ahora la implicación en juego es $E \supset (\neg F \supset \neg C)$. El procedimiento de Semmelweis es contrastar $\neg F \supset \neg C$. En esta ocasión el resultado es favorable. $\neg F \supset \neg C$ es verdadera y esto es tomado como una confirmación de E. Como señala Hempel, esto constituye el apoyo inductivo de la hipótesis E.

Con base en el caso de Semmelweis, mostraré un caso de NM1a y un caso de NM1b. El caso de NM1a está relacionado con el origen de la investigación de Semmelweis.

Así que es conveniente que lo exponga ahora. Semmelweis no hubiera efectuado su investigación si no hubiera algo que explicar. Su investigación se debe a que sus expectativas acerca de lo que debería haber ocurrido en una situación como la descrita no fueron satisfechas. Semmelweis esperaba que la División Primera tuviera menos mortandad que la División Segunda. Esto puede entenderse como un caso de no monotonicidad en una explicación que era considerada normal de acuerdo a los datos disponibles. Véamoslo con mayor detalle.

Sin tomar en cuenta los datos que causan la sorpresa de Semmelweis, es decir, si él no los conociera, y alguien le hubiera preguntado cuál de las divisiones tenía menores índices de mortandad, él hubiera contestado que la División Primera. Semmelweis podría explicar su respuesta diciendo que las capacidades científicas de los estudiantes de medicina superaban con mucho las de las comadronas y esto causaba que hubiera menos muertes en la división donde ellos practicaban que en la división en la que las comadronas lo hacían. Su explicación puede verse como el siguiente argumento:

(Medicina)

- a) Siempre que los pacientes son atendidos por personas que conocen lo suficiente acerca de medicina, los pacientes se mueren en menos casos que los pacientes atendidos por personas que no conocen suficientemente acerca de medicina.
- b) Los pacientes de la División Primera son atendidos por personas que conocen lo suficiente acerca de medicina,
- c) Los pacientes de la División Segunda son atendidos por personas que no conocen lo suficiente acerca de medicina.

[Deductivamente:]

- d) Los pacientes de la División Primera mueren en menos casos que los pacientes atendidos por la División Segunda.

Sin embargo, Semmelweis se ha enterado de que, en realidad, los pacientes de la División Segunda mueren en menos casos que los pacientes atendidos por la División Primera. Es muy plausible que Semmelweis haya tenido en mente algo como el razonamiento *Medicina* y los datos a,b,c. Consecuentemente, la inferencia natural en él es d. Semmelweis también asumía d, de no ser así no habría ningún motivo claro para realizar su investigación. Pero ahora esto no es compatible con la evidencia. Semmelweis ya no cree en d. Esto no puede modelarse adecuadamente si concebimos la relación inferencial en *Medicina* como una relación deductiva. Esto puede modelarse mediante esquemas de argumento no monotónicos. Después de su investigación, Semmelweis puede recurrir a lo que Hempel llamó "salvedad" para bloquear la inferencia del esquema *Medicina*. La relación inferencial en *Medicina* no se vería entonces como una relación deductiva:

(*Medicina-NM*)

- a) Siempre que los pacientes son atendidos por personas que conocen lo suficiente acerca de medicina, los pacientes se mueren en menos casos que los pacientes atendidos por personas que no conocen suficientemente acerca de medicina.
- b) Los pacientes de la División Primera son atendidos por personas que conocen lo suficiente acerca de medicina,
- c) Los pacientes de la División Segunda son atendidos por personas que no conocen lo suficiente acerca de medicina.

[Plausiblemente:]

- d) Los pacientes de la División Primera mueren en menos casos que los pacientes atendidos por la División Segunda.

En el razonamiento *Medicina-NM* la conclusión se sigue plausiblemente. Podemos entender esta noción de derivación como una que puede bloquearse si hay nueva evidencia relevante. Por ejemplo, el dato de que los estudiantes de medicina de la División Primera realizaban prácticas en la sala de autopsias justo antes de revisar a las pacientes

es relevante en este caso. Podemos entender "plausiblemente" como "en ausencia de evidencia en contra". Así, las prácticas en la sala de autopsias constituye evidencia en contra y, por lo tanto, la conclusión que antes se obtenía, no se obtiene más.

Medicina-NM es un caso de explicación no monotónica. Por supuesto, esto está estrechamente relacionado con el caso de las salvedades propuestas por Hempel²⁰⁹. De acuerdo a la formulación de *Medicina-NM*, puede haber excepciones para la conclusión d. El caso del magneto presentado por Hempel en su artículo posterior, puede representarse similarmente al caso *Medicina-NM*. El caso *Rec* de la ambigüedad epistémica también puede representarse de la misma forma. Una forma mucho más sofisticada de este tipo de razonamientos se verá en la siguiente sección y en el capítulo IV haré explícitas las diferencias de una representación como la expuesta aquí y la representación que explicaré en esa sección.

Ahora bien, obsérvese que si Semmelweis concibiera una explicación como *Medicina*, con una relación inferencial deductiva, tal vez concluyera que habría que abandonar tal argumento. Puesto que en realidad se trata de un argumento inválido. Esto no necesariamente sucede con la explicación *Medicina-NM*. Se ha bloqueado la inferencia pero esto no implica que tengamos que deshacernos de *Medicina-NM*. Si la información vuelve a cambiar, *Medicina-NM* puede volverse a usar. Esta es una flexibilidad que no posee la lógica clásica. Más aún, *Medicina-NM* continúa siendo una explicación plausible para el fenómeno representado por d. Desde mi perspectiva, algo es una explicación cuando propone una relación de pertinencia explicativa adecuada, respecto de cierto contexto. Desde el contexto de las creencias de Semmelweis, los datos a,b y c pueden establecer una relación causal respecto de d, independientemente de que de hecho no se relacionen causalmente. Si para Semmelweis la relación causal es una relación de pertinencia adecuada, entonces *Medicina-NM* sigue siendo una explicación en ese contexto, aunque d ya no se pueda inferir más, de nuevo, en ese contexto.

²⁰⁹Aunque en [Hempel 1965] aún no había caído en la cuenta.

Los casos NM1b sí tienen que ver con un cambio de calidad en los argumentos candidatos a explicación. Algo que era una explicación, deja de serlo. Veremos ahora cómo también el caso de Semmelweis nos proporciona un caso de NM1b. Recordemos el caso *Posición* y el caso *Contaminación*. Cuando Semmelweis descubre el factor de la materia cadavérica hay que evaluar de nuevo ambas explicaciones.

Ahora bien, ¿Cómo ha de considerarse el razonamiento *Posición* en relación a este nuevo argumento?, ¿Es o no una explicación? Por lo que se expuso en la sección 1.2, desde el punto de vista de Hempel *Posición* no puede considerarse una explicación, pues no cumple con el requerimiento empírico, a saber, que todas sus premisas deben ser verdaderas²¹⁰.

Sin embargo, supongamos ahora que la hipótesis A hubiera sido confirmada aunque fuera transitoriamente. Por ejemplo que, casualmente, en el mismo período en que las pacientes yacieron de costado por disposición de Semmelweis, los estudiantes hubieran decidido lavarse las manos con una solución de cal clorurada por alguna sugerencia supersticiosa. Por ejemplo si se hubiese corrido el rumor entre ellos de que uno debía lavarse bien las manos después de tocar un cadáver, para evitar así algún tipo de desgracia para ellos o para alguna persona que ellos tocaran después. Como es de suponerse, el hecho de lavarse las manos provocaría una disminución de la mortandad por fiebre puerperal. Si Semmelweis no se encontrara enterado de la nueva forma de prepararse para revisar a las pacientes después de haber ido a la sala de autopsias, lo cual es natural pues ningún estudiante tendría la osadía de explicar a Semmelweis los orígenes supersticiosos de este acto de limpieza, pensaría que la hipótesis A había sido confirmada.

Las confirmaciones transitorias suceden normalmente en la ciencia, aunque no sea por medio de estos hechos tan accidentales y de estas creencias supersticiosas. La historia

²¹⁰ A es considerada falsa, dada la consideración de la verdad de $\neg A$.

de la ciencia registra casos en los que hipótesis consideradas actualmente falsas fueron consideradas hipótesis confirmadas hasta no hallarse con problemas graves de adecuación empírica o de coherencia teórica. Obsérvese que, en última instancia, los problemas siempre son de cierto tipo de coherencia teórica, lo empírico no tiene efecto para falsar a menos que se incluya en la teoría nueva información acerca de ello. Esto es representable como un cambio en el conjunto de conocimientos. Esta representación significa ya un acercamiento a una concepción epistémica de la explicación. Lo interesante, sin embargo, es que si Semmelweis hubiera considerado confirmada la hipótesis A, lo cual no es un tipo anormal de proceder en la ciencia, entonces el razonamiento *Posición* debería haber sido considerado como una explicación.

El punto importante es que en muchas ocasiones trabajamos con información incompleta. En el caso de la superstición Semmelweis podía haber hecho que las pacientes yacieran de costado en los años sucesivos. Las explicaciones en el sentido en que Coffa y Railton las quieren caracterizar, generalmente no ocurren. Por lo general no es posible estar seguros de la verdad de las premisas de una explicación. Y, en general, no tenemos toda la información relevante a la ocurrencia del *explanandum*.

El modelo D-N supone información completa. Dada esta situación, a saber, que nuestras explicaciones en general se componen de información incompleta, si quisiéramos mantener un modelo de explicación que suponga información completa, gran parte del conjunto de cosas que consideramos explicaciones, sobre todo en el ámbito empírico, no serían capturadas. Por tanto, el modelo sería gravemente inadecuado.

En contraste, desde mi punto de vista ocurre algo distinto. Se está proponiendo que hay una relación causal entre el evento de la posición de las pacientes y el evento del incremento de la mortandad por fiebre puerperal. De acuerdo a las definiciones que dí al principio de esta sección si esta relación causal se supone como una relación de pertinencia explicativa, entonces se está proponiendo que hay una relación de explicación

entre el conjunto $\{A, B\}$ y el evento descrito por C. Es difícil considerar que la relación propuesta en *Posición* no sea una relación adecuada según Semmelweis, dado que él parece haber asumido que si se hubiese confirmado $\neg B \supset \neg C$, A hubiera estado confirmada también. Esta relación puede ser considerada bajo la perspectiva de Semmelweis como una relación adecuada en el sentido de que el tipo de eventos involucrados pueden mantenerla aunque de hecho no la mantengan. Otra cosa es decir que *Posición* no es en realidad una explicación *correcta* o, para decirlo en términos como los que usamos en la sección 1.3, que no es una buena explicación, aduciendo que es falso que exista la relación de pertinencia explicativa postulada. Pero esto es independiente de su calidad de explicación.

Una interpretación como la que propongo para el razonamiento *Posición* supone una concepción epistémica de la explicación. Nos coloca en una perspectiva desde la cual algo es una explicación a partir del tipo de postulaciones que involucra y no de la confirmación de dichas postulaciones.

¿Cómo es que, bajo el modelo Hempeliano de explicación, podría representarse un cambio de calidad en el razonamiento *Posición*? Es decir, ¿Cómo podemos representar el hecho de que *Posición* dejó de ser una explicación? Como dije pocos párrafos antes, en el marco conceptual propuesto por Hempel y Oppenheim lo que ocurre es que *Posición* nunca fue una explicación. No se cumple que toda premisa sea verdadera y, por tanto, nunca fue una explicación. En el modelo D-N, si algo es una explicación no puede dejar de serlo. Desde mi punto de vista, en cambio, *Posición* es una explicación pero esto no implica que no pueda dejar de serlo. El que algo sea o no una explicación depende de la información de nuestro contexto. Por ello, considero que el modelo I-S, lejos de ser una anomalía, es el resultado de un encuentro con la naturaleza epistémica de las explicaciones.

Desde mi perspectiva, un modelo de explicación tendría que capturar cómo un argumento que era considerado una explicación dejó de serlo. De acuerdo a mi perspectiva esto depende de si la explicación cumple con la postulación de una relación de pertinencia explicativa. Así, debemos distinguir dos cosas: la idea de que un argumento explicativo puede ser no monotónico, es decir, la idea de que puede no derivar su conclusión, de la idea de que ese argumento dejó de ser un argumento explicativo. En las siguientes líneas intentaré ilustrar esta distinción.

Hay una forma de interpretar el caso Semmelweis de una manera interesante que obedece también a una concepción epistémica. Desde el punto de vista de Hempel la explicación D-N no es posible si no introducimos una hipótesis general como premisa, por ejemplo, A o E. Dejando de lado un análisis acerca de si es posible construir explicaciones que no tengan hipótesis generales, la hipótesis, como el mismo Hempel se ha percatado²¹¹, puede que no establezca una relación estrictamente universal: la hipótesis puede no establecer una relación deductiva. Aún más, si consideramos la idea de que en nuestras explicaciones en general trabajamos con información incompleta, puede suceder que la hipótesis no esté completamente establecida. Las explicaciones pueden contener tan sólo la posibilidad de algunas hipótesis, y no las hipótesis mismas.

Supongamos que Semmelweis tiene la evidencia de que casi todas las pacientes de la División Primera, que han muerto, yacieron por lo general de espaldas. Así, postula la hipótesis de que cuando alguien yace de espaldas muere por fiebre puerperal. Supongamos ahora que la paciente número 13 murió por fiebre puerperal. Ante estos datos, Semmelweis concluirá que la causa es que la paciente número 13 yació por lo general de espaldas. Ante la evidencia provocada por la asepsia clandestina de los practicantes, Semmelweis pudo haber formulado la siguiente explicación más coloquial:

²¹¹Y por ello construyó el modelo I-S.

(Número 13)

G) La paciente número 13 de la División Primera por lo general yació de espaldas

A') En las circunstancias en que se encuentran las pacientes de la División Primera, *es posible* ♣{asumir} que cuando alguien yazca de espaldas muera por fiebre puerperal

H) La paciente número 13 de la División Primera murió por fiebre puerperal

Por supuesto, el razonamiento *Número 13* puede ser considerado por Hempel como una explicación incompleta. Pero hay un sentido en el que el razonamiento en cuestión es una explicación. En este caso, es muy parecido al caso de Paresia analizado en el capítulo 1. Subyace a los dos la idea de una baja probabilidad para la explicación del *explanandum*. Ante el hecho de que las otras pacientes, las de la División Segunda, yacen de costado, la apelación a la forma en que yacía la paciente número 13 es una explicación. Pues supone cierta relación causal entre la posición de la paciente y su muerte. En este sentido la proposición G es una razón para creer en la ocurrencia del hecho descrito por la proposición H.

Nótese sin embargo que en el razonamiento *Número 13* no se da por hecho una hipótesis general, tan sólo se postula como algo posible. Pero tampoco estamos suponiendo que tenemos, en el conjunto de proposiciones del contexto, la proposición de que tal relación de causalidad supuesta por A'. El razonamiento *Número 13* es independiente de que aceptemos la hipótesis A en el conjunto de nuestras creencias. Podemos interpretar al argumento como no deductivo. Observemos también que basta el dato más general de que no es incoherente pensar que dada una cierta posición, en ciertas circunstancias, esta posición podría estar vinculada causalmente con una muerte, para que el razonamiento *Número 13* se encuentre postulando una relación causal

adecuada y, por tanto, no parece haber nada que impida catalogarlo como una explicación desde mi perspectiva. Aunque quizá *Número 13* pueda considerarse como una mala explicación, no deja por ello de ser una explicación en el sentido que propuse en la sección 1.3. Ello no implica la aceptación de A como verdadera. A' no implica la verdad de A. G es una proposición acerca de una circunstancia más particular. La verdad de A puede aún dejarse en suspenso en espera de mayor confirmación. De hecho esto no está muy lejano de lo que hubiera pasado en una situación real, pues a Semmelweis la hipótesis A no le parecía muy plausible.

Volvamos ahora al caso ficticio de las creencias supersticiosas. Supongamos que, después de que se ha confirmado en algún grado la hipótesis A a través de los clandestinos actos supersticiosos, sucediera que Semmelweis descubriera estas acciones de los practicantes. En tal caso Semmelweis no tendría sólo evidencia a favor de la hipótesis A sino también a favor de la hipótesis E y a favor de una nueva hipótesis M que puede formularse así: "Si a los miembros de un grupo de pacientes parturientas, al someterlos a revisión, se les pone en contacto con un mal agüero, entonces el porcentaje de muerte estará entre el 4 y el 9 más que en grupos a los que no se les ha sometido a un mal agüero". Obsérvese que los hechos que funcionaron como evidencia no han cambiado: sigue siendo cierto que durante el período en que las pacientes yacieron de costado hubo una disminución en la mortandad por fiebre puerperal, aunque esto sólo haya sido casual. Ahora Semmelweis sabe algo más: durante el período en que los estudiantes se lavaron las manos con cal clorurada hubo una disminución en la mortandad por fiebre puerperal.

Por supuesto, en una situación como ésta Semmelweis comenzaría por aislar algunos factores y contrastarlos por separado. Supongamos que, en una decisión fatal, Semmelweis prueba primero con la hipótesis A. Prohíbe a los estudiantes lavarse las manos y ordena que las pacientes yazcan de costado. Es de esperarse que durante este período la mortandad por fiebre puerperal volviera a alcanzar su terrible nivel inicial.

Semmelweis tiene ahora un dato más: durante el período en que las pacientes yacieron de costado y los estudiantes no se lavaron las manos con una solución de cal clorurada, la mortandad por fiebre puerperal aumentó. Lo cual significa que Semmelweis posee evidencia en contra de la hipótesis A.

Omitamos, por simplicidad, la duda respecto de la veracidad de cualquier hipótesis auxiliar que pudiera suponerse. A pesar de estos nuevos datos, sigue existiendo la coherencia acerca del tipo de cosas que pueden mantener una relación causal. La postulación de la relación causal en *Número 13* sigue siendo adecuada pues postula una relación de pertinencia explicativa. Si Semmelweis no hubiera pensado que una relación causal entre el hecho de que alguien yazca de espaldas y el hecho de que muera por fiebre puerperal es posible, no la hubiera considerado como hipótesis. Así, aunque A es implausible, A' no. Esto se debe a que no estamos suponiendo que A' se encuentra en el conjunto de proposiciones del contexto, como algo que fuera verdadero. Sin embargo, no estaríamos dispuestos a derivar justificadamente H a partir de G. Esto puede interpretarse como un caso en el que la explicación se ha bloqueado por la introducción de nuevos datos a la teoría. Ante la nueva información de que el caso particular de relación causal descrito por A es falso, C no se infiere más. Esto aun en el caso de que *Número 13* sea todavía considerado como una explicación, mala por cierto.

Bajo esta interpretación epistémica la no monotonicidad que aparece en las explicaciones I-S puede también presentarse al menos en otro caso de explicación. Así, tenemos al menos dos casos fácilmente identificables de no monotonicidad en la explicación: las explicaciones I-S (*Rec*) y las explicaciones de probabilidad sin hipótesis generales (*Número 13*).

3.3.2 No Monotonicidad y Relación de Pertinencia Explicativa.

Ahora bien, continuando con la ficción, Semmelweis tendría que vérselas con la tarea de discernir lo que pasa con los siguientes dos argumentos que corresponden cada uno a casos de explicación sin hipótesis generales, que suponen relaciones causales vinculadas con las hipótesis E y M, separadamente.

(Número 13 científico (N13C))

G') A la paciente número 13 de la División Primera se le introdujo materia cadavérica en el torrente sanguíneo.

E') Es posible que si a alguien se le introduce materia cadavérica en el torrente sanguíneo, muera por fiebre puerperal.

H) La paciente número 13 de la División Primera murió por fiebre puerperal

(Número 13 mágico (N13M))

G'') A la paciente número 13 de la División Primera se le tocó con manos de mal agüero.

M'') Es posible que si a alguien se le toque con manos de mal agüero, muera por fiebre puerperal.

H) La paciente número 13 de la División Primera murió por fiebre puerperal

El problema con los argumentos *N13C* y *N13M* es que es difícil pensar en una contrastación que nos ayudara a discriminar uno en favor del otro: el mal agüero está basado en no haberse lavado las manos y esto es lo que constituye el evento que serviría

para constrastrarlos. Es decir, si Semmelweis ordena que los practicantes se laven las manos y que las pacientes permanezcan yaciendo de espaldas, ambas hipótesis E y M estarían confirmándose. Y, por el contrario, si Semmelweis ordena que los practicantes no se laven las manos y que las pacientes permanezcan yaciendo de espaldas, ambas hipótesis E y M estarían desconfirmándose. A primera vista, no hay cómo discriminar alguna de las dos hipótesis.

Sin soslayar lo anterior, estamos seguros que Semmelweis discriminaría la hipótesis M en favor de la E. Aún más, estamos seguros de que Semmelweis no consideraría a *N13M* como una explicación, ¿Por qué?. La razón es que en realidad *N13M* no sólo está proponiendo que se ha tocado a las pacientes con manos de mal agüero sino que, además, la propiedad de ser de mal agüero es la causa de la muerte. En el sistema de creencias de Semmelweis tal tipo de eventos mágicos no tienen la posibilidad de mantener relaciones causales con, por ejemplo, la muerte de alguien²¹². No sólo M no es plausible sino que M' es falsa: no es posible, desde la perspectiva de Semmelweis, admitir que el mal agüero pueda causar la muerte. Así, si consideramos que Semmelweis piensa que si no hay relaciones causales no hay relación de pertinencia explicativa entonces la relación de pertinencia explicativa postulada en *N13M* no es adecuada al contexto de Semmelweis. Consecuentemente, *N13M* no es una explicación desde mi perspectiva. Nótese la diferencia entre *Número 13* y *N13M*: en el primero hay una relación de pertinencia explicativa que es adecuada, en el segundo, no. Nótese también que, aun cuando *Número 13* es una explicación bajo la perspectiva de Semmelweis, podría quedar bloqueado ante nueva evidencia.

Finalmente, para notar una característica importante más, *Número 13* podría dejar de ser una explicación bajo la perspectiva de Semmelweis. Si él añadiera a su conjunto de creencias que no puede mantenerse una relación de causalidad entre la posición de

²¹²Esto muestra, entre paréntesis, lo importante que es pensar en términos de la intención de los elementos de una explicación. Una proposición que aparece como miembro de una explicación debe reflejar toda la intención que se propuso al postular la explicación.

alguien y su muerte. En tal caso la relación de pertinencia explicativa postulada por *Número 13* sería inadecuada. Por lo tanto *Número 13* no sería ya una explicación, dejaría de serlo bajo la perspectiva de Semmelweis por el hecho de haber aceptado cierta información. Ésto es un caso NM1b de no monotonicidad.

La historia basada en la anécdota de Semmelweis muestra tres cosas: primero, cómo nuestras explicaciones pueden ser concebidas epistémicamente; segundo, cómo cierta interpretación de esta concepción epistémica de la explicación puede en ocasiones dar lugar a no monotonicidad; tercero, cómo desde esta perspectiva epistémica la explicación vista como un argumento se encuentra muy relacionada con la representación de la ignorancia de la que hablé al principio de esta sección. Respecto del último punto tenemos que una explicación puede seguir siéndolo aún a pesar de tener premisas falibles, que su *explanandum* pueda dejar de derivarse, y que el contexto puede influir tanto en este bloqueo o permisión de la derivación del *explanandum* como en la calificación de un argumento como explicación o no.

3.4 Algunas reflexiones sobre la lógica default y la no monotonicidad.

He incluido aquí un breve examen de algunas cuestiones que convendría tener en cuenta para entender cabalmente la noción de no monotonicidad, con especial referencia a la lógica del razonamiento por default, de Raymond Reiter. y las relaciones de esta noción con la inteligencia artificial, la lógica clásica y la explicación científica. Todos los puntos que serán tratados en esta sección intentarán sólo señalar detalles de interés que no podré desarrollar completamente en este trabajo. Así, el desarrollo de la sección será un intento por reflejar la relación sistemática entre algunas nociones vinculadas con la no monotonicidad.

3.4.1 Inteligencia Artificial y Explicación

Los dos principales problemas que dan origen al desarrollo de lógicas no monotónicas para la investigación en inteligencia artificial son los siguientes:

- A) El problema del *frame*.
- B) El problema de la cualificación.

El problema del *frame*, puede pensarse como el problema de saber cuáles supuestos deben modificarse y cuáles no, con base en nueva información [McCarthy & Hayes 1969]²¹³. En el mismo tenor, el problema de la cualificación puede pensarse como el problema de saber cuáles son los supuestos posibles que deben tomarse en cuenta ante la resolución de un problema [McCarthy 1977]²¹⁴. Suponiendo una base de datos que pueda representar parte importante de la situación dada, ambos problemas pueden verse, respectivamente, como las siguientes cuestiones: cómo actualizar las asunciones de nuestra base de datos ante una nueva información y cómo saber lo que es asumible en una situación dada. Obsérvese cómo en esta formulación el parentesco entre los dos problemas se hace evidente. McCarthy clasifica ambos problemas, así planteados, con un sólo nombre: "el problema de la cualificación" [*Idem*]. De lo que parece tratarse es, en todo caso, de la forma en que determinamos, a partir de una situación dada (sea una situación inicial o el resultado del cambio de esta situación inicial), lo que es asumible.

En el fondo, el problema tiene que ver con hallar criterios más o menos firmes para determinar qué información posible, en un contexto dado, es pertinente para el manejo de una situación dada. Lo anterior es una descripción que, a mi parecer, es mucho más

²¹³Su primera formulación es el problema 1 que aparece al principio del artículo: [McCarthy & Hayes 1969] McCarthy, John y Hayes, P. "Some Philosophical Problems from the Standpoint of Artificial Intelligence", en B. Meltzer and D. Michie (eds.), *It Machine Intelligence* Vol. 4, Edinburgh, N.Y., 1969 p. 468. Reimpreso en [Ginsberg, 1987] Ginsberg, Matthew L., *Nonmonotonic Reasoning*, Morgan Kaufmann Publishers, Inc., E.U.A., 1987, pp. 26-45.

²¹⁴Puede consultarse la caracterización de McCarthy en: [McCarthy 1977] McCarthy, John, "Epistemological Problems of Artificial Intelligence", en *International Joint Conference on Artificial Intelligence (IJCAI)*, 1977, pp. 1038-1044. Reimpreso en [Ginsberg, 1987] Ginsberg, Matthew L., *Nonmonotonic Reasoning*, Morgan Kaufmann Publishers, Inc., E.U.A., 1987, pp. 46-52.

real de lo que está de fondo en ambos problemas. Así, usando la noción de pertinencia, la cuestión general puede señalarse del siguiente modo: ¿Cómo determinar lo que es pertinente para sacar consecuencias a partir de una situación dada?. Llamemos a este problema el "**problema de la pertinencia general**". Pueden plantearse dos casos de este problema: Cuando la situación se toma como punto de partida; es decir, sin considerar aumentos de información (**calificación**) y, el caso contrario, cuando debe tomarse en cuenta un aumento de información en el conjunto-base²¹⁵ de datos (**frame**).

El problema de la pertinencia general no puede caracterizarse cabalmente si no se menciona que la determinación que exige tiene que hacerse con base en información incompleta: el problema de la pertinencia general surge en el marco de una situación con información incompleta. Una situación así, si se aborda de modo racional, implica dos problemas más:

- 1) El problema de determinar estructuras lógicas que permitan saltar a la conclusión²¹⁶ o, en otras palabras, que permitan conjeturar, en un razonamiento.
- 2) El problema de manipular razonamientos basados en supuestos falibles.

En realidad, lo que está en juego en ambos problemas es cómo llegar a conclusiones de manera "sensata", a partir de premisas que son insuficientes. Si entendemos "conjetura" como un razonamiento que supone un salto a la conclusión, la cuestión central se traduce en cómo hacer y manipular *conjeturas sensatas*²¹⁷. Podemos llamar a este problema general, el "**problema de la conjetura**".

²¹⁵Pensemos en el conjunto-base como el conjunto inicial de datos.

²¹⁶La noción de salto a la conclusión debería ser analizada con detenimiento, pero en general parece designar la relación de inferencia supuesta en razonamientos que concluyen a partir de premisas con las que no se puede asumir ningún esquema válido que relacione a éstas con la conclusión.

²¹⁷Aunque la propiedad de sensatez atribuida a esta clase de razonamientos descansa en nuestras intuiciones de lo que es un razonamiento aceptable, y esto sea, a su vez, problemático, el problema de cómo hacer y manipular *conjeturas sensatas* queda relativamente bien acotado ya que la cuestión de fondo en este problema es, justamente, examinar nuestras intuiciones de lo que puede ser calificado como "razonamiento aceptable".

El problema de la conjetura es una de las formas de abordar, desde un punto de vista lógico, el problema de la pertinencia general. Desde esta perspectiva, es natural considerar que los problemas surgidos en inteligencia artificial sean considerados como problemas que involucran discusiones de lógica pero también de epistemología, pues el problema de la pertinencia general tiene profundas raíces en cuestiones de índole epistemológico.

El problema de la conjetura y el problema de la pertinencia general, no sólo tienen raíces epistemológicas, también son útiles para la comprensión de algunos problemas de epistemología de la ciencia. Este es el caso del problema de la caracterización de lo que es una explicación científica. Si se piensa a la explicación como un argumento, se hace claro que el problema de la conjetura está estrechamente ligado al problema de la caracterización formal de la explicación. Dicha caracterización suele abordarse desde dos perspectivas distintas: La perspectiva-proceso y la perspectiva-producto. Aclararé ambas perspectivas en el marco de la explicación como un argumento. Desde la primera perspectiva la explicación se concibe como el proceso que conduce a la construcción de un *explanans*²¹⁸. Es decir, cómo llegamos a construir una justificación para el *explanandum*. Desde la segunda perspectiva, se concibe como un producto terminado, esto es, algo así como el resultado del proceso anterior. Aunque un análisis de las relaciones entre ambas perspectivas podría dar como consecuencia el descubrimiento de algunas relaciones de dependencia entre la explicación entendida como producto y la explicación entendida como proceso, supondré que ambas perspectivas son independientes en el sentido de que caracterizar formalmente la explicación desde una de las perspectivas no necesariamente implica su caracterización desde la otra.

A menudo la justificación que se aduce para el *explanans* o para el *explanandum* puede ser tal sólo dentro de un razonamiento conjetural. En la perspectiva-proceso

²¹⁸ Llamaré *explanans* a la parte de la explicación que tiene la función de explicar y llamaré *explanandum* a la parte que esperamos explicar. Bajo la perspectiva de la explicación como argumento, el *explanans* tiene el papel del conjunto de premisas y el *explanandum*, el de la conclusión.

intentamos determinar los factores que son pertinentes para una situación dada; específicamente, intentamos determinar los hechos que guardan ciertas relaciones con el hecho que constituye el *explanandum*. Por supuesto que esto a veces es conjetural pues los factores aducidos pueden no ser los correctos o los que realmente ocurrieron. Cuando lo vemos desde la perspectiva-producto sucede algo distinto. No usamos la conjetura para que nos lleve a encontrar los factores pertinentes al *explanandum*. De cualquier forma, en esta última perspectiva si bien no intentamos caracterizar la forma de conjeturar en busca de lo que es pertinente para el *explanandum*, sí intentamos caracterizar la forma de conjeturar, a partir del *explanans*, el *explanandum*. Una relación conjetural entre *explanans* y *explanandum* ocurre cuando los factores que aducimos no nos aseguran la ocurrencia del *explanandum*²¹⁹. Por ejemplo, en el caso de la explicación Inductiva-Estadística de Carl G. Hempel [Hempel 1964]. El fin de este último intento de caracterización puede ser el de establecer mecanismos de reconocimiento para identificar algo como una explicación.

Si uno está convencido de que la modelación de la explicación como argumento debe hacerse mediante la construcción formal de una relación conjetural, tal modelación incluirá formas de representación de los cambios de situaciones en relación con los cambios de información. Así, en el caso de la suposición de la explicación como argumento, sería conveniente contextualizar las tentativas de modelar formalmente la explicación, en el marco de la conexión entre el problema de la pertinencia general y el problema de la conjetura.

3.4.2 No monotonicidad

Intentar dar solución al problema de la conjetura nos conduce a la construcción de esquemas de razonamiento no monotónicos, pero no necesariamente nos conduce a ello. Algunos lógicos de la no monotonicidad han propuesto que un razonamiento no

²¹⁹Por supuesto, también en esta perspectiva a veces la relación inferencial entre *explanans* y *explanandum* no es conjetural.

monotónico es, en términos generales, un razonamiento cuyo conjunto-base de información (las premisas) es incompleto [Brewka & Dix & Konolige 1997]²²⁰. Si tomásemos esto como una definición nos percataríamos de que no es exacta. Hacer este ejercicio de análisis nos llevará a entender mejor la noción de razonamiento no monotónico.

La idea de tener premisas suficientes es una idea muy arraigada en lógica clásica. Un razonamiento con premisas suficientes lo es, desde esta perspectiva, si y sólo si es un **razonamiento válido**. En lógica clásica los esquemas de razonamiento válidos son los únicos esquemas *aceptables*. Entenderé el hecho de que un esquema *sea aceptado en un sistema formal* como el hecho de que tal esquema es una regla estipulada para efectuar demostraciones en dicho sistema. Cualquier razonamiento que posea premisas insuficientes debe ser entonces calificado de inaceptable en una lógica clásica²²¹.

Algunos esquemas inaceptables en lógica clásica, son aceptados por sistemas formales no clásicos. Un sistema de lógica no monotónica es un sistema formal no clásico. Ahora bien, en lógica no monotónica no querríamos incluir cualquier tipo de razonamiento inválido. Basta para ello pensar en la falacia de afirmación del consecuente (AC) comparada con el esquema (I)²²²:

(AC)	(I)
$p \supset q$	$p \supset q$
q	q
Por lo tanto:	Por lo tanto:

²²⁰Esta es la estrategia en [Brewka & Dix & Konolige 1997]. Los autores, para hablar generalmente de "razonamiento no monotónico", basan su caracterización en la idea de información incompleta. Por ejemplo en p. 1 de su trabajo conjunto: *Nonmonotonic Reasoning*, CSLI Publications, E. U., 1997.

²²¹Aunque hay lógicas no clásicas en las que sucede lo mismo. Por ejemplo, la lógica paraconsistente de Newton da Costa.

²²²En la presente sección la forma de simbolización de los ejemplos no será uniforme. Prescindiré cuando pueda del lenguaje del cálculo cuantificacional y usaré en cambio el del cálculo proposicional. Pero usaré el primero de los lenguajes cuando lo crea pertinente.

p $\neg p$

Mientras que una instancia de *AC* resultaría *plausible* aunque inválida, una instancia de *I* sería calificada de *irracional* además de inválida. Ahora comparemos el esquema *AC* con el esquema *Def*:

(AC)	(Def)
$p \supset q$	$(p \supset q) \wedge q : \mathbf{Mp}$
q	Por lo tanto:
Por lo tanto:	p
p	

Def parece más racional que *AC*, aunque ambos sean razonamientos inválidos. En realidad, cuando yo uso "razonamiento conjetural" en vez de "razonamiento a partir de información incompleta" intento enfatizar una diferencia fundamental implícita en los desarrollos de lógica no monotónica: No todo razonamiento a partir de premisas insuficientes (es decir inválido) es un razonamiento sensato. Si esta diferencia no se aclara se puede incurrir en algunos errores. La definición hipotética de razonamiento no monotónico que he sugerido al principio del apartado puede provocar estos errores. Veámoslo detenidamente.

Un sistema de lógica clásica, decía, sólo acepta esquemas de razonamiento válidos. Cuando ocurre esto, el sistema será forzosamente deductivo. Esto puede expresarse diciendo que la noción de consecuencia del sistema (lo que normalmente se designa por el símbolo " \vdash ") es deductiva. Una noción de consecuencia deductiva cumple con la propiedad de monotonicidad. Consecuentemente, ante un aumento de información como

base para la derivación, dicha noción de consecuencia mantiene todas las conclusiones a las que llegaba sin la nueva información²²³.

Consideremos ahora que λ es una lógica cualquiera. Respecto de λ , R es el conjunto de esquemas de razonamiento aceptados y B es un conjunto-base de proposiciones²²⁴. " \vdash_R " designa la noción de consecuencia de la lógica λ . \vdash_R está relativizada sólo a los esquemas de razonamiento de R . Un esquema aceptado para λ no es, necesariamente, un esquema válido. Supongamos que R puede ser dividido en dos conjuntos: Un subconjunto RV de R contendría entonces a los esquemas de inferencia deductivamente válidos y un subconjunto RI , a los deductivamente inválidos. Si $RV = R$ entonces \vdash_R es una noción de consecuencia deductiva. Si $RI = R$ podría darse el caso de que la situación no cambie. Supongamos que $RI = R = \{AC\}$. En este caso, si $B = \{p \supset q, q\}$ entonces $B \vdash_R p$. Pero también sucederá que para cualquier conjunto de proposiciones G , tengamos $B \cup G \vdash_R p$. Es decir que la noción de consecuencia es también deductiva. Lo mismo sucedería para $RI = R = \{I\}$. Así, la invalidez no asegura no monotonicidad. Dado que todo razonamiento inválido es un razonamiento con premisas insuficientes, se sigue que no basta que un razonamiento parta de premisas insuficientes para que sea un razonamiento no monotónico.

La cuestión de caracterizar un razonamiento no monotónico es una manera legítima de abordar una caracterización más general: la no monotonicidad de una noción de consecuencia. De hecho, la propiedad de ser no monotónico es una propiedad que se atribuye más directamente a nociones de consecuencia. La pregunta es entonces qué es lo que produce la no monotonicidad en una noción de consecuencia. Podría pensarse que el problema consiste en que RI no contiene **esquemas de razonamiento en conflicto**; es decir, esquemas con conclusiones contradictorias. Pero esto no es así. Para $RI = R = \{AC, I\}$

²²³Es importante tener en cuenta que no todo sistema formal deductivo es clásico. Por ejemplo, la lógica paraconsistente de da Costa es deductiva y, sin embargo, no es clásica.

²²⁴Para el uso que hago de "esquema de razonamiento" y de "conjunto-base" véase, *supra*, sección 3.2.2.

en relación con B sucede también que $B \vdash_R p$ y que $B \cup G \vdash_R p$, para cualquier G. Sabemos que el conjunto de consecuencias resultante a partir de B y el de a partir de $B \cup G$ son, cada uno, inconsistente²²⁵ pues contienen también a $\neg p$. Pero esto no obsta para que \vdash_R sea monotónica.

Finalmente, ante un caso como el inmediatamente anterior, en el que tenemos esquemas en conflicto dentro de R, podría argüirse que hace falta un **criterio de discriminación** entre esquemas en conflicto. Supongamos el criterio siguiente: Ante dos esquemas de argumento en conflicto deberá aplicarse sólo el que posea la conclusión tal que el número de ocurrencias del functor de negación sea non. Bajo este criterio, en el mismo caso de $R = \{AC, I\}$ y $B = \{p \supset q, q\}$, tendríamos que se aplicaría el esquema I y no el AC. Sin embargo, ocurriría también que $B \vdash_R \neg p$ y que $B \cup G \vdash_R \neg p$. La noción de consecuencia \vdash_R respecto de ambos R y el criterio de discriminación, seguiría siendo monotónica. Por ende, al menos no todo criterio de discriminación entre esquemas de argumento es capaz de producir no monotonicidad en su noción de consecuencia correspondiente.

He sugerido que el esquema *Def* es más sensato que el esquema AC. En realidad una distinción entre grados de *sensatez* tendría que estar fundada en relación a ciertos marcos de racionalidad. De cualquier forma parece que, bajo la racionalidad occidental dominante, es más sensato inferir p a partir de *Def* que a partir de AC pues querríamos evitar construir con nuestros esquemas un conjunto de consecuencias inconsistente. *Def*, a diferencia de AC, contiene elementos que contribuyen al impedimento de la construcción de un conjunto de consecuencias inconsistente.

²²⁵En el sentido de [Peña 1993] Peña, Lorenzo, *Introducción a las Lógicas No Clásicas*, 1993, IIF-UNAM, México. Ambos conjuntos contienen, para al menos una fórmula, también la negación de esa fórmula.

La diferencia entre *Def* y *AC* es que *Def* posee una forma de bloquear algunas de las instancias del esquema. Es decir, *Def* posee una **forma de bloquear inferencias**²²⁶. Sabemos que incorporar *Def* en *R* es suficiente para hacer que la noción de consecuencia \vdash_R sea no monotónica²²⁷. No obstante, tener simplemente una forma para **perder consecuencias** no lo es.

El Principio de No Monotonidad para una noción de consecuencia cualquiera \vdash dicta, como he dicho en otras secciones de este capítulo, que hay casos en que:

(NM)

$D \vdash p$ y, aunque $D \subseteq F$, no necesariamente $F \vdash p$

(donde *D* y *F* son conjuntos de proposiciones y *p* una fórmula)

De manera que habría que incorporar la cláusula de que la pérdida de consecuencias se efectúa mediante el aumento de información en el conjunto-base. Así, si al menos en un caso se produce la pérdida de consecuencias cumpliendo con la cláusula antedicha, entonces tenemos no monotonicidad²²⁸.

Entender este tipo de pérdida de consecuencias es clave para entender la no monotonicidad. Es pertinente para ello una nueva formulación de la distinción entre esquemas de razonamiento aceptados que son inválidos y esquemas de razonamientos que producen no monotonicidad. La distinción puede exponerse más claramente con otra distinción: la distinción entre consecuencia **obtenida faliblemente** y consecuencia **retractable**.

²²⁶Aquí "inferencias" debe tomarse en el sentido de procesos inferenciales o de relaciones de derivación pero no, como también es usual, en el sentido de consecuencias.

²²⁷Recuérdense los desarrollos de la lógica *default* expuestos en la sección anterior.

²²⁸Nótese que, bajo la noción de no monotonicidad, lo que nos debería interesar es el análisis del hecho de perder la justificación de la información, más que el análisis del hecho de perder la información.

Supongamos que \vdash_X es una relación de consecuencia lógica relativizada a un conjunto de esquemas de razonamiento aceptados X . Podemos ver la distinción entre invalidez y no monotonicidad como sigue. Una consecuencia cualquiera p tal que $A \vdash_X p$ y que puede inferirse (o "derivarse") mediante un conjunto de esquemas aceptados $Y \subseteq X$, ha sido obtenida faliblemente (o "inválidamente") respecto de \vdash_X si y sólo si todo esquema inválido de \vdash_X está en Y y no ocurre que $A \vdash_{X-Y} p$. Una consecuencia p obtenida faliblemente es retractable respecto de \vdash_X si puede no ocurrir que $A \cup G \vdash_X p$, para algún conjunto G de proposiciones. Si X contiene esquemas de razonamiento que producen no monotonicidad en \vdash_X entonces hay al menos una consecuencia p que es retractable respecto de \vdash_X . Si la conclusión de un esquema Z es retractable en relación con un conjunto de esquemas entonces decimos que Z es retractable y si su conclusión no es retractable pero se infiere faliblemente decimos que Z es falible.

Los default, son un caso de esquemas que producen no monotonicidad mediante una forma de bloqueo de inferencias relacionado con el aumento de información. Poseen conclusiones que pueden ser retractables en relación con algún conjunto de esquemas aceptados. Los defaults pueden bloquear la inferencia mediante una clase de chequeo de consistencia y, cuando la información aumenta, puede ocurrir que se retracten algunas consecuencias. Puede concebirse, como se vio en la sección 3.2., que existan otras formas distintas para bloquear inferencias con base en nueva información y que sean distintas a la forma de hacerlo con los *defaults*. Usaré, el término "**salvedad**"²²⁹, en un sentido muy general, para designar cualquier forma de bloqueo en la aplicación de inferencias, del tipo que he estado describiendo. Esto es, bloqueos de inferencia debidos a aumento de información en el conjunto-base. Usaré el término "**excepción por inconsistencia**" para

²²⁹Por supuesto, el término es en honor a Hempel, véase [Hempel 1988]. Puede verse ya desde la estipulación de este término el camino que seguiré para interpretar el problema de las salvedades planteado por Hempel. A pesar de que trataré el problema en el capítulo IV quiero advertir que la forma en que Hempel usa el término es sutilmente distinta. Esta distinción la veremos con detalle en ese mismo capítulo. Por ahora el término puede sernos de gran utilidad para la discusión en esta sección.

designar la forma de salvedad de los *defaults*, basada sólo sobre un tipo de chequeo de consistencia.

Así, una noción de consecuencia que sea no monotónica está relacionada con alguna forma de salvedad. La salvedad puede expresarse en los esquemas mismos de razonamiento que son aceptados (como en el caso de los defaults) o bien puede estipularse de alguna otra forma que afecte la noción de consecuencia en cuestión.

Si se elucida lo que se entiende por noción de consecuencia se pueden destacar algunos puntos interesantes. Dos candidatos usuales para definir una **noción de consecuencia** \vdash en los sistemas de lógica son los siguientes:

(cc) El **conjunto de consecuencias** que produce \vdash a partir de un conjunto-base B.

(ce) El **conjunto de esquemas** de razonamiento²³⁰ aceptados que supone \vdash .

Veremos lo que pasa cuando evaluamos *cc* y *ce* con relación a NM. Hay dos intuiciones contrarias igualmente arraigadas en el uso del término "noción de consecuencia". (1) Una de ellas entiende a \vdash como una misma noción de consecuencia que puede ser aplicada a distintos contextos. (2) Otra entiende a \vdash como una noción relativizada a contextos en el sentido de que si el conjunto de proposiciones que aparece a su lado izquierdo cambia, cambia también la noción de consecuencia. Desde mi punto de vista es mucho más intuitivo entender el significado de NM en relación con (1), pues es más familiar pensar que para que el enunciado NM pueda tener un sentido interesante la primera y la segunda ocurrencia del signo " \vdash " en NM deben referir a lo mismo. De lo contrario, bajo la interpretación *cc* de " \vdash ", NM diría algo que puede parecer trivial: la noción de consecuencia definida por un conjunto de consecuencias que contiene a *p*, puede derivar a *p* a partir de *D* pero puede suceder que *otra* noción de consecuencia, cuyo

²³⁰Dentro de los esquemas de razonamiento considerados para la definición de " \vdash ", pueden figurar algunos axiomas. Por simplicidad, por ahora prefiero referirme sólo a reglas.

conjunto de consecuencias no contiene a p , no pueda derivar p a partir de F (tal que $D \subseteq F$)²³¹. El candidato ce de interpretación es suficiente para sostener el sentido que nos interesa de NM y, además, es de suma utilidad para dar cuenta de una distinción que puede ser pertinente en el uso de lógicas no monotónicas para la modelación de sistemas racionales de conocimiento: la distinción entre **proponer una afirmación** y **proponer justificadamente una afirmación**. Dicha diferencia depende, como veremos más adelante, de que no establezcamos una identificación entre cc y ce como interpretaciones equivalentes para la noción de consecuencia.

Interpretar a " \vdash " en el sentido de ce nos proporciona un claro sentido para NM: con la misma noción de consecuencia puede darse el caso de que perdamos algunas consecuencias ante un aumento de información. Esto puede leerse también como "Para una misma noción de consecuencia podemos tener conjuntos de consecuencia distintos para un conjunto-base A que para un conjunto-base A incrementado". Obsérvese que de la separación entre conjuntos de consecuencia y noción de consecuencia surge un sentido claramente interesante de NM: el punto sorpresivo en NM respecto de la lógica clásica es que cambia el conjunto de consecuencias a pesar de que mantenemos inmutable el conjunto de reglas. Desde mi perspectiva, entonces, una noción de consecuencia no será el conjunto de sus consecuencias.

Ahora bien, haciendo variar el conjunto-base, una clase de distintos conjuntos de consecuencia puede producirse a partir de una sola noción de consecuencia. Si la pérdida de consecuencias se debe a pérdida de información en el conjunto-base, esto no tiene nada de sorprendente. De igual modo, la pérdida de consecuencias que está relacionada con un cambio de los elementos del conjunto-base, es completamente trivial. Estos dos casos, junto con los aumentos de información sin pérdida de la misma, constituyen tres opciones de **cambio de información en el conjunto-base**. Por otro lado, cualquier pérdida de inferencias que no esté provocada por una modificación en la información, de

²³¹Por supuesto, el contenido del paréntesis es lo que hace que una interpretación bajo cc no sea realmente trivial.

cualquiera de estos tres tipos, sería completamente irracional. Esto se debe a que si el conjunto-base no ha cambiado, no tendríamos justificación para dejar de inferir lo que antes inferíamos.

Otro punto interesante es que puede suceder que distintas nociones de consecuencia tengan propiedades comunes. Una propiedad aplicable a una noción de consecuencia es la de ser monotónica. Además de la monotonicidad hay otras propiedades que pueden ser comunes a varias nociones de consecuencia. A estas propiedades se les llama a veces "**propiedades estructurales**" de una noción de consecuencia²³². Un conjunto de estas propiedades puede determinarnos un conjunto de nociones de consecuencia distintas. Esto es debido a que dos nociones de consecuencia, a pesar de estar determinadas por conjuntos distintos de esquemas aceptados, pueden tener las mismas propiedades estructurales.

El objeto del cual predicamos las propiedades estructurales es una noción de consecuencia. Será clarificador entonces establecer una noción, diferente a la noción de consecuencia, para la clase de nociones de consecuencia. Diremos que un conjunto de propiedades estructurales define un **tipo de derivación**. En este sentido extensional, un tipo de derivación no es una noción de consecuencia sino una clase que puede abarcar distintas nociones de consecuencia.

Si tomamos como único punto de referencia a la propiedad de monotonicidad, podemos distinguir dos clases de nociones de consecuencia: Las que corresponden a un tipo de derivación monotónico y las que corresponden a uno no monotónico. Una noción de consecuencia deductiva es siempre monotónica pero hay nociones de consecuencia no deductivas que no son monotónicas.

²³²Otras propiedades estructurales además de la monotonicidad son la de corte, la de compacidad, y la de reflexividad. El término "propiedades estructurales" está influido por el de "reglas estructurales".

Tocaré ahora el punto de la distinción entre **tener una afirmación** y **tenerla justificadamente**. Está visto que para un par de conjuntos de consecuencias S, T , aún a pesar de que $S = T$ puede darse el caso de que correspondan a distintas nociones de consecuencia \vdash_X y \vdash_Y . Si esto es así, ¿Sería legítimo considerar que \vdash_X es lo mismo que tener \vdash_Y como noción de consecuencia? Mi respuesta es no. En ocasiones nos interesará, además de tener una afirmación en nuestro conjunto de consecuencias, la justificación²³³ de tal afirmación. Consideremos el esquema AC en comparación con el esquema I' :

(AC)	(I')
$p \supset q$	$\neg p \supset q$
q	q
Por lo tanto:	Por lo tanto:
p	p

Supongamos que $X = \{AC\}$, $Y = \{I'\}$ y el conjunto-base $B = \{p \supset q, \neg p \supset q, q\}$. El conjunto de consecuencias de B mediante \vdash_X y el de B mediante \vdash_Y son iguales. Sin embargo, hay una diferencia importante en la justificación de p mediante \vdash_X respecto de la justificación de p mediante \vdash_Y . Si creemos que la justificación de una afirmación es importante para una cierta modelación, no querríamos identificar \vdash_X con \vdash_Y pues AC parece un esquema más sensato que I' . Como puede verse, la interpretación ce facilita la distinción.

Finalmente dos consecuencias de toda la perspectiva construida en este apartado, en relación con nuestra forma de entender la lógica de Reiter, son las siguientes: Primero, no interpretaremos la pérdida de inferencias debida sólo a un aumento de reglas en una lógica no monotónica, como un caso de no monotonicidad, puesto que esto sería, desde esta perspectiva, un cambio de noción de consecuencia y queremos que la interpretación

²³³Aquí el término "justificación" debe entenderse en su sentido más general. No debe entonces confundirse con las proposiciones llamadas "justificaciones" en la lógica *default*.

de NM mantenga uniforme esta noción. En el caso de la lógica de Reiter, este caso no sucede. Como lo muestra el teorema 3.2, una teoría *default* no puede perder información sólo por aumento de reglas en D ²³⁴. Esta propiedad Reiter la llamó "semimonotonicidad"²³⁵. Segundo, Podemos interpretar la función Γ de la lógica de Reiter²³⁶ no como la noción de consecuencia de la lógica de Reiter²³⁷ sino como una prescripción general para extender las teorías *default*. Así, sería una función que establecería entre otras cosas que la noción de consecuencia de cada teoría default incluye todos los esquemas válidos de la lógica clásica. Consecuentemente, desde esta perspectiva la lógica del razonamiento por *default*, de Reiter, puede verse como una lógica para el análisis de distintas nociones de consecuencia asociadas a teorías $\langle D, W \rangle$ y definidas mediante Γ y el conjunto D particular que le corresponda.

3.4.3 El Análisis Interno

3.4.3.1 La Regla *Default*

Modelar el "salto a la conclusión" no ofrece ningún problema desde un tipo de derivación deductiva, pero lo que nos interesa es realizar dicho salto justificadamente aunque esta justificación no pueda establecerse en base al tipo de derivación deductiva. Por ello será útil realizar un análisis del **mecanismo inferencial** de los defaults que nos lleva a hacer este tipo de inferencias justificadamente.

Cuando hablo de "mecanismo inferencial de los defaults" me refiero a las relaciones de interacción entre los elementos de los defaults. Por ejemplo, a las relaciones de interacción entre el prerequisite y las justificaciones. La estructura general de un *default* es la siguiente:

²³⁴Véase, *supra*, sección 3.2.

²³⁵Es interesante observar que, desde esta perspectiva, en una teoría default $\langle D, W \rangle$, el aumento de información en W no cambia la noción de consecuencia de la teoría; mientras que un aumento en D , sí lo hace.

²³⁶ Γ es la función definida mediante D1-D3 en la sección 3.2. Con ella, Reiter caracteriza formalmente una manera de construir las extensiones de una teoría default.

²³⁷Lo cual es también una interpretación posible.

$$\underline{\alpha : M_1 \dots M_n}$$

ω

En concordancia con la terminología que he estipulado, en un default: $\underline{\alpha}$ es el prerrequisito, " $M_1 \dots M_n$ " es la partícula del esquema, que expresa la salvedad y ω es la consecuencia retractable respecto de algún conjunto de reglas²³⁸. En términos simples, lo que sucede al inferir por default es lo siguiente: Si el prerrequisito está en una extensión de la teoría default, éste autoriza²³⁹ la aplicación del *default* y, si las justificaciones M_i son consistentes con la extensión entonces la excepción por inconsistencia se cumple y ω se infiere retractablemente para esa extensión. Examinaré con más detalle las relaciones entre el prerrequisito, la salvedad y la consecuencia retractable, de los *defaults*.

En la sección precedente²⁴⁰ hablé del prerrequisito como algo que autoriza la aplicación de un default. Si el prerrequisito no ha sido corroborado entonces no puede haber ni bloqueo ni aplicación del default. En este sentido, el prerrequisito hace que el default sea aplicable, lo cual no significa que sea aplicado de hecho para una cierta extensión. Esta interpretación del *default* es coherente con la forma en que Reiter introduce intuitivamente la relación entre el prerrequisito y la salvedad²⁴¹. Esta forma de ver el mecanismo inferencial del *default* resulta esclarecedora para algunas cuestiones que veremos más adelante. Sin soslayar lo anterior, estrictamente hablando el proceso de [autorización]-[aplicabilidad]-[aplicación o bloqueo] en el *default*, es una metáfora. Que se

²³⁸Las reglas son esquemas de razonamiento.

²³⁹Para la noción de autorizar véase, *supra*, sección 3.2., primera parte, "*Sentido común y No Monotonidad*".

²⁴⁰En la parte de "*La Lógica del Razonamiento por Default*".

²⁴¹Reiter muestra primero la necesidad de tomar en cuenta las excepciones en el antecedente de un condicional clásico que representa un hecho con excepciones. Luego intenta dar justificación a un salto a la conclusión. La salvedad del salto a la conclusión tiene varias interpretaciones coloquiales. Estas interpretaciones parecen suponer la ocurrencia previa del prerrequisito. Así lo indica el uso de las frases que refieren a las modalizaciones **M** de las justificaciones del *default* en su [Reiter 1980], Reiter, Raymond, "A Logic for Default Reasoning", *Artificial Intelligence*, Vol. 13, 1980, pp. 81-123. Esto es claro especialmente en las frases: "en ausencia de información al contrario" (p. 81, en G: 68), "[...] y es consistente creer [...]" (p. 83, en G: 69), y "[...] a menos que sea posible deducir otra cosa" (p. 85, en G: 70). Cuando Reiter las usa, en cada caso lo hace después de haber supuesto la ocurrencia del prerrequisito.

cumplan ambos, el prerequisite y la salvedad, es condición necesaria para la aplicación del *default*. Y no tiene sentido hablar de que uno se cumpla *antes* que el otro. Esto refleja algunos detalles interesantes.

La condición D3, en la definición de Γ , dicta que si un *default* cuenta con la autorización y el no bloqueo de la aplicación en relación con una extensión E, entonces ω debe estar en E. Que se cumpla la salvedad de un default (la excepción por inconsistencia) significa que no hay bloqueo para el default. Por ende, el cumplimiento (para una extensión E) del prerequisite y la salvedad son *suficientes* para tener una derivación de la consecuencia ω . Aunque dicho cumplimiento no es suficiente para la verdad de la proposición ω . Este es el tipo de relación que guarda el antecedente de una regla *default* (prerequisite y salvedad) con su consecuente. Ahora desglosaré esta misma relación.

La expresión de la salvedad " $\mathbf{M}_1 \dots \mathbf{M}_n$ " es leída por Reiter como: "es consistente asumir que $\mathbf{M}_1 \dots \mathbf{M}_n$ ". La expresión tiene dos componentes básicos, una secuencia $\mathbf{M}_1 \dots \mathbf{M}_n$, y el signo ":"²⁴². La secuencia está formada por fórmulas modalizadas \mathbf{M}_i que son proposiciones acerca de la posibilidad de algo. La lectura usual de estas proposiciones es "es posible que \mathbf{M}_i ". Cada miembro de la secuencia se descompone en dos: un operador de posibilidad \mathbf{M} y una fórmula \mathbf{M}_i llamada "justificación". Si la justificación es consistente con un conjunto de referencia, entonces puede aplicarse legítimamente el operador a dicha fórmula. Así, el sentido de \mathbf{M}_i es que la justificación \mathbf{M}_i es consistente con un cierto conjunto de referencia. Esto no significa que se afirme dicha justificación, sino tan sólo que se asuma. Aunque no está claro en el texto de Reiter, la asunción antedicha podría estar expresada mediante el signo ":". En el sistema de Reiter la posibilidad de una justificación está condicionada sólo a su consistencia con el conjunto de referencia, pero tal posibilidad

²⁴²El énfasis sobre este componente de la forma de salvedad lo debo a los comentarios de Raymundo Morado en su Seminario de Investigación y Tesis en Lógica y Semántica, UNAM, agosto-diciembre 1996.

otorga el derecho de una asunción y no de una afirmación. Por ello Reiter lo expresa como el hecho de que sea consistente asumir la justificación [Reiter 1980, p. 82]²⁴³.

Es importante enfatizar que el que sea consistente asumir la justificación no es equivalente a la afirmación de la justificación. Tampoco es equivalente a la afirmación de la posibilidad de la justificación. Si por "afirmación de la justificación" entendemos la pertenencia de la justificación J_i correspondiente a la extensión, entonces todas las justificaciones que contribuyeran a construir una extensión cualquiera estarían contenidas en ella. Si por "afirmación de la posibilidad de la justificación" entendemos la pertenencia de la proposición modal M_i a la extensión entonces todas las proposiciones modales que contribuyeran a construir una extensión cualquiera estarían contenidas en ella. Ninguno de los consecuentes de dichas opciones ocurre en la lógica de Reiter.

Además, el lenguaje objeto de las teorías default no contiene operadores modales. Estos sólo indican una relación de las justificaciones con el conjunto de referencia, para la activación de las reglas *default*. Por ello, la salvedad de un default debe ser entendida como una salvedad metalingüística. La posibilidad de la asunción es una propiedad que no se expresa en la extensión de una teoría sino una propiedad que podría predicarse de la extensión mediante el lenguaje de la lógica de Reiter. En consecuencia, la conjetura ω de un *default* en una extensión E depende entre otras cosas del cumplimiento de una condición cuya determinación es extra-teórica respecto de la teoría *default* Δ que da origen a E.

Esta es una de las características de la lógica default que la distinguen de las lógicas modales en [McDermott & Doyle 1980]²⁴⁴, en las que las proposiciones modalizadas son expresadas en el lenguaje objeto del sistema.

²⁴³"it is consistent to assume", Reiter 1980, *op. cit.*, p. 82. G: 68.

²⁴⁴ [McDermott & Doyle, 1980] McDermott, Drew & Doyle, Jon, "Non-monotonic logic I", *Artificial Intelligence*, vol. 13 num. 1-2, 1980, pp. 41-72.

3.4.3.2 Prerrequisito y Salvedad

Reiter escribe un ejemplo para mostrar una consecuencia de la diferencia entre su lógica y las lógicas de McDermott y Doyle. El ejemplo es la comparación de una teoría default con una teoría modal no monotónica. A pesar de que son análogos los elementos iniciales de ambas teorías, sus puntos fijos no coinciden [Reiter 1980. p. 93]²⁴⁵:

Teoría default $\Delta = \langle \{A : MB / B, C : MD / D\}, \{A \vee C\} \rangle$

única extensión: $\text{Th}(\{A \vee C\})$

Teoría no-monotónica $D^* = \{(A \& MB) \supset B, (C \& MD) \supset D, A \vee C\}$

único punto fijo: $\text{Th}(\{B \vee D\} \cup D^*)$

Dada la consistencia de B y D con D^* y con W (de Δ), se cumplen las respectivas salvedades contenidas en las reglas y los condicionales correspondientes, respectivamente, a cada teoría. El primer conyunto de cada condicional en D^* no está dado ni como dato inicial ni como resultado de ninguna derivación. Análogamente, no se dispone de los prerrequisitos de los defaults de . No se dispone de una regla tipo dilema constructivo que nos permita inferir la disyunción de los consecuentes de los default. En contraste, sí hay una regla tal²⁴⁶ para el caso de los condicionales de las lógicas no monotónicas modales. Pues en ambas teorías, Δ y D^* , se suponen todas las reglas de la lógica clásica.

Lo interesante del ejemplo que usa Reiter para señalar la diferencia entre ambos tipos de lógicas no monotónicas es que se encuentra directamente relacionado con una forma de ver el problema de la conjetura, que ha sido generalmente evadida por los lógicos de la no monotonicidad.

²⁴⁵*Ibid.*, p. 93. G: 74.

²⁴⁶El llamado "dilema constructivo".

No obstante que la diferencia entre una y otra lógica puede basarse en el uso de principios o reglas del tipo del dilema constructivo, esto no puede tomarse como la razón fundamental de la diferencia. Bastaría añadir un meta-principio análogo para los defaults de una teoría como para equiparar los resultados de sus puntos fijos. Podría pensarse que sería racional introducir un principio como éste para los defaults. Reiter no opta por esta solución y sólo enfatiza la diferencia entre una y otra lógica en base al ejemplo. Sugeriré la idea de que la negativa de Reiter a esta solución racional obedece al supuesto de una conexión más estrecha entre el prerrequisito y la salvedad en los defaults. Supuesto que, por otra parte, es anulado en el desarrollo formal de la lógica del razonamiento por *default*.

Hasta este punto del análisis conocemos, si no exhaustivamente, sí suficientemente la relación entre, por una parte, el **antecedente de un default**, es decir, el prerrequisito y la salvedad de un *default* y, por otra, su conjetura. Pensemos ahora en la relación entre el prerrequisito y la salvedad. Aparentemente ambos son presentados como independientes en la lógica del razonamiento por *default*. Pero la diferencia que Reiter enfatiza entre las lógicas modales no monotónicas y las lógicas *default* podría indicar algo contrario.

En D^* las salvedades son parte del lenguaje objeto. Esto permite realizar la derivación en combinación con la disyunción $A \vee C$. En cambio, en Δ la situación es distinta. Aún a pesar de que las salvedades se han cumplido, la disyunción $A \vee C$ puede no cumplir con el prerrequisito²⁴⁷. Mi hipótesis es que esto no es el problema real. El problema no consiste en carecer de un meta-principio o una meta-regla tipo dilema constructivo aplicable a las reglas *default* porque, aún teniéndolo, no sería *suficiente* para que también se cumplieran las salvedades correspondientes.

²⁴⁷Excepto en casos raros como que uno de los disyuntos sea tautológico o que ambos disyuntos sean el mismo.

Recordemos la metáfora [autorización]-[aplicabilidad]-[bloqueo o aplicación]. Dicha metáfora subyace a las explicaciones preliminares del texto de Reiter. Además de ello, la metáfora es coherente con una intuición lógicamente más precisa: la idea de que el conjunto respecto del cual debe tomarse el prerequisite, no es necesariamente la extensión final y casi nunca lo es. La intuición aparece más claramente en el teorema 2.1 de la lógica de Reiter [Reiter 1980, p. 89]²⁴⁸:

Teorema 2.1. *Sea $E \subseteq L$ un conjunto de fórmulas bien formadas cerradas, y sea $\Delta = \langle D, W \rangle$ una teoría default cerrada. Definamos*

$$E_0 = W$$

y para todo $i \geq 0$

$$E_{i+1} = \text{Th}_L(E_i) \cup \{ \omega \mid \alpha : \mathbf{M}\beta_1 \dots \mathbf{M}\beta_n / \omega \in D \text{ donde } \alpha \in E_i \text{ y } \neg\beta_1, \dots, \neg\beta_n \notin E_i \}$$

Entonces E es una extensión para Δ sii

$$E = \bigcup_{i=0}^{\infty} E_i$$

En este teorema la extensión se describe como la unión al infinito de una serie de conjuntos de consecuencias $E_1, E_2, \dots, E_\infty$ que guardan relaciones definidas entre sí, y tales que E_∞ es el punto fijo de la función Γ . En esta serie, para un conjunto de consecuencias E_n el prerequisite debe cumplirse respecto del conjunto E_{n-1} . Esto puede observarse en la definición del conjunto de consecuencias E_{i+1} . Tal conjunto de consecuencias está construido con todo lo que es capaz de producir el conjunto de consecuencias anterior en la secuencia, a saber, E_i , más lo que resulte de los *defaults* cuyo prerequisite esté en este

²⁴⁸"**Theorem 2.1.** Let $E \subseteq L$ be a set of closed wffs, and let $\Delta = \langle D, W \rangle$ be a closed default theory. Define

$$E_0 = W$$

and for $i \geq 0$

$$E_{i+1} = \text{Th}_L(E_i) \cup \{ \omega \mid \alpha : \mathbf{M}\beta_1 \dots \mathbf{M}\beta_n / \omega \in D \text{ where } \alpha \in E_i \text{ and } \neg\beta_1, \dots, \neg\beta_n \notin E_i \}$$

Then E is an extension for Δ iff

$$E = \bigcup_{i=0}^{\infty} E_i".$$
 [Reiter 1980], p. 89. G: 72.

mismo conjunto de consecuencias (E_i). Es interesante observar que esto no ocurre con el chequeo de las justificaciones de la salvedad. Su consistencia se checa con la extensión final, E . Así, aunque desde el punto de vista de la derivación el prerrequisito y la salvedad son, ambos, igualmente necesarios para aplicar el *default*, la descripción lógica de Reiter parece suponer la precedencia lógica del primero respecto del segundo para dicha aplicación.

Por supuesto, para una teoría *default*, la construcción de una extensión partiendo primero de la verificación de los prerrequisitos y luego checando la consistencia de las justificaciones, puede considerarse extensionalmente equivalente a una construcción que proceda primero por el chequeo de las justificaciones y después por la confirmación de los prerrequisitos. No obstante, psicológicamente hablando, en la decisión por la aplicación o no de un *default* en D de Δ , parece suponerse una precedencia lógica del prerrequisito respecto de la justificación. La precedencia lógica a la que me refiero es análoga a la precedencia lógica de los requisitos de una regla deductiva respecto de la consecuencia obtenida por dicha regla: Suponiendo que no hubiera otra forma de tener la consecuencia de una regla más que seguir la regla, la confirmación de los requisitos precede lógicamente a dicha consecuencia.

El nombre que Reiter otorga a la letra α del esquema general de un *default*, a saber "prerrequisito", delata esta carga psicológica que es lógicamente innecesaria. Coloca a la salvedad como un requisito de una regla, que depende de la confirmación de otro requisito: el *pre-requisito*.

Sin embargo, tal vez la idea de una regla que contenga **dos niveles de requisitos**, prerrequisito y salvedad, pueda ser útil para fines de modelación. La idea de la precedencia lógica del prerrequisito respecto de la salvedad sugiere la idea de que el prerrequisito autoriza la aplicación del *default*. Y en esta idea de **autorización** subyace el supuesto de una **conexión** estrecha entre el prerrequisito y la salvedad. En esta línea de

pensamiento, una vez que al *default* se le ha declarado aplicable entonces se procede a checar si su salvedad se cumple y, con ello, verificar si se aplica o se bloquea. Si esto fuera así, entonces la legitimidad del cumplimiento de una salvedad se vería condicionada por la precedente autorización del *default*.

En el caso de la diferencia entre D^* y Δ se tendría, bajo esta interpretación, lo siguiente. Habría una fuerte dependencia de cada salvedad con el cumplimiento de su correspondiente prerequisite en cada *default* en Δ . Y dado que la disyunción $A \vee C$ no afirma en especial ninguno de sus disyuntos no sería posible **legitimar** el cumplimiento de la salvedad, aún en el caso de que aquella (la disyunción) se cumpliera. Más precisamente: Si el prerequisite no ha sido dado no hay autorización para la aplicabilidad del *default*. Por tanto, no puede ser aplicado ya que ni siquiera tiene la propiedad de ser aplicable. Por el contrario, los condicionales retractables de las lógicas no monotónicas modales carecen de esta conexión de dependencia entre los antecedentes y las salvedades. La relación que establecen es, simplemente, una conjunción.

Independientemente de si Reiter mantuvo este supuesto o no, lo cual sabríamos después de una difícil y oscura búsqueda hermenéutica, he usado esta posibilidad como pretexto para encontrarnos con una nueva forma de entender el problema de la conjetura. En el caso de la estructura interna de un *default*, se trata de una nueva forma de entender la interacción entre el prerequisite y la salvedad, pero también entre la salvedad y la conjetura.

Como he apuntado antes, dicho problema inquiere por la forma de realizar conjeturas sensatas. La sensatez del razonamiento conjetural puede buscarse, en parte, en las **relaciones lógicas** que pueden exigirse entre los elementos de éste. Esta manera de abordar el problema se centra en la construcción de un esquema que modele el mecanismo inferencial de un razonamiento conjetural. Reiter ha tenido un éxito enorme en esta vía para resolver el problema de la conjetura.

Otra forma de abordar el problema es el intento por elucidar las relaciones no puramente lógicas entre los elementos de un razonamiento conjetural. Este tipo de relaciones son las que deberían guardar entre sí los elementos del razonamiento para evitar que, a pesar de cumplir con las relaciones lógicas entre ellos, el razonamiento permanezca desarticulado. Llamaré a éstas, "**relaciones de conexión**" y las entenderé como relaciones supuestas entre los sucesos referidos por las fórmulas del razonamiento o entre, por un lado, los sucesos que harían verdaderas algunas de dichas fórmulas con, por otro, los sucesos que harían verdaderas otras de ellas. En términos simples, las relaciones de conexión, como puede verse, son relaciones entre sucesos descritos por proposiciones y no entre proposiciones mismas.

Por supuesto, ambas formas de abordar el problema de la conjetura son compatibles. Pero la segunda manera de hacerlo implica un grado mayor de dificultad pues la determinación de las relaciones de conexión generalmente es dependiente del contexto. Para determinar las relaciones de conexión en el esquema general de un razonamiento conjetural habrá que tomar en cuenta el tipo de razonamiento que se intenta representar. Así, la generalidad que pueda lograrse del esquema del razonamiento conjetural será más limitada que la que resulta de la primera forma de abordar el problema. Veamos algunos ejemplos para ilustrar la diferencia entre las relaciones lógicas y las relaciones de conexión.

No hay nada en la estructura general de un *default* que nos indique una conexión entre el prerrequisito y la salvedad. Sin embargo, todos los ejemplos concretos del texto de Reiter parecen suponer una tal conexión [Reiter 1980, pp. 81-6]²⁴⁹. En todos sus ejemplos subyace la intuición de que es sensato pensar en la salvedad dado el prerrequisito, en un *default* determinado. En contraste, veamos el siguiente ejemplo:

²⁴⁹Cfr. [Reiter 1980], pp. 81-6. G: 68-70.

Ma : "Mi vecina es mexicana"

Ya : "Mi vecina es de Yucatán"

(Y)

Ma :MYa

Ya

En realidad, no hay razón para suponer que Ya a partir de Ma. Un *default* como Y es un esquema conjetural mucho menos sensato que los defaults presentados en el texto de Reiter. Pero la diferencia no estriba en su forma lógica sino en su correspondencia con nuestras expectativas acerca de las relaciones de conexión entre sucesos descritos por las fórmulas del razonamiento.

3.4.3.3 Prerrequisito y Conjetura

Pensemos ahora en la relación que guardan el prerrequisito y la conjetura. Con ello será fácil ver también una conexión entre la salvedad y la conjetura. El prerrequisito es condición necesaria para la aplicación del *default* que, por lo demás, nos lleva a la conclusión. Además de esta relación de derivabilidad, al menos en términos estrictamente lógicos, no existe relación entre el prerrequisito y la conjetura. No hay constreñimiento en el esquema general de un *default*, que nos indique que el prerrequisito deba tener alguna relación con la conclusión. El esquema (Y) también es un ejemplo para este tipo de casos: No hay conexión entre lo que dice el prerrequisito y lo que dice la conclusión al menos desde el punto de vista de nuestras expectativas. Veamos un ejemplo extremo:

Ma : "Mi vecina es mexicana"

Pp : "Pepe es un pingüino"

Db : "El mercado bursátil mundial se desestabilizará"

(P)

Ma :MPp

Db

Tanto *Y* como *P* pasarían como casos ajenos al sentido común. No obstante, para cada uno de estos *defaults* podemos imaginar contextos en los que alguien esperaría (por extraño que nos parezca) la conclusión de las premisas, a partir del prerrequisito y la salvedad. En este sentido, la lógica del razonamiento por *default* modela bien el razonamiento del sentido común aunque no restringe la estructura del mecanismo inferencial mediante relaciones de conexión. Lo modela *qua* razonamiento conjetural en términos muy generales.

El punto interesante es que, como puede suponerse, Reiter no usa este tipo de ejemplos extraños. Como he anticipado, el razonamiento que nos interesa, y el que se presupone en los sistemas de lógica no monotónica, es el razonamiento conjetural sensato. Así, en un *default* el prerrequisito debería estar en una relación más estrecha con la salvedad. El prerrequisito debería legitimar no sólo la salvedad sino también las justificaciones. Similarmente la salvedad y las justificaciones también deberían estar en conexión con la conjetura y, generalmente a través de dicha conexión, debería establecerse la conexión entre el prerrequisito y la conjetura.

Bajo la perspectiva de una conexión entre los elementos de un razonamiento conjetural, cambia la forma de entender la salvedad. En el caso de los *defaults* reinterpretados desde esta nueva perspectiva, la salvedad no sólo cumple la función de determinar la aplicación o el bloqueo del razonamiento sino también la de **legitimar**²⁵⁰ la conclusión *default*.

²⁵⁰He estado usando el término "legitimar" en esta sección en el sentido general que tiene "justificar" en epistemología: el efecto de dar apoyo racional para aceptar algo o para considerarlo como aceptable. Quiero con esto evitar confusiones con el término "justificaciones" que tiene un sentido distinto y muy preciso en la lógica *default*.

Así, aunque el prerrequisito y la salvedad guarden relaciones de conexión, si la salvedad no mantuviera ninguna conexión con la conjetura, el *default* aparecería aún desarticulado.

La idea de las relaciones de conexión en un *default* es muy importante para entender la diferencia entre los defaults de Reiter y la forma en que voy a proponer que se construya un modelo de explicación. La nueva forma de abordar el problema de la conjetura es una necesidad para quien desee modelar más interesantemente algún tipo específico de razonamiento conjetural. Sin duda los instrumentos formales construidos hasta ahora pueden servir como primer paso para una tarea como ésta. Sin embargo los lógicos de la no monotonicidad han prestado poca atención a las relaciones de conexión, y la cuestión de elucidar y poner en términos formales este tipo de relaciones, aun cuando hubiese sido sólo en forma general, no ha sido abordada tenazmente por ellos.

La nueva forma de entender el problema de la conjetura refleja también una nueva forma de entender el problema de la pertinencia general. La forma de determinación de los factores que guardan relaciones pertinentes ante una situación dada puede abordarse en un sentido operacional o bien en un sentido inquisitivo. En el primer caso lo que interesa solucionar es la eliminación de los hechos que no guardan pertinencia y estorban a la inferencia, dados ciertos supuestos generales acerca de la situación. Para ello basta con la construcción de un mecanismo inferencial que permita el salto a la conclusión, como por ejemplo el mecanismo de un default. En el segundo caso, estamos interesados no por los factores sino por la elucidación del tipo de relaciones que debería tener un hecho para ser pertinente a la situación. Nos interesan en este caso sobre todo sus conexiones con otros hechos de la situación. Un caso concreto de la forma inquisitiva de abordar el problema de la pertinencia general es el problema de las relaciones de pertinencia explicativa que mencioné en la introducción general.

Esta nueva forma de ver al problema de la pertinencia general y al problema de la conjetura tendrán que afrontarse al modelar la explicación. En esta dirección lo que interesa es elucidar las relaciones de pertinencia entre los hechos de una situación, específicamente, la situación de que el *explanandum* ocurra. Consecuentemente nos interesa reflejar estas relaciones de pertinencia a través de las conexiones internas de un argumento conjetural.

3.4.3.4 Inconsistencia entre Salvedades

Volvamos, por último, brevemente, a la legitimación que debería ofrecer la salvedad para la conjetura. Dicha legitimación está fundada, desde esta perspectiva, en relaciones de conexión entre las justificaciones y la conjetura. No obstante, el alcance de esta legitimación es todavía incierto. Puede, quizá, exigir la participación de otro tipo de factores. Por ejemplo, la de las salvedades de otros *defaults* en la misma teoría. Esto indicaría que la conexión entre prerrequisito y conjetura, establecida a través de la conexión entre salvedad y conjetura, podría no ser suficiente para una completa legitimación de la conjetura mediante el antecedente de un *default*. Veamos un ejemplo de esto.

Supongamos que el Sr. Goya es una persona sumamente adinerada. El Sr. Goya tiene un placer especial por ofrecer grandes y elegantes fiestas. En la última fiesta que ofreció, el 16 de septiembre, organizó un espectáculo de juegos pirotécnicos en el amplio jardín de su residencia. En el lapso en que duró el estruendo del espectáculo el Sr. Goya fue asesinado. Se ha contratado a un detective para que resuelva el caso y descubra al o a los asesinos. Los datos de que dispone el detective son los siguientes: De acuerdo al testamento, el sobrino del Sr. Goya es uno de los principales beneficiarios de la fortuna de este último; la guapa secretaria del Sr. Goya tenía una considerable deuda con un distribuidor de drogas; el silencioso mayordomo siempre estuvo enamorado de la secretaria; de acuerdo con la reconstrucción de los hechos, se suponen dos victimarios; la

secretaria no estuvo en la escena del crimen; el mayordomo es maniáticamente celoso; encontraron un arete de la secretaria en el departamento del sobrino del Sr. Goya; el sobrino del Sr. Goya y el mayordomo son grandes amigos; ninguno de los dos sentía verdadera simpatía por el Sr. Goya; para variar, la secretaria tampoco. De acuerdo con los datos anteriores, el razonamiento del detective pudo configurarse de acuerdo a los dos siguientes *defaults*:

(M)

Smg : "El mayordomo sentía verdadera simpatía por el Sr. Goya"

Tsb : "La secretaria tenía amoríos con el sobrino del Sr. Goya"

Am : "El mayordomo es uno de los asesinos materiales"

-Smg :M-Tsb

Am

(B)

Sbg : "El sobrino del Sr. Goya sentía verdadera simpatía por el Sr. Goya"

Tsb : "La secretaria tenía amoríos con el sobrino del Sr. Goya"

Ab : "El sobrino del Sr. Goya es uno de los asesinos materiales"

-Sbg :MTsb

Ab

M y *B* pueden sostenerse de acuerdo a los datos proporcionados. Obsérvese que la justificación en el default *M* se comprende si se tiene en cuenta la tendencia a los celos que tiene el mayordomo, y que la justificación en el default *B* se comprende considerando la deuda peligrosa de la secretaria.

Las conclusiones de *M* y *B* son consistentes entre sí. No obstante, las justificaciones para cada una de ellas no son consistentes entre sí. Dejando de lado la falta de perspicacia del detective, pues no consideró la posibilidad de que la secretaria haya usado al sobrino del Sr. Goya y teniendo planes ulteriores con el mayordomo, es irracional que el detective concluya que ambos el mayordomo y el sobrino del Sr. Goya fueron los asesinos. La razón es que ambas conclusiones tienen justificaciones contradictorias. La salvedad de *M* mantiene una conexión con su conjetura correspondiente mientras que la salvedad de *B* hace lo mismo con la suya. Ambas salvedades tienen justificaciones contradictorias y por tanto, la justificación racional a la que se aduce para la conclusión de *M* es incompatible con la de *B*.

La lógica del razonamiento por *default* permite ambas derivaciones (de *M* y de *B*) para una misma extensión. Esto indica alguna carencia en la legitimación que va desde el prerrequisito hasta la conjetura. El alcance de dicha legitimación va mucho más allá del papel de bloqueo o aplicación que tiene la salvedad. Quizá también más allá de las relaciones de conexión.

La lógica de Reiter incluye, en el papel de bloqueo o aplicación que tienen las salvedades, restricciones que hacen depender la legitimidad de una conjetura, de la relación con otros *defaults*. El hecho de que se fundamente este papel de las salvedades en el chequeo de consistencias implica que dos *defaults*, tales que la justificación de uno sea la negación de la conjetura del otro, no puedan aplicarse ambos en una misma extensión. De este modo la asunción de las justificaciones de un *default* juega un papel importante en la legitimación de las conjeturas. Pues no es legítima una conjetura que ha sido derivada mediante una justificación inconsistente con la justificación de otro *default* que también haya sido usado²⁵¹. Aunque Reiter permite este tipo de restricción en su

²⁵¹Por supuesto, no es legítima sólo en tanto su única legitimidad para aparecer en la extensión es ésa derivación. Esto es, aquí no tomo en cuenta el caso en que la conclusión ya pertenezca a *W* o pueda obtenerse mediante reglas deductivas.

lógica, no profundiza en la posibilidad de agregar una restricción para el caso de la inconsistencia entre salvedades.

3.4.3.5 Certeza

A pesar de la atenuada función legitimadora de las justificaciones respecto de las conjeturas, incluida en el sistema de Reiter, las justificaciones no son consideradas como fórmulas contenidas en la extensión. No poseen ningún grado de certeza.

En contraste, en una extensión, la conjetura de un *default* es asumida, al menos transitoriamente, con el mismo grado de certeza que las fórmulas contenidas en W. La **certeza transitoria** de las conclusiones de los *defaults* es una característica de las teorías default que corresponde, en gran medida, con nuestras intuiciones de la forma en que se procede en muchos de los razonamientos conjeturales. Observemos que la autorización de un *default* puede efectuarse sólo mediante una fórmula que ya ha sido derivada. Este tipo de fórmulas puede ser de dos clases: las contenidas en W o derivadas deductivamente, y las que han sido derivadas de otros *defaults*. No obstante que las últimas han sido derivadas mediante razonamientos conjeturales y son, en un sentido relativo, falibles, se toman transitoriamente como si fuesen verdaderas. Esta actitud ante nuestras conclusiones justificadas conjeturalmente, aunque conduce a hipótesis aún más aventuradas, permite el desarrollo de una estrategia heurística dados ciertos datos y ciertas reglas disponibles. En este sentido, se puede decir que el formalismo de Reiter es capaz de representar, mediante las extensiones, distintas **estrategias heurísticas**²⁵² para una misma situación inicial de datos y reglas. La circularidad para construir cada extensión no hace sino reflejar la naturaleza heurística del razonamiento conjetural: La construcción completa de cada estrategia heurística (es decir, de cada extensión) requiere de un ir y venir entre los resultados finales de dicha estrategia y las propias tentativas para construirla.

²⁵²Cada extensión representaría por lo menos una estrategia heurística, pero veremos después, en el capítulo IV, la conveniencia de distinguir la noción de extensión de la de estrategia heurística.

En las líneas contenidas en esta parte de la sección he hecho un análisis del mecanismo inferencial de las reglas del sistema de Reiter. El análisis nos ha llevado a una conclusión interesante: En relación con la modelación de algunos razonamientos conjeturales algunas relaciones de conexión importantes en esos razonamientos, no son capturadas por las reglas *default*. Dichas relaciones de conexión revelan una nueva forma de entender el problema de la pertinencia y el problema de la conjetura, que fueron planteados a través del estudio de la inteligencia artificial. La nueva forma de entender ambos problemas consiste en una forma de ver la legitimación de las conjeturas. Esto produce dos efectos en mi investigación: primero, clarifica la relación entre la lógica no monotónica y la modelación de la explicación y, segundo, servirá de motivación para

3.4.4 El Análisis Externo

3.4.4.1 El Contexto de un *Default*

En los dos apartados anteriores, y en diversos puntos de aplicación he intentado explorar distintos aspectos de lo que entendemos por "legitimación"²⁵³ en relación con los razonamientos conjeturales. He tratado la importancia de la legitimación de una conclusión *default* y la cuestión de la conferencia de legitimación que puede ocurrir entre los distintos elementos de un *default*. Como adelanté en el apartado anterior, el análisis interno de los *defaults* no basta para completar una perspectiva respecto de la legitimidad de una conjetura. Un caso problemático, por ejemplo, es la inconsistencia entre salvedades.

El mecanismo inferencial de un *default* particular depende fundamentalmente del contexto en que ha sido planteado dicho *default*. Una forma útil, aunque incompleta, de capturar el contexto es, usando la lógica de Reiter, interpretarlo como un conjunto de aseveraciones iniciales acerca del mundo y un conjunto de reglas particulares. Es decir, la

²⁵³De nuevo, al término "legitimación" le otorgo el sentido general del de "justificación" usado en epistemología.

teoría *default*²⁵⁴. Una teoría *default* Δ está definida por una tupla $\langle D, W \rangle$ que cumple con ciertas cualidades descritas en la sección precedente. La aplicación de un *default* en D depende del contenido de W y del resto del contenido de D . Veamos el siguiente ejemplo formal:

Supongamos la teoría cerrada $\Delta = \langle D, W \rangle$ tal que

$$D = \{r : \mathbf{M}f / \neg s; \neg q : \mathbf{M}s / \neg t\}$$

$$W = \{\neg q, r\}$$

La extensión para Δ es la siguiente:

$$E = \text{Th}(W \cup \{\neg s\})$$

Los prerequisites r y $\neg q$ hacen aplicables a ambos defaults en D . Las justificaciones f y s son consistentes con W . Sin embargo, si aplicamos el primer *default* la conclusión por default $\neg s$ hace que la justificación del segundo default sea inconsistente con la extensión. Así, esto impide que apliquemos el segundo *default*. Nótese que no es una cuestión de orden temporal en la aplicación de los *defaults*. Como he dicho antes, si intentáramos aplicar el segundo *default* antes que el primero, nos percataríamos de que eso no impide la conclusión del primero. Una conclusión default es tomada como verdadera para seguir con la estrategia heurística elegida. Ahora bien, una conclusión *default* es tomada como verdadera en tanto no se demuestre su inconsistencia con la teoría. Si la teoría se expande, la conclusión *default* puede dejar de inferirse. Esta es la forma de modelar la no monotonicidad bajo la lógica de Reiter.

Tomando en cuenta la teoría Δ anterior:

²⁵⁴Una forma de capturar el contexto, mediante la lógica del razonamiento por *default*, está directamente conectada con una forma de involucrar el contexto para clarificar las relaciones de pertinencia explicativa que no trataré en esta investigación.

Si tenemos además los conjuntos

$$D' = \{-q \wedge r : \mathbf{Mg} / \neg f\}$$

$$W' = \{\neg f\}$$

Podremos construir las teorías,

$$\Delta_D = \langle D \cup D', W \rangle$$

$$\Delta_W = \langle D, W \cup W' \rangle$$

Las extensiones resultantes serán,

Para Δ :

$$E = \text{Th}(W \cup \{\neg s\})$$

Para Δ_D :

$$E_D = \text{Th}(W \cup \{\neg f, \neg t\})$$

Para Δ_W :

$$E_W = \text{Th}((W \cup W') \cup \{\neg t\})$$

Como puede verse, en el ejemplo no se cumple que $E \subseteq E_D$ ni que $E \subseteq E_W$. Las conjeturas pueden perderse una vez que la teoría ha aumentado su información. No obstante, el contenido en W se preserva. En ninguna de las extensiones E , E_D , E_W , se perdió información acerca de W : Las aseveraciones iniciales acerca del mundo permanecen intactas. Estas tienen, en los límites de la teoría, **certeza absoluta**.

Por otro lado, observemos que $E_D = E_W$. En este caso las dos formas de incrementar la información de W llevan al mismo conjunto de consecuencias. Según lo que he discutido

antes, una noción de consecuencia, puede, en un contexto único, definirse, en relación con la propiedad de no monotonicidad, como un conjunto de reglas y no como un conjunto de consecuencias. El conjunto D de una teoría *default* Δ contribuye en este sentido a la definición de la noción de consecuencia de esa teoría. Así, de estos dos ejemplos de pérdida de información, E_D y E_W , el único que puede tomarse como un caso de no monotonicidad es en el que aumentamos la información en W y no en nuestro conjunto de reglas D . Es decir, E_W . La pérdida de la inferencia de $\neg s$, en el caso de E_D no es idéntica a la del caso de E_W . Para hablar bajo esta nueva terminología, en el primero tenemos un caso de cambio de noción de consecuencia y en el segundo un caso de cambio de información inicial.

De esta manera, la aplicación de una regla *default* depende de la forma en que está configurado su contexto entendido como la teoría que contiene a dicho *default*. Así, podemos entender el contexto como algo constituido por un conjunto de aseveraciones iniciales y una noción de consecuencia. La noción de consecuencia puede definirse como un conjunto de reglas de inferencia. En la lógica de Reiter se entiende que el conjunto D de estas reglas contiene sólo reglas *default*. Pero en la definición de una extensión para una teoría *default* está la idea de que el aparato inferencial de la lógica deductiva permanece activo constantemente en la construcción de la extensión, a través de la función de teoremicidad Th . Por ello, permaneciendo fieles a la idea principal de Reiter, puede considerarse que el conjunto de reglas que definen la noción de consecuencia de un contexto (una teoría en términos de Reiter) contiene todas las reglas deductivas de la lógica de primer orden y, además, algunas reglas *default*.

Desde esta perspectiva, la lógica de Reiter es una lógica sobre teorías que contienen reglas de inferencia *default*. En esta lógica el principal punto de análisis es la interacción entre una noción de consecuencia y una información. En este sentido, puede considerarse a la lógica del razonamiento por *default* una metalógica que versa sobre una clase especial de lógicas. Las nociones de consecuencia que son susceptibles de análisis

mediante la lógica de Reiter se encuentran **relativizadas** a conjuntos de aseveraciones tomadas como absolutamente ciertas. Dichos conjuntos interactúan con otro grupo de aseveraciones de menor grado de certeza y produce una extensión del conjunto inicial. La retractación, como he reiterado, no es permitida para el conjunto inicial (para W). Pero es importante enfatizar que la noción que se le atribuye al término **derivación** en estas teorías no está comprometida con un grado de certeza deductivo de la proposición derivada.

3.4.4.2 Inconsistencia en la Extensión

Un par de consecuencias interesantes de la forma de relativización que posee una noción de consecuencia, respecto de W , son los siguientes corolarios:

Corolario 2.2. *Una teoría default cerrada $\langle D, W \rangle$ tiene una extensión inconsistente sii W es inconsistente.*

Corolario 2.3. *Si una teoría default cerrada tiene una extensión inconsistente, entonces ésta es su única extensión. Una teoría default cerrada es consistente sii tiene una extensión consistente.*²⁵⁵

Dado el tipo de salvedad de los *defaults* y su consecuente forma de construir las extensiones para una teoría *default*, no es posible extender una teoría *default* inconsistentemente a menos que ya de antemano el conjunto W de la teoría original sea inconsistente.

El Corolario 2.2. refleja la imposibilidad de introducir nuevas inconsistencias a la teoría. Nuestras reglas no deben conducirnos a inconsistencias. Esta es lo que Hempel

²⁵⁵"**Corollary 2.2.** *A closed default theory $\langle D, W \rangle$ has an inconsistent extension iff W is inconsistent.*

Corollary 2.3. *If a closed default theory has an inconsistent extension then this is its only extension. A closed default theory is consistent iff it has a consistent extension.*" [Reiter 1980], p. 91.

llama "el requisito principal de una regla", como vimos en el capítulo I²⁵⁶. Un principio de racionalidad como éste es sumamente útil cuando se trata de conjeturar a partir de un contexto general. Lo llamaré "**Principio de Consistencia de la Noción de Consecuencia**".

Un principio como éste también es compatible con el estilo de abordar los casos de inconsistencia en los formalismos de da Costa, con los cálculos C_n , $1 \leq n < \omega$, que son finitamente trivializables. En los cálculos de da Costa es posible partir de un conjunto de fórmulas inconsistente pero no introducir, mediante los cálculos, nuevas contradicciones. Podemos así trabajar con contradicciones, pero no producir otras además de las que ya tenemos de antemano. Con ello, como sugerí en el capítulo mencionado, la paraconsistencia en la ciencia puede modelarse como una inconsistencia de transición. Es decir, un ambiente formal que bloquea la trivialización del contexto del que se parte, pero que no implica ningún compromiso epistemológico que esté a favor de la producción de contradicciones.

Por otro lado, sin embargo, el Corolario 2.3 explicita que si el contexto del que se parte es inconsistente entonces si es posible extenderlo sólo habrá un resultado posible: la trivialización. En tal caso, toda extensión de ese contexto será equivalente a cualquier otra extensión del mismo contexto.

Este otro aspecto contenido en el Corolario 2.3, de la forma de abordar el caso de inconsistencia en la lógica de Reiter, es claramente clásico. Así si partimos de un contexto inconsistente, partimos de un contexto potencialmente trivial y cualquier extensión resultante será efectivamente trivial. Sin embargo, hay casos en que las inconsistencias no juegan un papel completamente destructivo en las teorías.

La legitimación de una conjetura que proviene de una extensión trivial pierde todo valor. No así, desde una perspectiva paraconsistente, si proviene de una extensión

²⁵⁶Véase, *supra*, apartado "La Propuesta Hempel-Oppenheim". También apartado "La Ambigüedad Epistémica".

inconsistente. En la espera por la resolución de una contradicción en una teoría, la teoría puede todavía ser muy útil. Particularmente, algunas conjeturas a partir de una teoría inconsistente pueden seguir dando frutos para la teoría.

Por tanto, si queremos modelar también la inconsistencia de transición de un contexto general, será necesario evitar una consecuencia como la que implica el Corolario 2.3. de Reiter: la pérdida de legitimación de las conjeturas de una extensión inconsistente, a través de su trivialidad.

3.4.5 No monotonicidad y Transitividad

Como tuvimos oportunidad de ver, la Lógica Clásica no puede modelar no monotonicidad. Sin embargo, podría también ocurrir que la lógica no monotónica tenga algunas limitaciones respecto de la Lógica Clásica, en primera instancia, su relación de derivación podría no representar el grado de certeza que nosotros quisiéramos capturar. Debido a la disminuida certeza que puede ofrecer a su conclusión un razonamiento no monotónico algunas de las propiedades que podemos atribuir a reglas de inferencia de lógica clásica de primer orden, intuitivamente no las podríamos atribuir con una noción de derivación no monotónica. Este es el caso de la regla de transitividad. Abordaré, finalmente para esta sección, las implicaciones del Principio de Transitividad en relación con las reglas *default*.

Si la consecuencia de un *default* es la misma que el prerequisite de otro *default*, y no hay restricciones para sus respectivas justificaciones, parecería plausible construir otro *default* que tuviera como prerequisite el del primer *default* y como consecuencia la del segundo *default*. Reiter y Criscuolo, en su artículo de 1981, descubren un sentido en que esto es antintuitivo. El ejemplo es el siguiente [Reiter & Criscuolo 1981]²⁵⁷:

²⁵⁷1)"Typically high school dropouts are adults." 2)"Typically adults are employed." Por lo tanto: "Typically high school dropouts are employed", [Reiter & Criscuolo 1981] Reiter, Raymond & Criscuolo, "On Interacting Defaults", en Ann Drinan (ed.), *Proceedings of the seventh IJCAI Conference*, Vancouver, Canada, Agosto 1981, 1981, p.270.

δ_1)Típicamente los desertores de secundaria son adultos

δ_2)Típicamente los adultos tienen empleo

Por transitividad: Típicamente los desertores de secundaria tienen empleo

Al parecer aquí la transitividad debe ser bloqueada pues que los desertores de secundaria típicamente tengan empleo va contra nuestras intuiciones. El tipo de relación que generalmente es representado mediante una derivación no monotónica, no es transitiva. Esta es la razón por la que Reiter y Crisculo consideran que no debe efectuarse transitividad entre *defaults*.

Sin embargo, aún impidiendo la construcción de un nuevo *default* por esta especie de **transitividad**, el problema permanece. En una teoría *default* que contuviera como reglas a δ_1 y δ_2 , si se cumple una instancia del prerequisite de δ_1 tendríamos una instancia de la consecuencia de δ_2 . Hay, de cualquier forma, un individuo que **transita** de un *default* a otro. En estas circunstancias una de las extensiones de la teoría podría contener que un individuo es desertor de secundaria y tiene empleo. Para Reiter y Crisculo esto es inadmisibile. De modo que consideran que el problema está en la forma en que simbolizamos δ_1 y δ_2 . La simbolización usual es la siguiente:

(D)

δ_1) Sx :**MA**x / Ax

δ_2) Ax :**ME**x / Ex

Pero como esto produciría el problema, Reiter y Crisculo proponen en primera instancia:

(D')

δ_1') Sx :**MA**x / Ax

$$\delta_2') Ax : \mathbf{M}(\neg Sx \wedge Ex) / Ex$$

Pero en este caso δ_2' es un *default* no normal, de la clase llamada por ellos "seminormal". Su interés por trabajar sólo con *defaults* normales [Reiter & Criscuolo 1981, p. 271]²⁵⁸ los lleva a considerar la siguiente simbolización:

(D*)

$$\delta_3) Sx : \mathbf{M}Ax / Ax$$

$$\delta_4) Ax : \mathbf{M}\neg Sx / \neg Sx$$

$$\delta_5) Ax \wedge \neg Sx : \mathbf{M}Ex / Ex$$

La idea contenida en δ_5 es similar a la de δ_2' sólo que ahora necesitamos tener $\neg Sx$, para lo cual no basta sólo haberlo asumido. Así, la función de δ_4 es posibilitar el uso de δ_5 . Bajo las simbolizaciones D' y D* no se infiere que un individuo cumpla con S y que a la vez cumpla con E.

Dos problemas pueden desprenderse de estas propuestas de simbolización. Por una parte, con las soluciones D' y D* ¿No estamos volviendo a una nueva forma de nuestro problema original, a saber, el problema de cómo saltar a la conclusión sin tener que corroborar las negaciones de todas las excepciones al caso? y, por otra, ¿En qué sentido es indeseable tener instancias que cumplan con S y con E a la vez?

En la motivación que Reiter aduce para la construcción de la lógica del razonamiento por *default* se encuentra el problema de colocar, en el antecedente de un condicional deductivo, todas las negaciones de las excepciones para que se dé el

²⁵⁸"Ellas [las teorías con *defaults* no normales] algunas veces no tienen extensiones, carecen de semimonotonocidad, y su teoría de prueba llega a ser considerablemente más compleja que la de las teorías normales". "they sometimes fail to have an extension, they lack semi-monotonocity, and their proof theory appaers to be considerably more complex than that for normal theories". [Reiter & Criscuolo 1981], p. 271. Lo que está entre corchetes es mío.

consecuente. El mismo problema que Hempel llamó "el problema de las salvedades" en [Hempel 1981]. En el caso de D^* y D' la condición esencial en δ_5 y δ_2' es que dada una instancia de Ax tengamos en consecuencia una instancia de Ex . Añadimos la instancia de $\neg Sx$ para hacer "más racionales" ambos *defaults*. En δ_2' necesitamos tan sólo asumir esta instancia racionalizadora. En δ_5 se requiere más que eso: la derivación de la misma. Sin embargo, en ambos casos, la reconstrucción de la simbolización de δ_2 sólo está contemplando una de las excepciones posibles. ¿Qué pasa si alguien es adulto pero sufre de algún mal que le imposibilita trabajar? Bajo la filosofía subyacente en la simbolización de Reiter y Criscuolo tendríamos que añadir la negación de esta posibilidad en el prerrequisito (preferentemente) o en la salvedad del *default* original. Pero, como dijera Hempel, ¿Cómo podemos estar seguros de haber contemplado todos los casos de excepción?

Esencialmente, la idea original de la lógica de Reiter no se proponía esto. La idea de concluir, por vía de posibilidad basada en consistencia, la conclusión, produce la cualidad de que el *default* pueda saltar a la conclusión sin tener que contemplar todas las excepciones al caso. En un intento por recuperar este espíritu original, analicemos el segundo problema que he planteado. Un regreso a la perspectiva original de Reiter puede hacer desaparecer la ilusión de que el tránsito de instancias en *defaults* como los de D debería evitarse.

Recordemos que la lógica de Reiter es una lógica que estipula la forma en que puede ser extendida una teoría. La extensión de una teoría *default* es algo que depende de una cierta noción de consecuencia. Dicha noción de consecuencia se encuentra relativizada a un conjunto de reglas *default* y lo que podamos hacer con esta noción depende del conjunto de supuestos W del que partamos. Supongamos el siguiente caso de teoría $\langle D, W \rangle$:

$$W = \{Sa\}$$

$$D = \{\delta_1, \delta_2\}$$

Para esta teoría tendríamos la extensión E tal que:

$$\{Sa \wedge Ea\} \subseteq E$$

¿Tiene sentido, *dentro del contexto* de la teoría $\langle D, W \rangle$, decir que la conjunción $Sa \wedge Ea$ es antintuitiva? En realidad no. Si un agente racional contuviera sólo la teoría $\langle D, W \rangle$ no le parecería nada extraña la conclusión de que $Sa \wedge Ea$. No tiene los elementos para juzgar la extrañeza de dicha conclusión. Lo contraintuitivo de la conjunción en cuestión depende directamente del contexto desde el que se mira. Pero del contexto entendido como un conjunto de elementos pragmáticos extra-teóricos. Como Reiter y Criscuolo mencionan, nosotros no quisiéramos concluir tal conjetura [Reiter & Criscuolo 1981, p. 270], pero en realidad se olvidan de que nuestro conjunto de creencias es mucho más rico que $\langle D, W \rangle$.

Mi punto entonces es que no necesitamos evitar que una instancia de Sx sea también una instancia de Ex . El tránsito de una instancia en ambos *defaults* de D no es necesariamente dañino.

Sin soslayar lo anterior, hay casos en los que sí sería indeseable una conjunción similar a la que me he referido. Reiter y Criscuolo ponen un ejemplo que viene al caso:

δ_6)Típicamente los estudiantes universitarios son adultos.

δ_7)Típicamente los adultos son empleados.

δ_8)Típicamente los estudiantes universitarios no son empleados.²⁵⁹

²⁵⁹"Typically university students are adults.", "Typically adults are employed.", "Typically university students are not employed." [Reiter & Criscuolo 1981], p.272.

La aseveración δ_8 es una condición que haría antintuitivo un tránsito de instancias de δ_6 a δ_7 . Lo antintuitivo de una consecuencia *default* en este tipo de casos es siempre en relación con un contexto determinado. δ_8 es una determinación explícita que puede contradecir el resultado del tránsito de instancias de δ_6 a δ_7 . Esto es lo que lo hace diferente respecto del ejemplo anterior. En este caso no hay ambigüedad de lo que determina el contexto: lo antintuitivo se explicita claramente en la teoría.

Ahora bien, una simbolización usual de las tres aseveraciones quedaría:

(D1)

$$\delta_6) Ux : \mathbf{MAx} / Ax$$

$$\delta_7) Ax : \mathbf{MEx} / Ex$$

$$\delta_8) Ux : \mathbf{M-Ex} / -Ex$$

Y una simbolización de acuerdo a D* sería:

(D1*)

$$\delta_6^*) Ux : \mathbf{MAx} / Ax$$

$$\delta_{7a}^*) Ax : \mathbf{M-Ux} / -Ux$$

$$\delta_{7b}^*) Ax \wedge -Ux : \mathbf{MEx} / Ex$$

$$\delta_8^*) Ux : \mathbf{M-Ex} / -Ex$$

El problema con D1 es que para autorizar δ_8 es necesario tener una instancia de Ux pero este mismo hecho produciría el bloqueo de δ_7 . De manera que, para $W = \{Ua\}$, δ_8 se aplica para alguna extensión y no para otra. En una obtendríamos $(Aa \wedge Ua) \wedge -Ea$ y en

otra $(Aa \wedge Ua) \wedge Ea$. Lo preocupante según Reiter y Criscuolo²⁶⁰, es que toda fórmula contenida en una extensión de una teoría default debe ser considerada como una derivación aceptable de la teoría, y la fórmula $(Aa \wedge Ua) \wedge Ea$ no es aceptable. En D1*, en cambio, para $W = \{Ua\}$ tenemos como una consecuencia sólo que $(Aa \wedge Ua) \wedge \neg Ea$; lo cual es, según ellos, más intuitivo.²⁶¹

Sin duda, D1* cumple con su cometido. Pero a costa de suponer que uno puede reescribir el *default* cada vez que se percata de una excepción.

En realidad el problema no es evitar el tránsito de δ_6 a δ_7 sino poder elegir a una fórmula sobre otra. El problema que tenemos ante nuestros ojos no es un problema de transitividad sino un problema de **evaluación de extensiones**. Exploraré la forma de solucionar el problema de adecuación sin recurrir a la adición de nuevas excepciones.

Desde un punto de vista, no es intuitivo que a partir de $W = \{Ua\}$ D1 permita obtener $(Aa \wedge Ua) \wedge Ea$ y $(Aa \wedge Ua) \wedge \neg Ea$. Pero, desde otro ángulo, tampoco es intuitivo que D1* sólo permita obtener $(Aa \wedge Ua) \wedge \neg Ea$. Originalmente la consecuencia de ser una instancia de Ex depende indirectamente de ser una instancia de Ux. Pero esto sólo hace que ante $W = \{Ua\}$ la instancia de Ex sea más débil que la instancia de $\neg Ex$, dado que la instancia de $\neg Ex$ se infiere directamente de la instancia de Ux. Así, la consecuencia Ea es perfectamente legítima, pues con esta teoría no sabemos a ciencia cierta si un universitario es empleado o no. La disyunción de posibilidades está bien representada por el par de extensiones ortogonales que resultan de D1. No obstante, es más racional optar por una de las extensiones que por la otra. Es más racional optar por la que contiene la consecuencia $\neg Ea$ que por la que no la tiene. Podemos ver el problema como la cuestión

²⁶⁰*Idem.*

²⁶¹"Nótese que esta representación predice que cualquier individuo que sea simultáneamente una instancia de A [U] y B [A] será una instancia de no-C [no-E], en vez de una instancia de C". "Notice that this representation predicts that any individual which is simultaneously an instance of A and B will be an instance of not C, rather than an instance of C." *Idem.* Lo que se encuentra entre corchetes es mío.

de poder obtener $(Aa \wedge Ua) \wedge Ea$ y $(Aa \wedge Ua) \wedge \neg Ea$ y después ver cuál es la más plausible de ambas fórmulas en concordancia con nuestra teoría.

Reiteraré el argumento iniciando con un caso más simple. Un caso sencillo de ortogonalidad es el siguiente par de *defaults* y $W = \{Ua\}$:

δ_8) $Ux : \mathbf{M}\neg Ex / \neg Ex$

δ_9) $Ux : \mathbf{M}Ex / Ex$

Observemos que, en este contexto, ante sólo la información de que a es un caso de Ux , no tenemos derecho a decidir si a es un caso de Ex o no, en absoluto. Pero si Ea es inferido no directamente por δ_9 sino indirectamente a través de δ_6 y δ_7 , la predilección por $\neg Ea$ se hace plausible. Sin embargo, una derivación indirecta es una derivación. En lógica clásica las derivaciones indirectas tienen el mismo estatus que las derivaciones directas. Es claro que en lógica no monotónica esto no puede ser así: Intuitivamente, una fórmula obtenida mediante derivaciones indirectas pierde cierto grado de plausibilidad²⁶². A pesar de ello, la **pérdida de plausibilidad** de una fórmula en lógica del razonamiento por default no es equivalente a la pérdida de su calidad de consecuencia. Toda fórmula que pertenezca a alguna de las dos extensiones es una consecuencia legítima de una teoría y, consecuentemente, cualquiera de las dos extensiones es legítima. Pero esto no obsta para que haya mejores extensiones que otras. En palabras nuevas: puede haber varias estrategias heurísticas legítimas para extender una teoría, pero no todas tienen por qué ser igualmente buenas.

De este modo, el problema puede verse como el problema de calificar el **estatus de las extensiones**. Si lográramos identificar el tipo de derivaciones a las que se recurrió para obtener una fórmula, podríamos tener algún criterio para elegir entre las

²⁶²Esto mismo sucede en el caso de los razonamientos probabilísticos. Cuando la probabilidad de estos argumentos es menor que 1 y mayor que 0, la probabilidad de las consecuencias de argumentos concatenados se decrementa.

extensiones. El criterio podría basarse entonces en el tipo de prueba que se usó para la construcción de las extensiones. Por ejemplo, entre dos extensiones, E_a y E_b , tales que E_a contiene consecuencias obtenidas mediante derivaciones que *encadenan*²⁶³ dos o más *defaults* y E_b no contiene consecuencias derivadas de este modo, entonces, si se supone que todos estos *defaults* tienen el mismo grado de plausibilidad, E_b es una extensión más fuerte que la extensión E_a .

Con esto sugiero una forma de evaluar el estatus entre varias extensiones de una misma teoría *default*. Esta forma de evaluar está relacionada con el tipo de reglas usadas en la derivación. El **encadenamiento** de *defaults* al que me referí en el párrafo anterior no es otra cosa que la posibilidad, para una instancia a , de transitar entre estos *defaults*. El tránsito de instancias en reglas deductivas no tiene el mismo efecto que el tránsito en reglas *default* que no son casos límite²⁶⁴. La atribución de la nueva propiedad (E , por ejemplo) por medio de este tránsito tendría en estos casos menos certeza que la que tendría si se hubiera usado sólo un *default* para obtenerla; y mucho menos certeza aún, que la que tendría si sólo se hubieran usado reglas deductivas para obtenerla.

Así, por una parte he dicho que las conclusiones obtenidas mediante *defaults* son tomadas como si tuviesen una certeza transitoria, para seguir extendiendo la teoría. Una certeza similar a la conferida para las conclusiones que vienen solamente de reglas deductivas. Por otra parte, dije recientemente que la certeza de las consecuencias varía en relación al tipo de reglas que se usaron para obtenerlas. Podríamos entonces tener dos clases de certeza en una lógica no monotónica: la certeza transitoria, y la **certeza real**.

Por último, hablar de certeza transitoria y certeza real tanto como hablar de extensiones más fuertes que otras, nos coloca frente a la tentación de construir un criterio de preferencia entre extensiones y excluir las inferiores. No obstante, debemos resistirnos

²⁶³Como el caso de δ_6 y δ_7 .

²⁶⁴Consideraré que una regla *default* es un caso límite cuando su conjetura se sigue deductivamente del prerrequisito.

a esta tentación si queremos modelar la explicación bajo las ideas que he defendido. Excluir las extensiones inferiores sería abrazar directamente el ideal de buena explicación del que hablé en el capítulo I. Poder hacer una evaluación entre las extensiones de una teoría *default* tendría el objetivo, desde mi punto de vista, de intentar precisar el tipo de justificación que se puede aducir racionalmente a una estrategia heurística en relación con la teoría desde la que se ha partido. En este sentido, tendría el objetivo de precisar el peso que tienen entre sí las distintas posibilidades de que se dispone. Sin embargo, la preferencia por una extensión en particular, podría obedecer a otro tipo de factores que yo no consideraré en mi investigación. De cualquier forma, por los límites de esta investigación, la interpretación de los casos de transitividad entre *defaults* no serán incluidos en el modelo que presentaré después.

Hasta aquí, en general, he intentado en esta sección clarificar algunos elementos de la no monotonicidad, en el contexto de la Lógica del razonamiento por *default*. Las reflexiones que hice servirán de punto de partida para ofrecer algunas soluciones concretas en la modelación formal de la explicación. Por ello en esta sección he querido también aproximar la idea de no monotonicidad a la idea de explicación. Abordaré más estrechamente el vínculo entre la no monotonicidad y la explicación en el siguiente capítulo.

CAPÍTULO 4.

UN MODELO ARGUMENTAL DE EXPLICACIÓN CIENTÍFICA

4.1 Un modelo paraconsistente y no monotónico de explicación.

4.1.1 El contexto de esta propuesta formal.

La explicación es un fenómeno estudiado por al menos dos diferentes e interesantes disciplinas de conocimiento: Inteligencia Artificial y Filosofía de la Ciencia. Ambas perspectivas pueden ser vistas como alternativas complementarias en razón de clarificar esta función básica de la Ciencia. En el contexto de la Inteligencia Artificial, las explicaciones han sido vistas principalmente como un proceso para inferir hipótesis. Desde la Filosofía de la Ciencia, las explicaciones se asumen siendo una clase de producto de la actividad científica²⁶⁵. En esta tradición de investigación, la idea central es construir una teoría o un modelo formal que define adecuadamente las explicaciones en la Ciencia. Cuando la elección teórica resulta en la construcción de un modelo formal, la propuesta clásica es ver las explicaciones como estructuras argumentativas. La discusión central es la siguiente: ¿Qué clase de razonamiento puede clarificar a la noción de explicación?

El modelo que se construya se espera que sea un criterio para identificar explicaciones, un criterio de reconocimiento de explicaciones desde otros elementos en la Ciencia. En el núcleo, esta línea de investigación en Filosofía de la Ciencia asume que así uno podría contribuir a bosquejar una idea clara de la noción de racionalidad. El debate en Filosofía se

²⁶⁵ Los textos clásicos son [Hempel & Oppenheim 1948] Hempel, Carl, G. & Oppenheim, Paul, "Studies in the Logic of Explanation", *Philosophy of Science*, vol. 15, pp. 567-579, 1948. Y [Hempel 1965] Hempel, Carl, G., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press y Collier Macmillan, N.Y. y Londres 1965.

inició por Hempel y Oppenheim en un artículo muy famoso de 1948. Ellos propusieron lo que podríamos llamar los “modelos clásicos”: básicamente un par de esquemas argumentales tales que, si se les añaden algunas otras condiciones relevantes, se ven como una abstracción de la explicación científica. La base de estos esquemas es un argumento que contiene leyes y condiciones iniciales como premisas (*explanans*), y una descripción de un fenómeno como conclusión (*explanandum*). Esta clase de argumento, con algunas adiciones de condiciones de no vacuidad y algunas restricciones empíricas, es el núcleo de los modelos clásicos. Por ello, como hemos visto, un modelo de explicación se presenta como una estructura argumentativa:

Explanans, por lo tanto, explanandum

Aunque, como he dejado claro en el capítulo 2, la diferencia no se reduce a eso, una diferencia importante que se suele aducir para los dos modelos clásicos, son que uno, el D-N, usa leyes deductivas, y el otro, I-S, leyes estadísticas.

Como hemos visto también, hubo un duro debate acerca de la posibilidad de modelación de la explicación, de la década de los 60's a la de los 90's. Por un lado, y paralelamente a esta discusión filosófica, algunos filósofos de la explicación construyeron sus propios modelos de explicación alternativos. Algunos de estos modelos fueron modelos de razonamiento, pero algunos otros fueron modelos formales sin la asunción de que la explicación pueda ser modelada por un razonamiento²⁶⁶.

A través del análisis que he hecho a lo largo de la investigación, desde mi punto de vista, los problemas acerca de la explicación, en Filosofía de la Ciencia, pueden ser clasificados en tres tipos:

- I) Problemas contra las virtudes de la inferencia en el modelo²⁶⁷.

²⁶⁶ Por ejemplo, [van Fraassen 1980] van Fraassen, Bas, *The Scientific Image*, Oxford University Press, 1980.

²⁶⁷ Estoy pensando en los contraejemplos acerca de las características suficientes o de las necesarias del modelo, o acerca del llamado “problema de la ambigüedad epistémica” (Cf. Hempel, 1965), o también sobre

- II) Problemas acerca de la relación de pertinencia²⁶⁸.
- III) Cuestiones acerca del contexto de la explicación²⁶⁹.

En esta sección presentaré un modelo formal de explicación, que ayude a resolver estos problemas mediante una representación de los siguientes 4 items:

- A) Explicación como un *producto* de la actividad científica, que puede verse como una clase de razonamiento.
- B) Interacción entre explicación y una parte de su contexto teórico asociado.
- C) Cambio explicativo.
- D) Explicación en contextos teóricos inconsistentes.

Mi análisis del debate filosófico acerca de la explicación muestra que, si abandonamos la pretensión de capturar formalmente la mejor explicación en cada caso en cuestión y, en cambio, nos enfocamos en los aspectos mínimos de una explicación, entonces los principales problemas desaparecen.

La estrategia es tomar distancia de algunos ideales filofóficos acerca de las explicaciones. Como hemos dejado claro, esto va más allá de una mera desarticulación de los problemas, nos ayuda a representar la explicación en una forma más dinámica. Un costo filosófico de esta estrategia nos ha costado, entonces, abandonar el ideal de buena explicación para conseguir plantear a las explicaciones como propuestas epistémicas (de representación) que son elaboradas para elucidar el mundo. Son propuestas epistémicas que pueden pensarse en el marco de una posición no deductivista y centrada en la representación pragmática de esta función de la Ciencia. Son construcciones de la Ciencia que son falibles, flexibles y retractables, inventadas para la comprensión de la realidad. He defendido que

la también difícil representación de deductiva en Ciencia, propuesta independientemente en [Hempel 1988] y [Orayen 1976].

²⁶⁸Como los problemas alrededor de la indeterminación causal o alrededor de las asimetrías de la explicación. Para ambos tópicos puede verse [Achinstein 1985] Achinstein, P., *The Nature of Explanation*, Oxford University Press, 1985.

²⁶⁹ Para la noción de causalidad en diferentes contextos, véase [van Fraassen 1980].

este enfoque nos dejará abierta la posibilidad de representar una conexión más flexible y modular entre explicaciones, teorías científicas y lógicas subyacentes a esas teorías. Al mismo tiempo, nos ayudará a resolver problemas centrales en el debate filosófico.

Así, para mostrar la interacción entre estos aspectos de la explicación científica, y mostrar a su vez que es posible tal representación, una parte central de mi propuesta es la construcción de un modelo no monotónico y paraconsistente que pudiera capturar todo esto y que sea, además, una manera de representación argumental de la explicación. El marco formal del modelo es la combinación de la lógica del razonamiento por *default* de Raymond Reiter²⁷⁰ y la jerarquía C_n de cálculos paraconsistentes de Newton da Costa²⁷¹. La construcción del marco general está basada en la noción de sistema epistémico y un tratamiento modular de la reglas en un sistema dado²⁷².

4.1.2 Algunas asunciones filosóficas acerca de la explicación, que se hacen en el modelo.

Una razón profunda para dar el giro hacia una perspectiva epistémica y pragmática de la explicación es el grado de complejidad involucrado en este concepto. Trataré de resumir brevemente ciertas asunciones subyacentes al modelo que presentaré. Las asunciones muestran justamente el grado de complejidad que ha alcanzado en el ámbito de la Filosofía de la Ciencia. Las asunciones son también el producto de mi reflexión a lo largo de esta investigación, y serán importantes para entender cabalmente la aplicabilidad teórica del modelo.

²⁷⁰ [Reiter 1980] Reiter, Raymond, "A logic for Default Reasoning", in *Artificial Intelligence*, vol. 13, 1980, pp. 81-123.

²⁷¹ [Da Costa 1974] Da Costa, N.C.A., "On the theory of Inconsistent Formal Systems", en *Notre Dame Journal of Formal Logic*, Vol. 15, Num. 4, 1974. Para un panorama muy completo de la paraconsistencia véase también [D'Ottaviano 1990] D'Ottaviano, Ítala M., "On the Development of Paraconsistent Logic and Da Costas's Work", en *The Journal of Non-Classical Logic*, Vol. 7, Núm. ½, May/Nov 1990, pp. 89-152.

²⁷² Modularidad en el sentido de que estos ítems pueden ser combinados en diferentes composiciones particulares.

- 1) Por “relación de explicación” entenderé una propuesta para relacionar dos cosas de muy diferente tipo: un fenómeno de la realidad y un conjunto de proposiciones acerca de él.
- 2) Si sustituimos el fenómeno por una descripción de él, tenemos una explicación. Las explicaciones pueden ser representadas por un metaargumento complejo con restricciones de diferentes clases. Trataremos de usar este argumento para construir un criterio mínimo para reconocer explicaciones. Lo que cumpla las condiciones de este metaargumento no necesariamente será considerada una buena explicación pero sí una explicación después de todo.
- 3) Este argumento es falible en el sentido de que algunas de sus proposiciones pueden ser falsas y es retractable a partir de que la relación de inferencia es no monotónica. La retractabilidad de la inferencia debería hacer plausible la idea de que una relación correspondiente de explicación, asumida entre el conjunto de proposiciones y la realidad, no se cumple.
- 4) Como una propuesta epistémica, la relación de explicación puede cumplirse respecto de un contexto teórico dado pero respecto de otro, quizá no. Uno y el mismo conjunto de proposiciones miembros del explanans pueden tener diferente comportamiento respecto de diferentes teorías. Este importante hecho nos hace distinguir entre *explanans* directos y *explanans* contextual, es una diferencia entre, por un lado, las proposiciones que constituyen directamente el argumento explicación y, por otro, el contexto asumido para inferir el *explanandum*.
- 5) Representar el cambio explicativo implica la capacidad de capturar la interacción entre *explanantes*²⁷³ directos y *explanantes* contextuales. Así, necesitaremos cierto nivel de representación de contextos. La Inconsistencia que tienen ciertos tesis en el cambio explicativo necesitará también poder ser representada. Pero, lo que es aún más interesante, ciertos casos históricos requieren también que se pueda representar un contexto o una teoría que que son inconsistentes, aunque esto fuera transitoriamente.

²⁷³ Plural of “*explanans*”.

- 6) Finalmente, deberíamos distinguir dos cosas, asumidas pero generalmente confundidas en el debate filosófico sobre la explicación: lo que he llamado “relaciones de explicación” y lo que llamaré “relaciones de pertinencia explicativa”. La confusión ocurre principalmente en el tratamiento del problema de las asimetrías en la explicación. Recordemos, por “relación de explicación” entenderé una relación entre un *explanans* y su fenómeno *explanandum*. Por “relaciones de pertinencia explicativa” entenderé una relación entre los sucesos que harían verdadero el *explanans* y el suceso descrito por el *explanandum* (o como a veces lo he dicho “el fenómeno *explanandum*”).
- 7) El modelo debería ser suficientemente flexible para articular muy diferentes ítems: *explananda*²⁷⁴, fenómenos *explanandum*, *explanantes* directos, *explanantes* contextuales, ítems reales a los que se refieren estos *explanantes*, relaciones de explicación y relaciones de relevancia explicativa.

4.1.3 Un Marco Teórico Formal para el Análisis de las Explicaciones

Llamemos a este ambiente formal “Marco Formal para el Análisis de Explicaciones” o simplemente “Ex”. El vocabulario es el usual de la lógica de primer orden con identidad y su metalógica. La idea central es que este marco permita comparar unas explicaciones con otras en el marco de teorías y sus lógicas subyacentes. Para ello representaremos las teorías como tuplas de reglas y asunciones acerca de la realidad, y representaremos de modo similar a las lógicas. El comportamiento de las teorías es similar al comportamiento de una teoría *default*. La idea central, entonces será usar la lógica *default*, cambiando alternativamente la lógica subyacente asociada a una noción de teoremicidad particular. En lo que sigue esbozaré el marco formal. Asumiré algún conocimiento básico de dos sistemas lógicos: la lógica *default* y la jerarquía de cálculos paraconsistentes, como fueron presentados respectivamente en [Reiter 1980] y en [da Costa 1974].

²⁷⁴ Plural de “*explanandum*”.

Supondremos la Teoría de Conjuntos Zermelo-Fraenkel (ZF) y la presentación estará situada en el nivel general de la metateoría. La siguiente caracterización de las reglas y los sistemas epistémicos, aún a pesar de ser diferentes de los que se presentan en [Marek & Trusczyński 1993], está inspirada en ese trabajo, y también en la idea de lógica subyacente de [da Costa 1974].

4.1.3.1 Sistemas Epistémicos

Empezaremos suponiendo, como una base para el Marco Formal Ex, los siguientes conjuntos de vocabulario: constantes lógicas $V_1 = \{\neg, \wedge, \vee, \supset, \leftrightarrow, \forall, \exists\}$, constantes para individuos $V_2 = \{a, b, \dots, t, a_1, \dots, t_1, a_2, \dots\}$, variables para individuos $V_3 = \{u, v, w, x, y, z, u_1, \dots, z_1, u_2, \dots\}$, constantes predicativas $V_4 = \{P, Q, R, \dots, Z, P_1, \dots, Z_1, P_2, \dots\}$, y variables proposicionales $V_5 = \{\alpha, \beta, \gamma, \dots, \zeta, \alpha_1, \dots, \zeta_1, \alpha_2, \dots\}$.

Tomando en cuenta que nuestro lenguaje será $\cup_{i=1}^5 V_i$, supondremos el conjunto de fórmulas bien formadas (*fbfs*), *FOR*. Construimos las *fbfs*, como se hace usualmente.

Definición 1 Sean $A, B, K_1, K_2, K_3 \subseteq FOR$ Y $\gamma \in FOR$. Dadas las tuplas $((A, \gamma), B)$, y $((K_1, K_2), K_3)$, decimos que $((A, \gamma), B)$ es **una regla r sobre FOR** sii se cumple que:

Si $A \subseteq K_1$ Y No para toda $\alpha \in FOR \cup K_2 \vdash \alpha$, Entonces $\gamma \in K_3$.

Llamaremos a cada $\alpha \in A$ “un **prerrequisito** en *r*”, γ se llamará “la **conclusión** de *r*”, A “la **condición de presencia** de *r*”, cada $\beta \in B$ “una **justificación** en *r*”, B “la **condición de ausencia** de *r*”, finalmente, $A \cup B$ se llamará “el **requisito** de *r*”.

Definición 2. Una regla **estándar (sr)** sobre *FOR* es una regla $((A, \gamma), B)$ sobre *FOR* tal que $B = \emptyset$. En este caso escribiremos (A, γ) en vez de $((A, \gamma), B)$. Por el contrario, una **regla no estándar** sobre *FOR*, *nsr*, es una regla $((A, \gamma), B)$ sobre *FOR* tal que $B \neq \emptyset$.

Definición 6. Un **sistema epistémico (eps)** Σ es una tupla $\langle \Psi, \Omega \rangle$, tal que $\Psi \neq \emptyset$ es un conjunto de reglas y donde $\Omega \subseteq FOR$. Los elementos de Ω se pueden llamar **asunciones**. Las *fbfs* pueden clasificarse en dos conjuntos: el primero de ellos contiene todas las *fbfs* que no tienen miembros de V_4 , el segundo de ellos, contiene las fórmulas que incluyen por lo menos un miembro de V_4 . Llamaremos **asunciones abstractas** a las asunciones en el primer conjunto, y **asunciones concretas**, a las del segundo.

A. Un sistema epistémico (eps) $\langle \Psi, \Omega \rangle$ es una **lógica Λ** (“lambda”) sobre *FOR* sii:

- a) Cada $\alpha \in \Omega$ es una asunción abstracta.
- b) Cada θ que ocurre en $((A, \gamma), B) \in \Psi$, es una asunción abstracta.

B. Un sistema epistémico $\langle \Psi, \Omega \rangle$ es una **teoría T** (letra griega “Tau”) sobre *FOR* sii:

- a) No toda $\alpha \in \Omega$ es una asunción abstracta, y $\Omega \neq \emptyset$.
- b) No toda θ que ocurra en $((A, \gamma), B) \in \Psi$, es una asunción abstracta.

Podemos, artificialmente, construir un **sistema epistémico complejo (ceps)**, como sigue:

Definición 7. Un **sistema epistémico complejo (ceps)** es una tupla $\langle \Sigma_1, \Sigma_2 \rangle$, tal que Σ_1 es una lógica y Σ_2 es una teoría. Denotaremos también un **ceps** como una tuple $\langle \Lambda, T \rangle$ tal que Λ es una lógica y T es una teoría. Es posible, *prima facie*, que no exista ninguna relación entre Λ y T. Llamaremos a Λ la lógica subyacente a la teoría T. Esto es una generalización de la idea básica de lógica subyacente de Newton da Costa.

Definición 8 Una **derivación** de una *fbf* ϕ **de un conjunto de fórmulas Z en un ceps** $\langle \langle R, S \rangle, \langle P, W \rangle \rangle$, es una secuencia de *fbfs* $\langle \gamma_1, \dots, \gamma_n \rangle$ tal que:

1) $\gamma_n = \phi$.

2) Para cada γ_i en la secuencia, se cumple por lo menos una de las siguientes condiciones:

a) $\gamma_i \in W \cup S \cup Z$.

b) γ_i es la conclusión de alguna regla $((A, \gamma_i), B) \in R \cup P$ tal que $((A, \gamma_i), B)$ es aplicable en γ_{i-1} .

Decimos que $((A, \gamma_i), B)$ es aplicable en γ_{i-1} *sii* las condiciones para que $((A, \gamma_i), B)$ dé como resultado γ_i , requieren miembros γ_k de la secuencia, tal que $k < i$ en la secuencia.

Si $\Sigma = \langle \Lambda, T \rangle = \langle \langle R, S \rangle, \langle P, W \rangle \rangle$, decimos que ϕ es una **$\langle \Lambda, T \rangle$ -consecuencia** o una Σ -consequence a partir del conjunto Z , o una **Σ -derivation** a partir de Z , *sii* hay por lo menos una $\langle \Lambda, T \rangle$ -derivación de ϕ a partir de Z , esto es, una Σ -derivación de ϕ a partir de Z . Escribiremos en símbolos que ϕ es una **Σ -consecuencia** a partir de Z , como $Z \vdash_{\Sigma} \phi$.

Así, definiremos la noción de consecuencia sintáctica relativa a diferentes lógicas Λ para diferentes teorías T . Decimos que $Cn^{\Sigma}(Z)$ es el conjunto de Σ -consecuencias a partir de Z . Así, la estrategia central para comparar explicaciones en contextos teóricos es clara: podemos tomar T como una teoría default que represente una teoría científica y tomar Λ como cualquier lógica de la jerarquía de cálculos de predicados paraconsistentes de primer orden con identidad C^{*}_n , $1 \leq n < \omega$, de da Costa. Consecuentemente, una vez que establezcamos estos cálculos estaremos estableciendo que podemos usar cualquiera de ellos. No obstante, para efectos prácticos, en las aplicaciones del modelo, supondremos en muchos casos C^{*}_1 , aunque sabemos que los resultados pueden generalizarse también para los otros cálculos de la jerarquía C^{*}_n , $1 \leq n < \omega$.

Definición 9. El conjunto de consecuencias, denotado por $Cn^{\Sigma}(Z)$, es tal que

$$Cn^{\Sigma}(Z) = \{\phi \mid Z \vdash_{\Sigma} \phi\}.$$

Nótese que las definiciones de sistemas epistémicos pueden ser dadas similarmente para un sistema T y para un sistema Λ .

Definición 10 Dada una *nsr* $((A, \gamma), B)$ en un ceps Σ , y dados $K, Z \subseteq FOR$, decimos:

- A) K es el **contexto de derivación respecto de A** (condición de presencia) en $\Sigma+Z$ (es decir en Σ extendido por la unión de Z y sus conjuntos de asunciones) *sii* para que $((A, \gamma), B)$ pueda cumplirse es necesario confirmar que $A \subseteq K$ y que $K \subseteq Cn^\Sigma(Z)$.
- B) K es el **contexto de asunción respecto de A** en $\Sigma+Z$ *sii* para que $((A, \gamma), B)$ pueda cumplirse es necesario confirmar que $A \subseteq K$ y que $K = K - Cn^\Sigma(Z)$.
- C) K es el **contexto de derivación respecto de B** (condición de ausencia) en $\Sigma+Z$ *sii* para que $((A, \gamma), B)$ pueda cumplirse es necesario confirmar que $Cn^\Sigma(B \cup K) \neq FOR$ y que $K \subseteq Cn^\Sigma(Z)$.
- D) K es el **contexto de asunción respecto de B** en $\Sigma+Z$ *iff* para que $((A, \gamma), B)$ pueda cumplirse es necesario confirmar que $Cn^\Sigma(B \cup K) \neq FOR$ y que $K = K - Cn^\Sigma(Z)$.

Esto nos permitirá construir diferentes características de las reglas, que pueden presentarse gradualmente en relación con el tipo de requisitos que ellas exigen. Podemos distinguir, por ejemplo, reglas válidas y reglas inválidas y, dentro del conjunto de reglas inválidas, podemos distinguir reglas retractables de reglas no retractables. De este modo, podemos caracterizar diferentes clases de interacción entre lógicas y teorías dependiendo del comportamiento de la noción de consecuencia pensada esta última como una relación gobernada por cierta clase de reglas. Es mediante esto que es posible reconstruir las teorías default como las de Reiter, pero podremos hacerlo en una forma más general que nos permita una combinación exitosa con un cálculo que admita inconsistencias. Veámos esto con detalle.

4.1.3.2 Construyendo una Teoría Default

Primero describiremos la jerarquía cuantificacional paraconsistente de da Costa, $C^{*=n}, 1 \leq n < \omega$, con identidad mediante la estructura axiomática de los sistemas de da Costa en la jerarquía $C_n, 1 \leq n < \omega$ [da Costa 1974]²⁷⁵, más los axiomas y reglas del cálculo de predicados de primer orden de la Lógica Clásica que están relacionados con los cuantificadores, y, además, los tres axiomas Q1-Q3 escritos más abajo, y los dos axiomas usuales de la identidad (=). Observemos que el lenguaje de los cálculos $C^{*=n}, 1 \leq n < \omega$, es el lenguaje de los cálculos C_n acrecentado con el símbolo del predicado de identidad, y los cuantificadores \forall y \exists .

Los axiomas especiales son los siguientes:²⁷⁶:

Axioma Q1. Si α y β son fórmulas congruentes [Kleene 1952] o una se obtiene de la otra por eliminación de cuantificadores vacuos, entonces $\alpha \leftrightarrow \beta$.²⁷⁷ .

Axioma Q2. $\forall x(P(x))^{(n)} \supset (\forall xP(x))^{(n)}$

Axioma Q3. $\forall x(P(x))^{(n)} \supset (\exists xP(x))^{(n)}$

Si tomamos el cálculo $C^{*=1}$ y añadimos el esquema $\neg(\alpha \wedge \neg\alpha)$ como un postulado más, obtenemos la Lógica Clásica de Primer Orden con identidad, que denotaremos, como hace da Costa, por $C^{*=0}$.

Una observación importante es que en los cálculos $C^{*=n}, 1 \leq n \leq \omega$, no son válidos los siguientes esquemas²⁷⁸:

²⁷⁵ Esto se describió en el Capítulo III.

²⁷⁶ [D'Ottaviano 1990], p. 116.

²⁷⁷ *Idem.*

²⁷⁸ *Idem.*

$$\neg\exists x\neg A(x) \leftrightarrow \forall xA(x)$$

$$\neg\forall x\neg A(x) \leftrightarrow \exists xA(x)$$

$$\neg\exists xA(x) \leftrightarrow \forall x\neg A(x)$$

$$\exists x\neg A(x) \leftrightarrow \neg\forall xA(x)$$

Ahora presentaremos una generalización natural de la Lógica *Default* asumiendo C^*_0 y la concepción general de las reglas que describimos en la última subsección.

Definición 11. Sea $((A, \gamma_i), B)$ una regla de Σ , y $Z, K^\epsilon, K^\epsilon \subseteq FOR$, y K^ϵ es el contexto de derivación de A , y K^ϵ el contexto de asunción de B , ambos en $\Sigma+Z$. Entonces decimos que $((A, \gamma_i), B)$ está **cerrada bajo el conjunto Z** en Σ sii:

- 1) Si $A \subseteq K^\epsilon$ y $Cn^\Sigma(B \cup K^\epsilon) \neq FOR$, Entonces $\gamma_i \in Cn^\Sigma(Z)$.

Usamos $T_\Sigma \subseteq FOR$ como el conjunto de todas las tautologías de Σ . Y $\perp = FOR$, como el conjunto de todas las *fbfs* con $\cup_{i=1}^5 V$. Así, tenemos:

Definición 12 Si una regla $((A, \gamma), B)$ es tal que Si $A = \{r_1, r_2, r_3, \dots, r_n\}$ Entonces $(r_1 \wedge r_2 \wedge r_3 \wedge \dots \wedge r_n \supset \gamma) \in T_\Sigma$, decimos que $((A, \gamma), B)$ es **débilmente válido** en Σ . En este caso también decimos que γ es **relativamente infalible** para $((A, \gamma), B)$.

$((A, \gamma), B)$ es **fuertemente válido** sii If cada $\theta \in A$ es verdadero es verdadero Entonces γ es verdadero. En este último caso decimos que γ es **absolutamente infalible** para $((A, \gamma), B)$.

Definición 13. Una *sr* (A, γ) en una **regla clásica (cr)** de un *eps* Σ sii:

- a) Para cada $Z \subseteq FOR$, (A, γ) está cerrada bajo el conjunto $Cn^\Sigma(Z)$.
- b) (A, γ) es débilmente válida in Σ .

Definición 14.

Sea $E, Z \subseteq FOR$, el *eps* $\Sigma = \langle \Lambda, T \rangle = \langle \langle R, S \rangle, \langle P, W \rangle \rangle$. Si:

$$E_0 = W \cup Z$$

Y para toda E_i $i \neq 0$:

$$E_{i+1} = \text{Cn}^\wedge(E_i) \cup \{ \gamma \mid ((A, \gamma), B) \in \text{RUP} \text{ tal que es aplicable en el punto } E_i. \}$$

Entonces decimos que E es una **extensión** de $\langle \Lambda, T \rangle + Z$ (i.e. $\langle \langle R, S \rangle, \langle P, W \cup Z \rangle \rangle$) *sii*:

$$E = \bigcup_{i=0}^{\infty} E_i.$$

Llamamos a cada E_i una **semiextensión** de E .

Definición 15 Una **derivación clásica** de una *fbf* ϕ a partir de un conjunto Z de fórmulas en un cpes $\Sigma = \langle \langle R, S \rangle, \langle P, W \rangle \rangle$, es una secuencia de *fbfs* $\langle \gamma_1, \dots, \gamma_n \rangle$ tal que:

- 1) $Z \vdash_{\Sigma} \phi$ (ϕ es una Σ -consecuencia a partir de Z)
- 2) Cada $\gamma_i \in \{ \gamma_1, \dots, \gamma_n \}$ es la conclusión de una *cr* $((A, \gamma), B) \in \text{RUP}$ or $\gamma_i \in S \cup W$.
- 3) Cada $((A, \gamma), B) \in \text{RUP}$ es aplicable en el punto γ_i *sii* para cada $\theta \in A$, $\theta \in \{ \gamma_1, \dots, \gamma_{i-1} \}$.

Dado un conjunto $K \subseteq \text{FOR}$, llamaremos a K un **conjunto consistente de fórmulas** *sii* no existe ninguna *fbf* θ tal que $\theta \in K$ y también $\neg \theta \in K$. En el caso contrario, llamaremos a K un **conjunto inconsistente de fórmulas**. Llamaremos a K un **conjunto trivial de fórmulas** *iff* $K = \perp$.

Así, dado un *eps* Σ , llamamos a Σ un **eps explosivo** *sii*: para todo K inconsistente, $\text{Cn}^{\Sigma}(K)$ is trivial (i.e. $\text{Cn}^{\Sigma}(K) = \perp$)

Teorema 1. Dadas las siguientes condiciones: $\Lambda = \text{C}^*_{=0}$, toda *sr* en $\langle \Lambda, T \rangle$ es una *cr*, and $Z \subseteq \text{FOR}$, Entonces $\langle \Lambda, T \rangle + Z$ tiene una única extensión.

Teorema 2. Dadas las siguientes condiciones: $\Lambda = \text{C}^*_{=0}$, toda *sr* in $\langle \Lambda, T \rangle = \langle \langle R, S \rangle, \langle P, W \rangle \rangle$ es una *cr*, y $Z \subseteq \text{FOR}$, Entonces $\langle \Lambda, T \rangle + Z$ tiene una extensión inconsistente *sii* $Z \cup S \cup W$ es inconsistente.

Teorema 3. Dadas las siguientes condiciones: $\Lambda = C^*_{=0}$, toda sr in $\langle \Lambda, T \rangle = \langle \langle R, S \rangle, \langle P, W \rangle \rangle$ es una cr , and $Z \subseteq FOR$, Entonces $\langle \Lambda, T \rangle + Z$ tiene una extensión inconsistente E *sii*:

- 1) E es la única extensión de $\langle \Lambda, T \rangle + Z$
- 2) $Cn^{\langle \Lambda, T \rangle}(Z) = \perp$.

Definición 16 Una regla $((A, \gamma), B)$ es **una regla no monotónica de nivel 0 (nmr)**, o **regla retractable** en Σ *sii* existe un par de conjuntos $Z, R \in FOR$, tal que:

- 1) Existe por lo menos una Σ -derivation $\langle \gamma_1, \dots, \gamma_n \rangle$ de γ a partir de Z (Consecuentemente, $Z \vdash_{\Sigma} \phi$) tal que $((A, \gamma), B)$ es aplicable al punto γ_{n-1} .
- 2) Toda Σ -derivation $\langle \gamma_1, \dots, \gamma_n \rangle$ of γ de $Z \cup R$ es tal que $((A, \gamma), B)$ no es aplicable al punto γ_{n-1} .

Cuando $R \neq \emptyset$, podemos decir que $((A, \gamma), B)$ es una **regla nomonotónica de nivel 1 (nmr1)**. Podemos decir acerca de una *nmr1*, que es una **regla nomontónica de nivel 2 (nmr2)**, *iff* $R \neq Z$. Finalmente, una *nmr2* es una **regla no monotónica de nivel 3 (nmr3)** *sii* $R - Z \neq \emptyset$.²⁷⁹

Nótese que una regla puede ser no monotónica a nivel $n > 0$ en Σ y ser débilmente o fuertemente válida. La validez, es independiente de cualquier nivel de no monotonicidad. Consecuentemente, una conclusión γ de una regla a no monotónica puede ser relativamente o absolutamente infalible, o podría ser también una consecuencia falible. Nótese que también podemos extender los niveles de nomonotonicidad en una regla, mediante restricciones en las interacciones entre las reglas o entre los elementos de las reglas y otros elementos del sistema.

Un *eps* Σ es **retractable** *sii*: para algunos $Z, R \subseteq FOR$, $Z \vdash_{\Sigma} \gamma$ y No ocurre que $Z \cup R \vdash_{\Sigma} \gamma$.

Un *eps* Σ es **relevantemente** (i. e. $n > 0$) **no monotonioc** *sii*: es retractable y $R \neq \emptyset$.

²⁷⁹ Es interesante notar que podemos todavía pensar en un nivel 4 para la no monotonicidad, más racional que los niveles 1-3, con una restricción acerca de qué clase de información introduce R en el sistema. Por lo menos para fines epistémicos, R no sólo debe contener información sino que, su información debe realmente ser relevante para producir una retractación en la inferencia.

Un **default de Reiter** es una $((A, \gamma), B)$ tal que A (contexto de presencia) solamente tiene contexto de derivación y B (contexto de ausencia) solamente tiene contexto de asunción. La estructura precisa de un *default* puede presentarse como $(\{\alpha\}, \omega, \{\beta_1, \dots, \beta_n\})$. Reiter presenta un *default* como:

$$\alpha : \diamond \beta_1, \dots, \diamond \beta_n / \omega$$

Que puede leerse: “ α , such that it is consistent to assume every β_i in the sequence β_1, \dots, β_n , so infer ω ”. Aquí el diamante está interpretado como posibilidad en el sentido de ser consistente respecto de un conjunto de proposiciones.²⁸⁰

Como podemos ver, nuestra noción de regla es una generalización de una noción del *default*. Por otro lado, nuestra presentación de un *default* podría ser llamada **default de Reiter generalizado (default-RG)**. La razón es que nuestra noción de regla posee una generalización basada en trivialidad en la condición de ausencia B de una regla. Este cambio nos permitirá integrar, de una forma natural, paraconsistencia y razonamiento default en la siguiente sección.

Definición 17. Una regla $((A, \gamma), B)$, i.e., $(\{\alpha\}, \omega, \{\beta_1, \dots, \beta_n\})$ en Σ , es un **default de Reiter generalizado** *sii*, dado un conjunto $Z \subseteq \text{FOR}$,

- A) Ambos A and B tienen solamente contextos de derivación: $K^\in, K^\notin \subseteq \text{FOR}$.
- B) $((A, \gamma), B)$ está cerrada bajo Z en Σ .
- C) $K^\in = K^\notin = \text{Cn}^\Sigma(Z)$.

Esta estructura argumentativa es ya interesante para el problema de representacion de la explicación. Yo propuse *defaults* como un mecanismo básico de representación en mi tesis de maestría precisamente por su capacidad de capturar las explicaciones como argumentos con baja probabilidad y con casos de excepción indefinibles.

²⁸⁰ Este proceso de chequeo, no es necesariamente tan fácil como puede parecer. La clave es considerar el final de la extensión para este proceso.

Definición 18 Una teoría *default* es un *eps* $\Delta = \langle D, W \rangle$, tal que D es una teoría y W es un conjunto de reglas *default*.

Definición 19. Un *sistema de Reiter* es un *coceps* $\Sigma = \langle \Lambda, \Delta \rangle$ donde $\Lambda = C_0^*$, i.e., lógica clásica, y Δ es una teoría *default*.

Teorema 4. Un sistema de Reiter $\Sigma = \langle \Lambda, \Delta \rangle + Z$ tiene una extensión inconsistente *sii* $W \cup Z$ es inconsistente.

Teorema 5. Para todo par de extensiones de un sistema de Reiter, E_1 y E_2 , $E_1 \cup E_2$ es inconsistente. En este caso decimos que E_1 y E_2 son ortogonales.

Corollary 6. Si un sistema de Reiter $\Sigma + Z$ tiene una extensión inconsistente entonces ésta es su única extensión.

4.1.3.3 Una Teoría *default* paraconsistente.

Esta interesante forma de producir inferencias nos conduce a varias extensiones de las conclusiones, dependiendo de qué defaults apliquemos, y puede capturar la retractabilidad de la explicación científica: la noción resultante de consecuencia es no monotónica. Nuestra estrategia central para comparar explicaciones en contextos teóricos paraconsistentes será: tomar T como una teoría *default* representando una teoría científica y Λ como una lógica de la jerarquía de cálculos C_n de da Costa. Sin embargo, dado que una derivación *default* está basada normalmente en la observación de consistencia, esto produce una dificultad para conectar defaults y lógicas paraconsistentes. Por esta razón fue conveniente la modificación de la definición de un *default*.

Con nuestro esquema general de regla podemos exponer definiciones independientes de extensión y diferentes clases de *default*, y también, con los elementos desarrollados en las

secciones anteriores, podemos definir nociones diferentes de extensiones. Para ello asumiremos los axiomas de la jerarquía de cálculos paraconsistentes de da Costa's C_n^* , $1 \leq n < \omega$.

Recuérdese que la base del chequeo en los *defaults* no es ya la consistencia, sino la trivialización de la extensión. Esto es un cambio que nos permite usar lógica paraconsistente asociada a una teoría *default* mediante el operador de consecuencias (/teoremicidad) C_n .

Podemos ahora construir un sistema epistémico no monotónico y paraconsistente, la clase de sistema epistémico que queremos usar para modelar la interacción entre explicaciones y contextos teóricos:

Definición 20. Definimos un *sistema Reiter-da Costa* (o un *sistema-RC*) como un $eps \langle \Lambda, \langle D, W \rangle \rangle + Z$ donde Λ es algún cálculo de la jerarquía C_n^* , $1 \leq n < \omega$, y D contiene sólo *defaults*-RG.

Teorema 7. Si E es una extensión de un sistema-RC $\Sigma + Z$, podría pasar que E sea inconsistente pero $E \neq \perp$.

Teorema 8. Un sistema-RC $\Sigma + Z$ tiene una extensión trivial *sii* $W \cup Z$ es trivial.

Teorema 9. Todo par de extensiones de un sistema-RC E^1 y E^2 , son ortogonales entre sí.

Definición 20 Un sistema RC Σ es *consistente* *sii* tiene una extensión consistente.

Definición 21 Decimos que un par de extensiones de un sistema RC, E^1 y E^2 , son *paraortogonales* entre sí *sii* $E^1 \cup E^2 = \perp$.

Teorema 10 Todo par de extensines de un Sistema-RC, E^1 y E^2 , son *paraortogonales* entre sí.

Hemos modificado los mecanismos de inferencia en las reglas *default*, y con este cambio ya no sólo podemos representar inferencias con baja probabilidad sino también inferencias no monotónicas en contextos paraconsistentes. Podemos representar no sólo contextos teóricos con una *W* inicial inconsistente, sino también inferencias por *default* que producen extensiones inconsistentes con *fbfs* contradictorias aceptadas. Con esta estructura general de sistemas-RC, podemos ver explicaciones como argumentos especiales dentro de teorías, representados por una clase particular de *defaults*. Estos argumentos pueden pertenecer a *D* dentro de la estructura de sistema-RC. Con estos sistemas epistémicos ya podemos capturar un fragmento importante de la dinámica de la explicación científica. Podemos:

- 1) Representar la explicación científica como un argumento sin el ideal de buena explicación. Una explicación puede apoyar su conclusión con baja probabilidad. Además, puede ser falible y retractable y puede representarse aún cuando tenga excepciones indefinibles o, incluso sin leyes determinadas. Si basamos las explicaciones en la estructura de una regla tipo *default*, podremos proponer explicaciones en un sentido muy pragmático y con una noción de inferencia más débil.
- 2) Representar el cambio explicativo. Una explicación puede parar de derivar un *explanandum* si tenemos información adicional en el contexto teórico. Su comportamiento es no monotónico.
- 3) Representar explicaciones como predicciones falibles (*explananda* como conjeturas) que pueden ser retractables en una forma no monotónica (nivel $0 < 1$).
- 4) Representar relaciones entre lógicas, teorías y explicaciones. Podemos examinar si una teoría puede aplicar las mismas explicaciones con dos diferentes lógicas paraconsistentes. En realidad, uno podría extender el tipo de lógica, incluso más allá de la jerarquía de cálculos de da Costa. Podemos examinar si una explicación

(*default rule*) de una teoría puede aplicarse coherentemente en otros diferentes contextos que podrían contener diferentes asunciones acerca de la realidad.

5) Representar relaciones entre los *explanantes* directos (conjuntos de proposiciones) and diferentes *explanantes* contextuales (teorías o sistemas epistémicos complejos).

6) Representar explicaciones como objetos que funcionan en contextos teóricos inconsistentes.

7) Analizar el comportamiento de diferentes clases de conjuntos de conclusiones a partir del mismo conjunto de condiciones iniciales.

8) Analizar explicaciones con predicciones en conflicto (con *explananda* mutuamente contradictorios) en diferentes extensiones de una misma teoría.

9) Representar, en uno y el mismo contexto teórico, diferentes estrategias heurísticas (diferentes secuencias de *defaults* aplicados) para alcanzar una misma conjetura (conclusión) o el mismo conjunto de conclusiones.

10) Representar combinaciones entre *explanantes* directos, *explanantes* contextuales, lógicas y *explananda*.

11) Analizar una explicación, sus interacciones, dentro de su contexto.

Estos son ya puntos buenos a favor de esta nueva representación de la explicación científica entendida en el paradigma de un producto de la actividad científica, sobretodo en el sentido de que ofrece un análisis dinámico de esta noción y un giro interesante en las asunciones generales que normalmente atribuimos a la racionalidad científica. Con este Marco Formal Ex, estamos resolviendo los problemas de representación A-D (presentados al principio de esta sección) y, estamos también contribuyendo así, mediante herramientas lógicas paraconsistentes y otras usadas principalmente en Inteligencia Artificial, a resolver los problemas de tipo I-III, mediante la defensa de una concepción epistémica y pragmática de las explicaciones.

Pero las características de una explicación científica no son fáciles de capturar. No desde el punto de vista de la Filosofía de la Ciencia. Los problemas que giran en torno a la relación de pertinencia explicativa, son tan importantes como los otros y tienen que ver con la dificultad de precisar qué clase de relación uno exigiría en un argumento explicativo. La discusión en Filosofía de la Ciencia nos dirige a muchas restricciones a la inferencia que modificarían y harían muy complejo un argumento para que pueda realmente representar una explicación.

Necesitamos, como establecimos en el capítulo III, una conexión más profunda entre prerrequisito y conclusión (o conjetura) in la regla *default*.

Pero, ya hemos definido condiciones nuevas para la inferencia, ¿qué otros elementos deberíamos añadir? ¿De qué están hablando los filósofos cuando exigen estas condiciones? Notemos que ya hemos caracterizado relaciones entre explicaciones y otras cosas en el contexto. No obstante no hemos caracterizado lo que puede marcar una diferencia en el contexto teórico, entre una explicación y otra clase de argumentos. ¿Toda regla *default* en D es una explicación? ¿Qué clase de relación de pertinencia explicativa uno debe suponer cumpliéndose en una explicación? esta última cuestión es difícil porque apunta a la interconexión entre dos cosas: las condiciones que supuestamente se cumplen en la inferencia y, por otro lado, las condiciones que supuestamente se cumplen entre los sucesos o eventos involucrados en la explicación. Como he defendido en el capítulo II, hay una confusión de estas dos clases de elementos en la discusión acerca de los problemas de las asimetrías explicativas. Algunos filósofos intentan proponer una clase de relación de pertinencia explicativa, y otros amplían irrestrictamente la gama de posibles relaciones de este tipo. Pero, en la discusión, parecería, como expliqué en el capítulo II, que toman en un sentido vago la determinación de estas relaciones. Además, la elección de relaciones de pertinencia explicativa es muy difícil de establecerse fundadamente. Desde el punto de vista de algunas posiciones filosóficas, es imposible encontrar una relación de pertinencia explicativa universalmente válida para toda teoría²⁸¹. De cualquier forma, a pesar de lo que ya hemos avanzado en la elucidación de la explicación (capturando la

²⁸¹ Por ejemplo Achinstein o Van Fraassen.

dinámica de la explicación), si nosotros queremos una idea más incluso de cómo funciona la explicación (sus funciones internas y externas) debemos añadir elementos que puedan tomar en cuenta los problemas de la relación de pertinencia explicativa. Yo defendí en el capítulo II y en parte del III, que hay por lo menos una estrategia que podría proponer sólo dos clases de relación de pertinencia explicativa intentando con esa partición capturar todas las clases de relaciones que pueden defenderse como relaciones intuitivas de pertinencia explicativa. El análisis supone la distinción entre relación de explicación, explicación y relaciones de pertinencia explicativa que he recordado al principio de este capítulo. La relación de pertinencia explicativa es una que se da entre los fenómenos involucrados en las proposiciones que constituyen una explicación²⁸².

Defendí también, al final del capítulo III, que una representación fuerte de una noción de consecuencia, que realmente pueda tener aplicaciones de modelación en Ciencia, tendría que tomar en cuenta la diferencia entre inferir e inferir justificadamente, en relación con ciertos fines epistémicos. Así, por lo pronto, algo que mejoraría la racionalidad de nuestro modelo de explicación sería construir vínculos más fuertes entre diferentes *defaults* usados en una inferencia. Por ejemplo, un *RG-default* no exige coherencia entre diferentes *provisos*.

Definición 21. Sea una función $X(\langle \Lambda, \Delta \rangle + Z)$ tal que resulta en el conjunto de todas las extensiones E^i de $\langle \Lambda, \Delta \rangle + Z$.

Definición 22. Definamos además una función H que aplicada a una extensión $E^h \in X(\langle \Lambda, \langle D, W \rangle \rangle + Z)$, donde $\langle \Lambda, \langle D, W \rangle \rangle$ es un sistema-RC, y $D = \{d_1, d_2, \dots, d_n\}$, resulta en el conjunto de secuencias $\{s_1, s_2, \dots, s_m\}$ tal que, para toda s_i :

- a) $s_i = d_{i1}, d_{i2}, \dots, d_{ik}$ es un orden en el cual estos *defaults* $(d_{i1}, d_{i2}, \dots, d_{ik})$ pueden ser aplicados para producir E^h .

- b) cada $d_i \in D$.
- c) No necesariamente $n = k$.
- d) No necesariamente $k \neq 0$.
- e) Cada s_i mantiene la siguiente relación con E^h :

$$E_0 = W,$$

$$E_1 = \text{Cn}^\wedge(E_0) \cup \{\omega_1 \mid \omega_1 \text{ es una conjetura de } d_{i1} \text{ en } s_i\}$$

$$E_2 = \text{Cn}^\wedge(E_1) \cup \{\omega_2 \mid \omega_2 \text{ es una conjetura de } d_{i2} \text{ en } s_i\}$$

.

.

.

$$E_k = \text{Cn}^\wedge(E_{k-1}) \cup \{\omega_k \mid \omega_k \text{ es una conjetura de } d_{ik} \text{ en } s_i\}$$

$$E_k = E^h.$$

Decimos que $H(E^i)$ es el conjunto de **estrategias heurísticas** de $\langle \Lambda, \langle D, W \rangle \rangle$ **para** E^i .
Decimos que toda $s_i \in H(E^i)$ es una estrategia heurística de $\langle \Lambda, \langle D, W \rangle \rangle$ para E^i .

Definición 23 Sea $\langle \Lambda, \Delta \rangle$ un sistema-RC and sea $E^i \in X(\langle \Lambda, \Delta \rangle)$. Definimos una función J que aplicada a una estrategia heurística s_k de $\langle \Lambda, \Delta \rangle$ para E^i , esto es, $J(s_k)$, resulte en el conjunto de todas las justificaciones de cada *default* en s_k .
Definición 24. Sea $\Sigma = \langle \Lambda, \Delta \rangle + Z$ un sistema epistémico complejo donde Δ es una teoría *default* tal que $\Delta = \langle D, W \rangle$. Decimos que un *default* $d = (\alpha : \diamond \beta_1, \dots, \diamond \beta_n / \omega) \in D$, i.e., un $((A, \gamma), B) = (\{\alpha\}, \{\beta_1, \dots, \beta_n\})$, es un *default-RGS* (*default* de Reiter Generalizado con coherencia de *provisos*) *sii* para cada extensión E , en toda estrategia heurística s_k de Σ para E :

- A) A tiene solo contexto de derivación: $K^\epsilon \subseteq \text{FOR}$.
- B) B tiene contexto de derivación, $K^{\neq d} \subseteq \text{FOR}$, y contexto de asunción, $K^{\neq a} \subseteq \text{FOR}$,
- C) Si $A \subseteq K^\epsilon$ y $\text{Cn}^\Sigma(B \cup K^{\neq d} \cup K^{\neq a}) \neq \text{FOR}$, entonces $\omega \in E$.
- D) $K^\epsilon = K^{\neq d} = E$.

$$E) K^{\neq a} = J(s_k).$$

Toda E construida con estas condiciones será llamado una **extensión-RGS de $\Sigma+Z$** .

Nótese otra vez que el chequeo de posibilidad está basado en trivialización pero incluyendo justificaciones. Esto es un cambio que contribuye a generar mayor racionalidad en la construcción de una extensión, como lo discutimos en el capítulo III.

Una teoría *default* $\Delta = \langle D, W \rangle$ cuyas reglas $A/\omega \in D$ son default-RGS puede llamarse una **teoría default-RGS**. Y un sistema epistémico complejo que es igual a un sistema-RC excepto porque en vez de tener una teoría *default*-RG, tiene una teoría *default*-RGS, puede llamarse un **sistema-RCS**.

En contextos inconsistentes puede ser posible que tengamos, para un mismo sistema-RCS, en la misma estrategia heurística s_k , dos *default*-RGS con prerequisites mutuamente contradictorios. Un asunto que tratamos ya en el capítulo III y que con este Marco Formal podría ser representado.

4.1.3.4 El Modelo de Explicación

En esta sección introduciremos el modelo de explicación integrado al marco formal, de modo que nos sirva para revisar si un argumento en un sistema epistémico cumple con una noción de explicación mínima, revisar a su vez las interacciones de los elementos de estas explicaciones mínimas y las interacciones de las explicaciones mínimas con su contexto.

El refuerzo que establecimos entre el vínculo entre justificaciones y conjeturas, no es todavía suficiente para resolver los problemas acerca de la relación de pertinencia explicativa (problemas a los que he estado haciendo referencia en esta sección).

Introduciremos ahora las dos clases de relación de pertinencia explicativa que he defendido en el capítulo II como relaciones que podrían usarse en una estrategia para capturar desde un enfoque contextualista muchas relaciones de pertinencia explicativa que son intuitivas y, sin embargo, no relativizar sin restricciones esta clase de restricción en los argumentos que modelan explicaciones.

Primero quiero recordar algunas reflexiones que hice en el capítulo II, conectadas ahora con este marco formal, y que dan una justificación de la partición de relaciones de pertinencia explicativa que proponemos. Si encontramos aspectos de un argumento *default*, tales que podamos caracterizar con ellos una noción de explicación, entonces querríamos que esos aspectos pudieran mantenerse en un nivel distinto para nuestros sistemas epistémicos complejos. Desde un enfoque no relativista, deberíamos desear que ciertos aspectos de las explicaciones pudieran mantenerse como propiedades fijas, más allá de la ontología supuesta por diferentes sistemas epistémicos. Además, esto sería conveniente porque hacerlo tendría por lo menos tres ventajas: a) tendríamos un criterio externo en la determinación de si un par <conjunto de proposiciones, *explanandum*> es o no una explicación. Consecuentemente, tendríamos razones externas (externas respecto de todo eps particular) acerca del hecho de que un candidato es una explicación en un dado sistema-RCS Σ_1 mientras que no lo es en otro sistema-RCS Σ_2 . b) Podríamos tener un mecanismo de generación de explicaciones, un mecanismo específico para generar *defaults*. Podríamos usar este criterio externo para producir *defaults* dentro de un eps. Por ejemplo, si un par <conjunto de proposiciones-*explanandum*> es una explicación respecto de un sistema-RCS $\langle \Lambda, \Delta \rangle$ entonces podemos “copiar” ese par construyendo un *default* dentro de Δ . Por supuesto que la cuestión de cómo necesitaremos restringir esta construcción²⁸³ de *defaults* tendrá que ser pensada en detalle. c), Creemos que un estructura extrasistémica para reconocer explicaciones es un enfoque similar a la práctica de la actividad científica, no sólo en el sentido de que los científicos usan exitosamente

²⁸³ Debido a que hay razones para pensar que no podemos racionalmente construir un *default* sobre la base dos *defaults* encadenados (el silogismo hipotético no es siempre una opción en lógica deductiva). Véase, *supra*, sección, 3.4.5.

explicaciones como evidencia para confirmar una teoría (en este caso una teoría *default* en un eps) pero también en el sentido de que algunos científicos usan diferentes explicaciones de diferentes sistemas epistémicos a la hora de intentar resolver un mismo problema.

Una cosa que debemos tener en mente es que estamos intentando usar la relación de pertinencia explicativa capturándola como un elemento unificador, no sólo como un elemento que restringe, en la inferencia. Usarlo como un elemento que conecte el prerrequisito con la conjetura en un *default*.

El primer concepto que introduciremos ahora es el concepto de cadena de fenómenos o “[est]”. Una cadena es una estructura que conecta sucesos de nuestra realidad mediante relaciones de cierta clase. La idea intuitiva de una cadena es que cada individuo de la cadena está conectado con otro, que está conectado con otro y así sucesivamente. Así, podemos decir que el primer individuo está conectado, en cierta forma, con el último de los individuos de la cadena. Nuestra propuesta es que esta clase de estructura puede ser lo que la gente tiene en mente cuando formulan una explicación. La gente, en el propósito de evaluar algo como una explicación, o de entenderla como tal, usualmente busca la conexión del *explanandum* en una estructura más general y también busca la conexión entre algo que se dice en el *explanans* y algo que se dice en el *explanandum*.

Podemos poner esta idea en el *proviso* de un *default*. i.e., en su condición de ausencia, e indicar que el *explanandum* está conectado con cierta cadena. Para hacer esto es necesario tener en el Marco Formal la capacidad de hablar de sucesos de la realidad. Necesitamos ser capaces de hablar de sucesos como individuos de una cadena [est].

Una dificultad es que, desde nuestra perspectiva de la explicación, estamos concibiendo una relación de explicación como una relación entre un conjunto de proposiciones y un fenómeno particular de la realidad. Por lo tanto, necesitamos establecer un vínculo con

una cadena, en el nivel de la realidad, en la forma de una relación de pertinencia explicativa y, por otro lado, necesitamos algo que involucre esta conexión en un nivel de las proposiciones relacionadas con el *explanans*. ¿Cómo podemos hacer esto? Podemos usar una descripción de la cadena en el *explanans*, de los elementos que pertenecen a una cadena. Aquí tenemos la idea central del modelo como un trasfondo teórico para observar las interacciones de las explicaciones. Los detalles son más complicados si consideramos que algunas veces, la descripción de una estructura (cadena) en la cual queremos ubicar al *explanandum* no es derivable del *explanans* directo solo. Y esto es supuesto en muchos casos de explicación desde un punto de vista pragmático. Por lo tanto, parece que el prerrequisito de una explicación no debería ser simplemente una fácil verificación de un conjunto de fórmulas en un contexto de derivación. Una explicación puede ser capturada más bien en términos de un argumento muy complejo que relaciona un conjunto de proposiciones, un fenómeno particular, un contexto teórico (sistemas epistémicos con lógicas específicas), y una relación de pertinencia entre diferentes clases de eventos. Todas estas cosas en el ambiente no monotónico y paraconsistente que ya hemos logrado capturar en algún grado. Estas últimas líneas pueden servirnos como un breve bosquejo de cómo el modelo resultante es más complejo de lo que podíamos esperar, puede verse como un metaargumento que relaciona todos estos elementos. Véamos esto más detalladamente.

Informalmente, la idea general es que si podemos derivar, relativamente a un sistema epistémico Σ (el *explanans* contextual), a partir de un conjunto A (el *explanans* directo), una *descripción* de una cadena de fenómenos relacionados, y es posible en el contexto de Σ que el fenómeno *explanandum*, digamos e , sea el último elemento de la cadena, entonces A y e mantienen una relación de explicación. Esta es idea clave que proponemos en este modelo. Definamos algunos símbolos para exponerla formalmente.

Definición 25. Decimos que una fórmula ϕ of Ex es paraconsistentemente posible en el contexto del sistema-RCS $\Sigma = \langle \langle R, S \rangle, \langle D, W \rangle \rangle + A$, en símbolos, $\diamond_{\Sigma} \phi$, *sii* existe una extensión E con una estrategia heurística s_k of Σ such that:

$$\text{Cn}^{\Sigma}(\{\phi\} \cup A \cup J(s_k)) \neq \perp.$$

Otra manera de ver la posibilidad paraconsistente es la siguiente. En el contexto de un alguna lógica en la jerarquía C_n^* , $1 \leq n < \omega$, y suponiendo A en Σ , tenemos que $\diamond_{\Sigma} \alpha$ *sii* existe una extension E con una estrategia heurística s_k of Σ , tal que:

$$\alpha \in \text{Cn}^{\Sigma E}(J(s_k)) \text{ O } \neg \alpha \notin \text{Cn}^{\Sigma E}(J(s_k))$$

Si no suponemos la verdad de la disyunción anterior, tendremos 4 casos:

i) *Caso de Redundancia:* cuando α está en $\text{Cn}^{\Sigma E}(J(s_k))$ y es una fórmula bien comportada.

$$\alpha \in \text{Cn}^{\Sigma E}(J(s_k)) \text{ Y } \neg \alpha \notin \text{Cn}^{\Sigma E}(J(s_k)).$$

ii) *Caso de Trivialización:* Cuando α no está en $\text{Cn}^{\Sigma E}(J(s_k))$ Y es una fórmula bien comportada.

$$\alpha \notin \text{Cn}^{\Sigma E}(J(s_k)) \text{ Y } \neg \alpha \in \text{Cn}^{\Sigma E}(J(s_k)).$$

iii) *Caso de Paraconsistencia:* Cuando $\neg \alpha$ está en $\text{Cn}^{\Sigma E}(J(s_k))$ Y es una fórmula mal comportada.

$$\alpha \in \text{Cn}^{\Sigma E}(J(s_k)) \text{ Y } \neg \alpha \in \text{Cn}^{\Sigma E}(J(s_k)).$$

iv) *Caso de Ignorancia:* Cuando $\alpha, \neg \alpha$ no están en $\text{Cn}^{\Sigma E}(J(s_k))$.

$$\alpha \notin \text{Cn}^{\Sigma E}(J(s_k)) \text{ Y } \neg \alpha \notin \text{Cn}^{\Sigma E}(J(s_k)).$$

α es paraconsistentemente posible en los casos i, iii y iv. El caso de trivialización hace la disyunción falsa. Escribimos $\diamond_{\Sigma E}(A)$ para denotar que toda $\beta \in A$ es paraconsistentemente posible en Σ .

Como recientemente dijimos, los sucesos son importantes como elementos en interacción con una explicación. Ex debe ser capaz de capturarlos para analizar un candidato a ser una explicación. En razón de representar eventos y relaciones entre ellos, introduciremos la noción de conjunto de fenómenos y un mecanismo para considerar estos conjuntos como teniendo una estructura interna. Esta estructura puede ser vista como una representación de la existencia de relaciones de pertinencia explicativa entre relaciones entre los fenómenos. En razón de obtener esta clase de representación introduciremos los siguientes símbolos y usaremos definiciones comunes de la Teoría de Grafos para representar una estructura dentro de los conjuntos de fenómenos.

Definición 26. Decimos que:

- a) $est_i n$ es un conjunto i de n fenómenos, $[est]_i n$ una cadena constituida de elementos de $est_i n$, y $D(est_i n)$, la descripción en proposiciones, de los elementos de $est_i n$.
- b) $[est]_i n$ es una cadena de fenómenos *si* es una secuencia de eventos que guardan una estructura (relaciones entre ellos) tal que esa estructura puede ser representada por un digrafo no cíclico, no nulo, no vacío, no trivial, sin aristas paralelas, tal que:
 - b1) El conjunto V of vértices es $est_i n = \{e_1, e_2, e_3, \dots, e_n\}$.
 - b2) El conjunto E de aristas (edges) is $\{(e_1, e_2), (e_2, e_3), \dots, (e_{n-1}, e_n)\}$.
- c) Leemos $e_j \succ [est]_i n$ como “ e_j es el próximo miembro de la cadena $[est]_i n$ ”, y significa que existe una cadena $[est]_i n+1$ tal que difiere de $[est]_i n$ sólo en su conjunto de aristas E' :

$$E': \{(e_1, e_2), (e_2, e_3), \dots, (e_{n-1}, e_n), (e_n, e_j)\}.$$
- d) Decimos que $D(e)$, $D(est_i n)$, $D([est]_i n)$, $D(e \succ [est]_i n)$ son respectivamente la descripción de la existencia de un evento e , la descripción de la existencia de los

elementos de $est_i n$, la descripción de la existencia de la cadena $[est]_i n$, la descripción de la existencia de la cadena $[est]_i n$ tal que e es el último miembro de la cadena.

Hypothesis 27 Como una hipótesis metodológica, consideraremos que:

HM1) Sólo existe en in Ex estas clases de fenómenos: fácticos o no fácticos.

HM2) Sólo existen en Ex dos clases de relación de pertinencia explicativa (capturadas como aristas en un dígrafo que representa una cadena de fenómenos: relaciones causales entre fenómenos fácticos y relaciones inferenciales entre fenómenos no fácticos).

Ahora estamos listos para caracterizar los complejos aspectos mínimos de la explicación desde el punto de vista de los problemas de representación en Filosofía de la Ciencia.

Definición 28. Hay una relación de explicación entre un conjunto de proposiciones A y el fenómeno e relativo a un sistema-RCS Σ *sii* existe una E de Σ y un conjunto $est_i n$, tal que A y e mantienen una relación descrita como sigue:

Ex1) O bien $A \vdash_{\Sigma E} D(est_i n)$ o bien $A \mid \sim_{\Sigma E} D(est_i n)$

(la descripción de los elementos del conjunto de fenómenos $est_i n$, $D(est_i n)$, es derivada a partir de A en una forma deductiva, mediante una Σ -derivation clásica de E , o en una forma no monotónica, mediante una Σ -derivation no clásica de E).

Ex2) $\diamond_{E^{\Sigma+A}}(D(e \times [est]_i n))$

(Es paraconsistentemente posible en el contexto E de Σ que e sea el próximo miembro de una cadena $[est]_i n$).

Ex3) $\diamond_{E^{\Sigma}}(A)$

(Todo elemento de A es paraconsistentemente posible en el contexto E de Σ).

Ex4) $\diamond_{E^\Sigma}(D(e))$

Ex5) Si e es un fenómeno fáctico entonces todo vínculo de la cadena (arista) $[est]_i^n$ (representada en su dígrafo correspondiente por una arista) es un vínculo causal. Si, en cambio, e no es un fenómeno fáctico, entonces todo vínculo de la cadena es inferencial.

Nótese que en esta perspectiva la relación de explicación se cumple entre un conjunto de proposiciones A y un fenómeno e mediante la verificación de la posibilidad de una relación de relevancia explicativa entre fenómenos que son descritos por el *explanans* A (en la descripción de est_i^n) en el contexto teórico de Σ . Un par $\langle A, e \rangle$ puede ser evaluado desde Ex para determinar si sus elementos guardan entre ellos una relación de explicación, bajo algún sistema-RCS. Nótese también que la conexión entre el *explanans* A y el *explanandum* e está dada mediante una descripción de los elementos de alguna cadena de fenómenos. Finalmente, la noción de posibilidad está introducida para mantener una coherencia con los postulados que gobiernan los sistemas-RCS, desde fuera de estos sistemas.

Las condiciones que debe cumplir una explicación en Ex son extraepistémicas en el sentido de que necesitamos verificar la posibilidad de ciertos ítems que no necesariamente están explícitamente dentro del sistema epistémico. Un esquema argumentativo debería cumplir condiciones extraepistémicas en razón de calificar como una explicación.

Definición 29. Así, podemos identificar una explicación como perteneciendo a un sistema RCS $\Sigma = \langle \Lambda, \langle D, W \rangle \rangle$, como un RGS-*default* $((A, \omega), B) \in D$ tal que hay una relación de explicación $\langle A, e \rangle$ and ω es $D(e)$.

En esta forma, podemos distinguir tres importantes cosas en el modelo: **relación de pertinencia explicativa, relación de explicación y explicación**. las relaciones de pertinencia explicativa son relaciones particulares supuestas como cumpliéndose entre fenómenos. Una relación de explicación es una relación que se supone que se cumple entre un conjunto de proposiciones y un fenómeno particular. Las explicaciones son esquemas de argumento que cumplen condiciones extraepistémicas que involucran relaciones entre el conjunto de proposiciones y una descripción del fenómeno *explanandum*. Podemos identificar, desde Ex, explicaciones en un sistema-RCS particular Σ , por medio de una identificación de relaciones de explicación que se cumplen bajo el contexto de Σ .

También podemos representar una explicación por medio de un metaargumento, fuera de los sistemas epistémicos, constituido por los elementos que acabamos de clarificar.

Podemos observar algunas propiedades interesantes relativas a este marco teórico formal para representar las interacciones de explicaciones. Sean A, A' diferentes conjuntos de proposiciones; e un fenómeno; $[est]_i m, [est]_i n$ diferentes cadenas de fenómenos y Σ, Σ' diferentes sistemas RCS.

e1) Si hay una relación de explicación entre A y relativa a Σ , y pasa que A es verdadero (en el sentido de que todo elemento de A es verdadero), entonces esto apoya la verdad de $D(e)$ en el contexto de Σ .

e2) Si es el caso de que A es verdadero y $D(e)$ es verdadero en Σ , no necesariamente hay una relación de explicación entre A y e .

e3) El hecho de que A mantenga una relación de explicación con e en el contexto Σ , es independiente respecto de si A es verdadero o no.

e4) El hecho de que A mantenga una relación de explicación con e en el contexto Σ por medio de $[est]_i m$, no implica que extrasistemáticamente (esto es, fuera de Σ y A) la relación de pertinencia explicativa entre e y los fenómenos que hacen

verdadero A, realmente se mantenga en realidad, es decir no implica que la se mantenga una relación entre e y $[est]_i m$.

e5) Algunas veces $\langle A, D(e) \rangle$ podría ser un par *explanans-explanandum* en el contexto de Σ mientras que podría no serlo en el contexto de Σ' .

e6) Algunas veces $D(e)$ podría ser explicado por A pero también por A' , en el contexto del mismo Σ .

e7) Algunas veces $D(e)$ podría ser explicado mediante $[est]_i m$ pero también por medio de $[est]_i n$, en el contexto del mismo Σ .

e8) Algunas veces, si $\langle A, D(e) \rangle$ es un par *explanans-explanandum*, podría usar $[est]_i m$ como parte de la explicación pero también podría usar $[est]_i n$, en el contexto del mismo Σ .

e9) Algunas veces si $\langle A, D(e) \rangle$ es un par *explanans-explanandum* en el contexto de Σ , y la información de Σ crece, entonces, a pesar de que el prerequisite se cumple en algunas explicaciones que pertenecen a Σ , para ese par, podría pasar que el *proviso* no se cumpla más.

e10) Algunas veces, si $\langle A, D(e) \rangle$ es un par *explanans-explanandum* en un contexto Σ dado, cuando la información crece el prerequisite no mantiene ya más el vínculo con la relación de pertinencia explicativa (a partir de A no podemos inferir más $D(est; n)$), y, por lo tanto, $\langle A, D(e) \rangle$ ya no es más un par *explanans-explanation*. En este caso, $A/D(e)$ no es más una explicación.

e11) Algunas veces, si $\langle A, D(e) \rangle$ es un par *explanans-explanandum* en un contexto Σ dado, cuando la información crece, a pesar de que el prerequisite mantiene el vínculo con la relación de pertinencia explicativa, la cadena no es paraconsistentemente posible ya después del aumento de información y, por lo tanto, $\langle A, D(e) \rangle$ ya no puede ser un par *explanans-explanandum*. En este caso, $A/D(e)$ no es ya una explicación en ese contexto.

e12) Algunas veces en un contexto Σ dado, si $\langle A, D(e) \rangle$ es un par *explanans-explanandum*, cuando la información crece, podemos no inferir más, mediante $A/D(e)$ y la verdad de A , la descripción $D(e)$.

e13) Algunas veces en un contexto Σ dado, si $\langle A, D(e) \rangle$ no es un par *explanans-explanandum*, cuando la información crece, podemos inferir con $A/D(e)$ y la verdad de A , la descripción $D(e)$.

Estos importantes hechos muestran el carácter epistémico de este Marco Formal que modela la explicación, que integra herramientas formales de lógica paraconsistente y herramientas formales usadas en Inteligencia Artificial. Los hechos muestran también que pueden el marco formal puede contribuir a resolver algunos de los principales problemas planteados por la Filosofía de la Ciencia contra la tradición de intentar construir un modelo argumentativo de la explicación científica. Como hemos visto, el modelo está basado en dos clases de relación de pertinencia explicativa: la relación causal y la relación inferencial. Ambas relaciones han sido capturadas por una misma clase de estructura abstracta: la cadena de fenómenos. Pero la cuestión de si estas son relaciones adecuadas para definir la explanation, no es discutida aquí. De cualquier forma, esta cuestión, creemos, es independiente de las ventajas de clarificación que el modelo ofrece. Presenta una versión de cómo las relaciones entre explicaciones y otros diferentes aspectos de sus contextos teóricos puede ser entendido y esto, creemos, ya provee un valor epistémico para el modelo.

4.2 El Modelo y la Explicación.

Introducción

En la sección anterior presenté el marco teórico formal **Ex**. **Ex** es capaz de relacionar varios elementos. Principalmente, teorías, lógicas, *explananda*, *explanantes*, relaciones de pertinencia explicativa, explicaciones y pares *explanans-explanandum*. Todos estos elementos constituyen una noción particular de explicación.

La construcción del modelo comprendido por el marco formal **Ex** siguió directrices trazadas al principio de esta investigación. En el capítulo 1, discutí algunos problemas clásicos relacionados con la modelación de la explicación. La discusión expuesta ahí no dió como resultado una elaborada teoría sobre la explicación. Sin embargo sí sentó las bases para la construcción de una teoría de la explicación concebida sobre todo como una entidad epistémica y falible. El cambio de concepción propuesto eliminaría, según he dicho, algunos de los problemas básicos.

Presenté después dos tipos peculiares de casos de explicación. Los que contienen *explanans* inconsistentes y los que suponen una relación inferencial no monotónica. Cada tipo de caso planteó independientemente problemas para la modelación de la explicación aún bajo la nueva propuesta. Y la idea de conformar un modelo uniforme de explicación planteaba todavía otros problemas prácticos.

El modelo puede dar cuenta de muchos de los problemas básicos y prácticos planteados a lo largo de esta investigación. El modelo es una elucidación de la explicación en el sentido de que clarifica los elementos que la componen y su estructura, es decir, las relaciones entre dichos elementos entre sí y las relaciones entre estos elementos y algunos otros elementos externos. En la presente sección volveré a algunos de los

ejemplos paradigmáticos que hicieron problemática la construcción de un modelo formal de explicación basado en el supuesto argumental (SA) y que fueron presentados en el capítulo 2. A través del examen de algunos de estos ejemplos expondré dos cosas: el grado de adecuación y el funcionamiento formal del modelo presentado en la sección anterior. Dividiré la sección en 4 apartados: Pertinencia explicativa, Ambigüedad Epistémica y No Monotonidad, Paraconsistencia. Además, a lo largo de toda la sección pero sobre todo en el segundo de los apartados, expondré la forma en que mi modelo es un ejemplo de modelación de la explicación a través de argumentos.

4.2.1 Pertinencia Explicativa

4.2.1.1. Asimetrías

Expuse antes los casos de asimetría como casos de lo que llamé contrapertinencia, es decir, como casos de contraejemplo contra la idea de que la confirmación de una relación deductiva de derivación es suficiente para catalogar a un candidato como una explicación. Una reformulación de un ejemplo famoso de asimetría es el siguiente:

(*Marte*)

(*P*)

α) Todos los objetos en el firmamento emiten luz constante si y sólo si están cerca de la Tierra,

β) Marte está cerca de la Tierra.

χ) Marte emite luz constante.

(I)

α) Todos los objetos en el firmamento emiten luz constante si y sólo si están cerca de la Tierra,

χ) Marte emite luz constante.

β) Marte está cerca de la Tierra.

La asimetría produce la apariencia de que *Marte-P* es una explicación mientras que *Marte-I* no lo es. Como dije antes, la idea que subyace a esta distinción es que en el primer argumento está implicada alguna relación de pertinencia explicativa que está ausente en el segundo caso. La relación de pertinencia explicativa supuesta en los casos de asimetría es la relación de causalidad. Una manera de identificar los dos lados de una asimetría es, como afirmé, determinar claramente el tipo de eventos a los que pretende referirse con uno y otro argumento. Así, una vez identificado el tipo de evento, uno puede ver los casos de asimetría como casos en los que se mezclan dos tipos distintos de explicación: la explicación de eventos fácticos y la explicación de eventos no fácticos.

Una empresa interesante que debería poder lograrse mediante un modelo formal de explicación es lograr la capacidad de subsumir uno y otro lado de la asimetría en un mismo esquema formal uniforme para la explicación y, a la vez, discernir entre uno y otro lado de la asimetría. Esto implica la modelación de la ocurrencia de dos tipos de relación de pertinencia entre eventos. Como propuse antes, la causal y la inferencial, pueden dar cuenta de eventos fácticos y de no fácticos respectivamente. Capturar el papel que juegan estos dos tipos de relaciones de pertinencia es algo que también debería comprender el modelo formal.

Veamos cómo las reglas especiales del sistema **Ex** capturan el caso de asimetría *Marte*.

Marte-P es una explicación en **Ex** si y sólo si se entiende como un esquema $A/D(e)$, tal que cumple con condiciones específicas que relacionan a A y $D(e)$. Tales condiciones aparecen al final de la sección anterior. Las condiciones que relacionan a A con $D(e)$, pueden expresarse a través del robustecimiento del esquema $A/D(e)$ con las condiciones Ex1-Ex5 descritas en la sección anterior. Ese robustecimiento podría ponerse en términos de un esquema más complicado, como el siguiente:

(MP)

$$\frac{A1 \mid A1 \vdash_{E^\Sigma (\neg\text{vac})} D(\mathbf{est}_{in}) : \diamond_{E^\Sigma} D(e) \wedge [\mathbf{est}]_{in} \diamond_{E^\Sigma} (A), \diamond_{E^\Sigma} (D(e))}{D(e)}$$

donde:

- a) \vdash representa una relación derivacional deductiva o no deductiva,
- b) Σ es un sistema epistémico RCS,
- c) $A1 = \{\alpha, \beta\}$,
- d) Si χ describe un evento fáctico entonces $D(\mathbf{e})$ debe interpretarse como la descripción proposicional de éste, si χ describe un evento no fáctico entonces $D(\mathbf{e})$ debe interpretarse como describiendo la verdad de la proposición χ .

En tal caso el esquema MP dice que: Si se cumple $A1$ y se cumple que de $A1$ es no vacuamente Σ -derivable la descripción de un conjunto \mathbf{est}_{in} , tal que es paraconsistentemente posible 1) la descripción de que el evento \mathbf{e} es el último eslabón de la cadena $[\mathbf{est}]_{in}$, 2) A , y 3) la descripción del evento \mathbf{e} ; entonces infiérase plausiblemente la descripción de \mathbf{e} .

Un modo de resumir esta lectura de MP es: Si se cumple $A1$ y se cumple que de $A1$ es no vacuamente Σ -derivable la descripción de un conjunto \mathbf{est}_{in} , tal que es

paraconsistentemente posible la **descripción fuerte** de que el evento **e** es el último eslabón de la cadena $[\mathbf{est}]_{in}$, (lo cual incluye que 2) es paraconsistentemente posible A, y 3) es paraconsistentemente posible la descripción del evento **e**), entonces infiérase plausiblemente la descripción de **e**. Lo cual podría ponerse como el siguiente esquema:

$$\frac{A1 \mid A1 \vdash_{E^\Sigma (-vac)} D(\mathbf{est}_{in}) : \diamond_{E^\Sigma} DF(\mathbf{e}) \wedge ([\mathbf{est}]_{in})}{D(\mathbf{e})}$$

Así, *Marte-I* es una explicación en **Ex** si y sólo si se entiende como un esquema:

(MI)

$$\frac{A2 \mid A2 \vdash_{E^\Sigma (-vac)} D(\mathbf{est}_{im}) : \diamond_{E^\Sigma} DF(\mathbf{e}) \wedge ([\mathbf{est}]_{im})}{D(\mathbf{e})}$$

donde:

- a) \vdash representa una relación derivacional deductiva o no deductiva,
- b) Σ es un sistema epistémico RCS,
- c) $A2 = \{\alpha, \chi\}$,
- d) Si β describe un evento fáctico entonces $D(\mathbf{e})$ debe interpretarse como la descripción proposicional de éste, si β describe un evento no fáctico entonces $D(\mathbf{e})$ debe interpretarse como describiendo la verdad de la proposición β .

En tal caso el esquema MI dice que: Si se cumple A2 y se cumple que de A2 es no vacuamente Σ -derivable la descripción de un conjunto \mathbf{est}_{im} , tal que es paraconsistentemente posible la **descripción fuerte** de que el evento **e** es el último eslabón de la cadena $[\mathbf{est}]_{im}$, infiérase plausiblemente que $D(\mathbf{e})$.

En cualquier caso, ambos esquemas MP y MI son instancias de un esquema de regla en un sistema epistémico Σ . La respuesta acerca de qué lado de la asimetría es una explicación es dependiente principalmente del tipo de evento al que nos estemos

refiriendo al proponer cada uno de los argumentos *Marte-P* y *Marte-I*. La razón de identificar una asimetría es pensar que χ en MP y β en MI, ambos describen eventos fácticos. Así, en el caso en que β en MI es interpretado como describiendo un evento fáctico normalmente no habría ninguna red con la que el evento descrito por β pudiera establecer una conexión causal. Sin embargo, con mi modelo se ve claro que ocurriría igual en el caso de que χ fuera interpretado como describiendo un evento no fáctico, si no hubiera tampoco ninguna red con la que el evento descrito por χ pudiera establecer una conexión causal. El modelo que propongo, además, puede considerar la posibilidad de que *Marte-I* también sea una explicación, en el caso en que interpretamos a β como describiendo un evento no fáctico más bien que uno fáctico. En tal caso la conexión del evento con la cadena correspondiente tendría que ser, según la definición del predicado de conexión χ , de tipo inferencial.

Así, el modelo está basado en dos tipos de relación de pertinencia explicativa: la relación causal y la relación inferencial. Ambos tipos de relación de pertinencia han quedado subsumidos en un mismo tipo de estructura: la cadena o, en casos más específicos, una cadena causal o inferencial.

Una explicación es un argumento que hace plausible la inferencia de la descripción del evento que se quiere explicar, es decir, el *explanandum*. El argumento debe ser una instancia de esquemas como los robustecidos en los párrafos anteriores. La relación de pertinencia que captura mi modelo, en sus dos versiones, hace que la actividad científica de la explicación se distinga de una simple justificación. Esto se debe a que el tipo de argumento construido no supone que la relación de pertinencia esté basada sólo en una relación derivacional, como sucede en el caso de los modelos clásicos (los de Hempel u Oppenheim). Según mi modelo, los argumentos explicativos exigen además de una relación derivacional, que sea posible que la cadena, directamente relacionada con el *explanans*, dé cuenta de la ocurrencia del *explanandum* en cierto tipo de contexto. Este contenido adicional a la derivación está capturado en el modelo mediante el

cumplimiento de la relación de explicación entre el conjunto *explanans* y el evento a explicar, en la salvedad de los esquemas resultantes del robustecimiento del esquema que es candidato a explicación, digamos, A/D(e).

4.2.1.2 Explicaciones y Relaciones de Explicación

Tomando en consideración los ejemplos de las asimetrías, analizaremos brevemente la relación entre **explicaciones y relaciones de explicación**. Para que exista una relación de explicación basta que exista un vínculo especial entre un conjunto de proposiciones y un evento. El vínculo especial está dado por las condiciones Ex1-Ex5. Así podemos decir que:

A1 = { α , β } mantiene una relación de explicación con el evento **e1** descrito por χ en el sistema RCS Σ , en símbolos, que **RE(A1,e1) $_{\Sigma}$** , si y sólo si:

Hay una cadena [**est**] $_{in}$ tal que,

R1) O bien A1 $\vdash_{E^{\Sigma}(-vac)}$ **D(est) $_{in}$** o bien A1 $\vdash_{\sim E^{\Sigma}(-vac)}$ **D(est) $_{in}$** .

R2) $\nexists_{E^{\Sigma}} \mathbf{DF}(\mathbf{e}) \nexists$ [**est**] $_{in}$).

A2 = { α , χ } mantiene una relación de explicación con el evento **e2** descrito por β en el sistema RCS Σ , en símbolos, que **RE(A2,e2) $_{\Sigma}$** , si y sólo si::

Hay una cadena [**est**] $_{in}$ tal que,

R1) O bien A2 $\vdash_{E^{\Sigma}(-vac)}$ **D(est) $_{in}$** o bien A2 $\vdash_{\sim E^{\Sigma}(-vac)}$ **D(est) $_{in}$** .

R2) $\nexists_{E^{\Sigma}} \mathbf{DF}(\mathbf{e}) \nexists$ [**est**] $_{in}$).

Puede observarse que si un conjunto mantiene una relación de explicación con un evento, esto supone el hecho de que tal conjunto *postula* una relación de pertinencia explicativa entre el evento **e** y los contenidos en [**est**] $_{in}$. Esta relación de pertinencia tiene que ser coherente con el sistema Σ involucrado, debe ser paraconsistentemente posible en Σ .

Así, una explicación puede entenderse como un argumento por *default* $((A, \omega), B)$ cuya condición de ausencia exige el cumplimiento de una relación de pertinencia explicativa. Pero "cumplimiento" significa en este caso sólo una posibilidad de un tipo definido de conexión entre el evento-*explanandum* y una cadena. La cadena funciona como el vínculo entre el *explanans* y el *explanandum*. El *explanans* directo, en el contexto de un sistema epistémico complejo, hace que se pueda derivar la descripción del conjunto de elementos que la cadena supone. Así, a través de la cadena se establece la conexión que le da plausibilidad al *explanandum* dado el *explanans*. La plausibilidad está basada en la relación de pertinencia explicativa. Una explicación de **e** es un argumento que tiene por conclusión la descripción de un evento **e** (la conclusión es el *explanandum*) y cuyas condiciones antecedentes son dos: 1) que se cumpla una relación de explicación entre **e** y un conjunto de proposiciones (*explanans*), y 2), que sus condiciones de presencia y de ausencia se verifiquen.

Veamos cómo funciona en el modelo la relación de pertinencia explicativa con relación a la explicación. En el caso de una asimetría, tenemos dos versiones de explicación: *Marte-P* cuando el evento a explicar es un evento fáctico y *Marte-I* cuando el evento a explicar es un evento no fáctico. Bajo las otras interpretaciones del evento a explicar es posible que no hubiera explicación pero esto, extraepistémicamente hablando, es dependiente del contenido del sistema Σ que se tome en consideración.

Supongamos que *Marte-P* es una explicación tal que χ es un evento fáctico.

Si este es el caso entonces la cadena $[\mathbf{est}]_{in}$ es tal que hay una relación de explicación entre $\{\alpha, \beta\}$ y χ como sigue:

$$R1) \{\alpha, \beta\} \vdash_{\Sigma(-vac)} \mathbf{D}(\mathbf{est}_{in}) \text{ o bien } \{\alpha, \beta\} \vdash_{E^{\Sigma}(-vac)} \mathbf{D}(\mathbf{est}_{in}).$$

$$R2) \diamond_{E^{\Sigma}} \mathbf{DF}(\mathbf{e}) \nexists [\mathbf{est}]_{in}.$$

Supongamos que R1 se cumple. Dado que se trata de un evento a explicar que es fáctico, tenemos los siguientes dos casos generales para R2:

- i) Cuando $[\mathbf{est}]_{in}$ no es una cadena causal.
- ii) Cuando $[\mathbf{est}]_{in}$ es una cadena causal.

En el caso i, como hemos supuesto que hay una explicación y por tanto una relación de explicación, el conjunto \mathbf{est}_{in} puede ser, por ejemplo, el evento de la cercanía de Marte respecto de la Tierra o el evento de que Marte refleja la luz de una estrella cercana a la Tierra, a saber, el Sol, o el par de eventos anteriores, o el evento de que Marte está cerca de la Tierra y refleja la luz del Sol.

En el caso ii, por la misma razón que en el caso anterior, la el conjunto \mathbf{est}_{in} puede ser, por ejemplo, el evento de la cercanía de Marte respecto de la Tierra junto con el evento de que Marte refleja la luz del Sol causan que la luz reflejada por Marte recorra en poco tiempo la distancia entre Marte y la Tierra.

Ahora supongamos que *Marte-I* es una explicación tal que β es un evento no fáctico.

Si este es el caso entonces la cadena $\diamond_{E^2} \mathbf{DF}(e \nrightarrow [\mathbf{est}]_{in})$.es tal que hay una relación de explicación entre $\{\alpha, \chi\}$ y β como sigue:

$$R1') \{\alpha, \chi\} \vdash_{\Sigma(-vac)} \mathbf{D}(\mathbf{est}_{in}) \text{ o bien } \{\alpha, \chi\} \vdash_{\sim E^2(-vac)} \mathbf{D}(\mathbf{est}_{in}).$$

$$R2') \diamond_{E^2} \mathbf{DF}(e \nrightarrow [\mathbf{est}]_{in}).$$

Supongamos que R1' se cumple. Dado que se trata de un evento a explicar que es no fáctico, tenemos los siguientes dos casos generales para R2':

- i) Cuando $[\mathbf{est}]_{in}$ no es una cadena causal.
- ii) Cuando $[\mathbf{est}]_{in}$ es una cadena causal.

En el caso i, por ejemplo, $[\text{est}]_{\text{in}}$ puede ser la proposición que describe el evento de que Marte emite luz constante, o la proposición que describe el evento de que Marte refleja la luz de una estrella cercana a la Tierra, a saber, el Sol, o ambas proposiciones anteriores, o la proposición que describe el evento de que Marte luz constante y refleja la luz del Sol o cada una de estas opciones junto con la proposición α .

En el caso ii, por ejemplo, $[\text{est}]_{\text{in}}$ puede la cadena inferencial abierta: De la proposición χ se infiere que la luz reflejada por Marte recorra en poco tiempo la distancia entre Marte y la Tierra.

4.2.1.3 Causalidad

La relación de pertinencia explicativa para eventos fácticos, la relación causal, tenía que enfrentar aún dos problemas más expuestos en el capítulo 2: el problema de la sobredeterminación epistémica causal y el problema de la elección causal. En el caso de la sobredeterminación epistémica causal el problema es el siguiente: hay casos de explicación en los que el *explanans* es verdadero, el *explanans* describe factores causales adecuados y la derivación es correcta y, a pesar de toso esto, el *explanans* no describe *los* factores que realmente produjeron el *explanandum*. En el caso de la elección causal el problema es: hay casos en los que un mismo *explanandum* es el resultado de varias cadenas causales que pueden ser todas verdaderas y que responden a distintos aspectos sobre el mismo evento.

La forma en que yo abordé ambos problemas fue abandonando lo que llamé "el ideal de buena explicación". Para el caso de la sobredeterminación epistémica causal respondí aseverando que para que algo sea una explicación no debería ser necesario que el *explanans* describa los factores causales que realmente produjeron el *explanandum*. Para el caso de la elección causal dije que para que algo sea una explicación no debería ser necesario que los factores causales que describe el *explanans* se ajusten completamente a

las expectativas del contexto en el que se exige la explicación. Esta idea resultó en la conformación de una concepción epistémica, pragmática y mínima de la explicación. El modelo de la sección anterior captura una parte necesaria de las buenas explicaciones, el núcleo básico de toda explicación independientemente de las restricciones externas (como las que supone la identificación de la causa real) e independientemente de algunas restricciones internas (como las que supone la identificación de la causa adecuada al contexto en el que se exige la explicación). En esta concepción que he estado construyendo a lo largo de la investigación, una explicación puede no corresponder a los hechos reales involucrados con la ocurrencia del evento-*explanandum* y puede no ser completamente adecuado al contexto en el que se exigió tal explicación.

Sin embargo, la relación de pertinencia no puede ser completamente irrestricta. Esto sería como darle el nombre de "relación inferencial" y el de "relación causal" a cualquier cosa. No queremos una relación de pertinencia que no tenga ningún cánón de adecuación.

En el capítulo 1 me referí a un tipo de adecuación contextual que difiere significativamente de la adecuación contextual *à la* van Fraassen. La adecuación que exige mi modelo es una adecuación que llamé en esa sección "adecuación teórica". Ahora, con elementos formales más precisos, debe aclararse que se trata de una adecuación respecto de una teoría y una lógica subyacente a dicha teoría. Como hemos visto, la adecuación es respecto de un sistema epistémico complejo como los que he descrito en la sección anterior.

En los ejemplos anteriores, la explicación en mi modelo postula una relación de pertinencia explicativa entre eventos, pero no supone que dicha relación ocurra realmente. Tan sólo tenemos que hay una descripción de la cadena en el sistema Σ más el *explanans* A (en $\Sigma+A$) y que a esta cadena puede estar conectado como último eslabón el evento-*explanandum*. Pero no sabemos si esto corresponde o no con alguna realidad

fuera del ámbito de Σ . Esto refleja el carácter epistémico de la concepción de explicación que el modelo supone. Algo es una explicación siempre en relación a un sistema de conocimiento que funciona como contexto. Σ es el punto de referencia para checar la adecuación de la relación de pertinencia postulada en la explicación.

En el caso de la relación inferencial, la adecuación depende completamente del sistema epistémico complejo. Es posible que dos proposiciones mantengan una relación inferencial sólo si es derivable bajo el sistema epistémico complejo en cuestión. Aquí tenemos una relativización completa a los sistemas epistémicos. Sin embargo, el límite a cualquier tipo de inferencia está dado desde que los sistemas epistémicos que se asocian a las explicaciones son sistemas RCS. Todo sistema RCS usa una lógica de da Costa y todo sistema asociado a una explicación es un sistema RCS. Así, el significado de "relación inferencial" es aproximadamente uniforme en todo sistema.

En el caso de la relación causal es diferente. El problema es que una adecuación completamente relativizada al sistema epistémico complejo, en este caso, podría dar lugar a la trivialización de la noción de relación causal. Esto se debe a que no hay necesariamente restricciones intensionales para el significado de una relación causal dentro de un sistema RCS. Cada sistema Σ podría llamar "relación causal" a absolutamente cualquier tipo de relación. La restricción aquí viene intensionalmente desde fuera del sistema y extensionalmente desde dentro del sistema. Similarmente a como dije antes, la solución para restringir moderadamente nuestra relación de pertinencia explicativa para eventos fácticos es mantener una definición intensional fija para todo sistema Σ asociado a una explicación mientras dejamos que la extensionalidad de la relación causal varíe de sistema a sistema. El rol de la definición intensional fija está representado por la interpretación que demos a un predicado primitivo **C** que puede incluirse en el marco teórico **Ex**. Esta restricción intensional permite concebir de una única forma la relación de pertinencia para eventos fácticos mientras que permite, a la vez, flexibilidad en la concepción que tenga cada sistema epistémico complejo Σ respecto de

los tipos de pares de eventos que pueden mantener una relación **C** como la estipulada. Permite distintas postulaciones de lo que puede relacionarse causalmente y no permite que todo tipo de relación sea pertinente causalmente.

Ahora bien, deberíamos poder extraer, de formulaciones que son instancias de EC algunas otras formulaciones. Por ejemplo, deberíamos poder obtener en una teoría que asume la existencia de relaciones causales, del dato de que, digamos, **C(| α |, | β |)** y la afirmación α , la afirmación β . Es decir, deberíamos poseer una serie de reglas que permitan transformaciones entre fórmulas que involucran causalidad. Por supuesto, este conjunto de reglas define al menos parcialmente la noción de causalidad. Como he propuesto que la definición de causalidad sea uniforme en todo contexto teórico, debemos suponer que todo sistema epistémico complejo posee exactamente las mismas reglas que tienen que ver con causalidad. No me detendré a examinar cuál sería un conjunto adecuado de reglas para modelar la relación de causalidad pues esto excedería los límites de mi investigación, como he dicho antes, la definición de la relación causal no formará parte de este trabajo.

Las reglas causales pueden representarse adecuadamente mediante esquemas de regla *default* que estén contenidos en la teoría del sistema epistémico complejo RCS. Un ejemplo de regla causal obvia que se podría asumir como correcta en todo sistema epistémico complejo es la siguiente:

$$\underline{Ax : \Diamond[Bx \wedge C(|Ax|, |Bx|)]}$$

$$Bx \wedge C(|Ax|, |Bx|)$$

Este tipo de esquema nos permitiría pasar de la afirmación de un evento $|Ax|$ y la posibilidad de una conexión causal **C(| Ax |, | Bx |)** de este evento $|Ax|$ con otro evento $|Bx|$, a la afirmación de la relación causal y el evento $|Bx|$. El esquema sería

particularmente útil para permitir que una descripción de una cadena causal pueda derivarse a partir de un *explanans*.

4.2.2 Ambigüedad Epistémica y No Monotonicidad.

Hempel consideró la ambigüedad epistémica como un problema, debido a dos consecuencias de las inconsistencias inductivas: la debilitación de la noción de derivación de un argumento explicativo y la dificultad para tratar con conclusiones mutuamente contradictorias en un mismo conjunto de datos.

El problema de la ambigüedad epistémica constituyó un fuerte argumento contra la idea de que podemos modelar la explicación con base en argumentos de cierto tipo. Otro argumento viene de la idea de Salmon de que no es posible establecer criterios fundados para los límites probabilísticos que admitiría la relación de derivación en un argumento probabilístico. Básicamente, la idea que yo propuse, coherentemente con la idea de un modelo mínimo y epistémico de explicación, fue abandonar el supuesto de alto grado de certeza para la conclusión. Admitir consecuentemente argumentos sin restricciones de grados de certeza, una idea que defendí en mi tesis de maestría. Como hemos visto en el apartado anterior, esto no significa que los argumentos sean completamente irrestrictos. Queremos un grado de contextualismo, no relativismo. En el caso de mi modelo, la relación de pertinencia explicativa controla la relación de derivación.

La idea de concebir la explicación como una entidad epistémica no surge sólo de las ventajas que ofrece esta concepción para abordar los problemas de las asimetrías y, en general, los problemas relacionados con la causalidad. También surge de sus ventajas en el tratamiento de los problemas relacionados con la ambigüedad epistémica.

En esta concepción epistémica es básico el abandono del ideal de buena explicación y el abandono del supuesto de alto grado de certeza. Esto permite pensar en un tipo de explicación relacionada con una noción de derivación débil. Permite pensar en la explicación como un argumento falible. Por otro lado, facilita la introducción en el modelo de elementos contextuales que juegan un papel muy importante en la explicación, esto es, la relativización a ciertos cuerpos de conocimiento, como se desarrolló en el capítulo 3 de esta investigación. Dicha relativización, a la vez que incorpora en el modelo el papel del contexto, hace posible, como hemos visto en el apartado anterior, una solución del problema de las relaciones de pertinencia explicativa.

La concepción epistémica, pragmática y mínima de la explicación nos capacita para observar con precisión tres aspectos muy interesantes: la consideración de A) el *explanans* como algo postulado, B) el *explanandum* falible y C) la explicación como algo dinámico. Modelar estos aspectos es lo que he llamado "representación de la ignorancia". La representación de la ignorancia supone una concepción no deductivista de la inferencia. La concepción no deductivista que asumo propone una noción de derivación que es contraria al supuesto de alto grado de certeza y que tiene tres características básicas: 1) la capacidad de saltar a la conclusión, 2) la plausibilidad basada en las relaciones de pertinencia explicativa, y 3), la propiedad de no monotonicidad.

Los aspectos A-C de la representación de la ignorancia incluyen las características 1-3 de la posición no deductivista supuesta por mi concepción epistémica de la explicación. En este apartado mostraré cómo funciona mi modelo en los tres aspectos de la representación de la ignorancia. En el aspecto C tan sólo completaré lo ya expuesto sobre el salto a la conclusión y sobre la no monotonicidad.

4.2.3 Representación de la Ignorancia

4.2.3.1. Explanans Postulado

Recordemos un ejemplo de Achinstein: *Arsénico*. En este ejemplo Achinstein supone, fatalistamente, que Juan ha ingerido arsénico pero que muere en un accidente automovilístico de camino al hospital. En una interpretación normal, la explicación de la muerte de Juan es la explicación de un evento fáctico.

Supongamos que χ es el evento de que Juan murió. Tal como Achinstein nos presenta el ejemplo, el *explanans* puede estar constituido de dos proposiciones:

α : "Juan comió una cucharada de arsénico"

β : "Cualquiera que tome una cucharada de arsénico se muere"

Desde mi perspectiva $\{\alpha, \beta\}$ explica χ respecto de un conjunto de conocimiento Σ si se cumple que existe una cadena $[\mathbf{est}]_{in}$ tal que hay una relación de explicación entre $\{\alpha, \beta\}$ y χ como sigue:

$$R1) \{\alpha, \beta\} \vdash_{\Sigma(-vac)} \mathbf{D}(\mathbf{est}_{in}).$$

$$R2) \diamond_{E^{\Sigma}} \mathbf{DF}(\chi) \nexists [\mathbf{est}]_{in}.$$

donde:

\vdash representa una relación derivacional deductiva o no deductiva,

Σ es un sistema epistémico RCS.

Si χ es considerado un evento fáctico, el hecho de que la proposición χ se pueda inferir de $\{\alpha, \beta\}$ no basta para explicar χ . Lo que importa en un caso como éste es que se pueda establecer una red que nos lleve causalmente al evento *explanandum* χ , como bien indica la condición R2. R2 refiere a una clase de compatibilidad entre los elementos de la

cadena y el evento *explanandum*. De modo que el punto de referencia para checar la posibilidad paraconsistente contenida en R2 es la derivación de la que habla R1.

Para el *explanans* como postulado, hay dos aspectos importantes que observar: la falibilidad del *explanans* y la inadecuación del *explanans* (que falle al explicar). Primero daré algunos ejemplos introductorios y después pasaré a los que ilustran estos dos aspectos importantes. Construyamos entonces algunos ejemplos basados en el ejemplo de Achinstein. Consideremos las siguientes proposiciones:

Cx: "x come una cucharada de arsénico",

Tx: "El torrente sanguíneo de x se contamina de arsénico",

C(|Cx|, |Tx|): "El que x coma una cucharada de arsénico causa que el torrente sanguíneo de x se contamine de arsénico",

Px: "x tiene arsénico acumulado en sus pulmones",

C(|Tx|, |Px|): "El que el torrente sanguíneo de x esté contaminado de arsénico causa que x tenga arsénico acumulado en sus pulmones",

Bx: "Ocurre una combinación de arsénico y oxígeno en x",

Fx: "x presenta asfixia",

Mx: "x muere",

C(|Cx|, |Mx|): "El que x coma una cucharada de arsénico causa que x muera",

Ax: "x tuvo un accidente automovilístico",

C(|Ax|, |Mx|): "El que x tenga un accidente automovilístico causa que x muera".

Y consideremos los siguientes *defaults*:

(d1)

$Px : \diamond[Bx \wedge C(|Px|, |Bx|)]$

$Bx \wedge C(|Px|, |Bx|)$

(d2)

$$\underline{Bx : \diamond[Fx \wedge C(|Bx|, |Fx|)]}$$

$$Fx \wedge C(|Bx|, |Fx|)$$

(d3)

$$\underline{Fx : \diamond[Mx \wedge C(|Fx|, |Mx|)]}$$

$$Mx \wedge C(|Fx|, |Mx|)$$

(d4)

$$\underline{Cx : \diamond[Mx \wedge C(|Cx|, |Mx|)]}$$

$$Mx \wedge C(|Cx|, |Mx|)$$

(d5)

$$\underline{Ax : \diamond[Mx \wedge C(|Ax|, |Mx|)]}$$

$$Mx \wedge C(|Ax|, |Mx|)$$

Ejemplo *Arsénico 1*

Supongamos que:

$$\Sigma 1 = \langle \Lambda 1, \Delta 1 \rangle,$$

$$\Delta 1 = \langle D1, W1 \rangle,$$

$$D1 = \emptyset,$$

$$W1 = \emptyset,$$

$$\alpha = (\forall x)(Cx \supset Mx),$$

$$\beta = Cj = \text{"Juan comió una cucharada de arsénico"},$$

En este caso tenemos que si $|\chi|$ es el evento fáctico de que Juan murió, entonces $\{\alpha, \beta\}$ explica $|\chi|$ respecto de $\Sigma 1$. Es decir, $\mathbf{RE}(\{\alpha, \beta\}|\chi|)_{\Sigma 1}$. La razón es que se puede derivar deductiva y no vacuamente una descripción de Cj y dado que $W1$ es

vacío entonces los supuestos causales de Σ_1 son compatibles con alguna cadena causal que vincule C_j con M_j .

Ejemplo *Arsénico 2*

Supongamos que:

$$\Sigma_2 = \langle \Lambda_2, \Delta_2 \rangle,$$

$$\Delta_2 = \langle D_2, W_2 \rangle,$$

$$D_2 = \emptyset,$$

$$W_2 = \{ \text{"Morir no puede tener más que una sola causa: la voluntad de Dios"} \},$$

$$\alpha = (\forall x)(Cx \supset Mx),$$

$$\beta = C_j = \text{"Juan comió una cucharada de arsénico"},$$

En este caso tenemos que a pesar de que se puede inferir χ (= M_j) del *explanans* $\{\alpha, \beta\}$ y de que también podemos obtener la descripción de una red constituida sólo por C_j , $\{\alpha, \beta\}$ no explica $|\chi|$ respecto de Σ_2 . Es decir, no ocurre que **REU**($\{\alpha, \beta\}, |\chi|$) $_{\Sigma_2}$. La razón es que no podemos establecer ninguna conexión causal entre C_j y M_j dado el contenido de los supuestos causales de Σ_2 . Aunque podemos cumplir con R1 no podemos cumplir con R2. Dicho del modo que nos interesa destacar: la capacidad de un conjunto de proposiciones para derivar el *explanandum* no es suficiente para que ese conjunto sea considerado un *explanans* del *explanandum*.

Ejemplo *Arsénico 3*

Supongamos:

$$\Sigma_1 = \langle \Lambda_1, \Delta_1 \rangle,$$

$$\Delta_1 = \langle D_1, W_1 \rangle,$$

$$D_1 = \emptyset,$$

$$W_1 = \emptyset,$$

$$\phi = (\forall x)(Ax \supset Mx),$$

$\gamma = A_j = \text{"Juan tuvo un accidente automovilístico"}$,

En este caso, como en *Arsénico 1*, tenemos que $\{\phi, \gamma\}$ explica $|\chi|$ respecto de Σ_1 . Es decir, $\mathbf{RE}(\{\phi, \gamma\}, |\chi|)_{\Sigma_1}$. Ambas formas de explicar la muerte de Juan son admitidas como explicaciones en mi modelo y, como puede notarse, esto puede suceder respecto de un mismo sistema epistémico complejo. Esto es una consecuencia del abandono del ideal de buena explicación. Una consecuencia adicional de esta perspectiva es que si nuestro candidato a *explanans* fuera $\{\alpha, \beta, \phi, \gamma\}$, tendríamos dos explicaciones a partir de un mismo *explanans* y un mismo sistema epistémico complejo

Ejemplo *Arsénico 4*

Supongamos que:

$$\Sigma_4 = \langle \Delta_1, \Delta_4 \rangle,$$

$$\Delta_4 = \langle D_4, W_4 \rangle,$$

$$D_4 = \{d_1, d_2, d_3\}$$

$$W_4 = \{(\forall x)[C_x \supset (\mathbf{C}(|C_x|, |T_x|) \wedge T_x)], (\forall x)[T_x \supset (\mathbf{C}(|T_x|, |P_x|) \wedge P_x)]\}$$

$$\beta = C_j$$

Este caso representa una explicación dependiente de una derivación no deductiva de una cadena de varios factores causales y cuyo *explanans* se constituye de una sola proposición, a saber, la proposición singular β . Aquí se cumple que $\mathbf{RE}(\{\beta\}, |\chi|)_{\Sigma_4}$.

La única extensión de Σ_4 es:

$$E^{\Sigma_4} = C_n^{\wedge 1}(W)$$

La única extensión de $\Sigma_4 + \{C_j\}$ es:

$$E^{\Sigma_4 + \{C_j\}} = C_n^{\wedge 1}[W \cup \{C_j\} \cup \{[\mathbf{C}(|C_j|, |T_j|) \wedge T_j], [\mathbf{C}(|T_j|, |P_j|) \wedge P_j], [B_j \wedge \mathbf{C}(|P_j|, |B_j|)], [F_j \wedge \mathbf{C}(|B_j|, |F_j|)], [M_j \wedge \mathbf{C}(|F_j|, |M_j|)]\}]$$

De aquí que $\{C_j\} \vdash_{\Sigma_4(-vac)} [\mathbf{C}(|C_j|, |T_j|) \wedge \mathbf{C}(|T_j|, |P_j|) \wedge \mathbf{C}(|P_j|, |B_j|) \wedge \mathbf{C}(|B_j|, |F_j|)]$,

Si designamos $\mathbf{C}(|Cj|, |Tj|) \wedge \mathbf{C}(|Tj|, |Pj|) \wedge \mathbf{C}(|Pj|, |Bj|) \wedge \mathbf{C}(|Bj|, |Fj|)$ con $\mathbf{D}([\mathbf{est}]_{in})$, podemos abreviar como sigue:

$$\{Cj\} |_{\sim \Sigma 4(-vac)} \mathbf{D}(\mathbf{est}_{in}).$$

Tenemos así que una descripción de una cadena, específicamente, una descripción de elementos de una cadena causal, se deriva del *explanans* en el contexto del sistema epistémico complejo $\Sigma 4$.

Obsérvese que la conjetura del *default* 3, $[Mj \wedge \mathbf{C}(|Fj|, |Mj|)]$, sí pertenece a la extensión pero no aparece en la descripción que nos interesa que se derive. Por supuesto que también es posible derivar una descripción donde aparezca como uno de los conjuntos, pero no es lo que queremos. La razón es que para una explicación nos basta con que podamos conectar, en el sentido del predicado, la cadena descrita con el evento que se describe en el *explanandum*. Nos interesa que la cadena termine con $\mathbf{C}(|Bj|, |Fj|)$ y que se cumpla que $\diamond_{E^2} \mathbf{DF}(|Mj|) \nrightarrow [\mathbf{est}]_{in}$, (es decir, R2).

De este modo, es claro que si d3 no perteneciera D4 en el sistema epistémico complejo $\Sigma 4$, de cualquier forma se cumpliría que $\mathbf{RE}(\{\beta\}, |\chi|)_{\Sigma 4}$. Esto se debe a que para una misma extensión E^k tenemos que $\diamond_{E^2} \mathbf{DF}(|Mj|) \nrightarrow [\mathbf{est}]_4$ y $\{Cj\} |_{\sim \Sigma 4(-vac)} \mathbf{D}(\mathbf{est}_4)$. Esta extensión común E^k es la extensión $E^{\Sigma 4 + \{Cj\}}$. Hay una cadena $[\mathbf{est}]_{4c}$ que es una continuación de la cadena $[\mathbf{est}]_4$, tal que el último tramo de $[\mathbf{est}]_{4c}$ es $\mathbf{C}(|Fj|, |Mj|)$. La descripción de $[\mathbf{est}]_{4c}$ es así la fórmula que resulta de añadir como conjunto $\mathbf{C}(|Fj|, |Mj|)$ a $\mathbf{D}(\mathbf{est}_4)$.

Nótese por último que no es necesario que la derivación de $\mathbf{D}(\mathbf{est}_4)$ sea no deductiva.

Ahora bien, todo este análisis ha tenido como objetivo básico, familiarizarnos con la forma en que el modelo captura la explicación pero también anticipar la perspectiva epistémica con que se encuentra comprometido. Este último punto es el que me propongo profundizar en este apartado, que puede traducirse como la representación de la ignorancia que produce el modelo.

Pensemos de nuevo en el *explanans* como algo postulado. Como dije antes de dar los ejemplos *Arsénico 1-4*, hay dos cosas por tratar: la falibilidad del *explanans* y el hecho de que el *explanans* falle en explicar al *explanandum*. Aprovecharé el ejemplo anterior para tratar primero el segundo aspecto: lo que he llamado también "la inadecuación del *explanans*".

Normalmente, se pueden identificar al menos dos clases de inadecuación. Cada una de ellas corresponde a una forma de compromiso con el ideal de buena explicación. En mi modelo la única inadecuación real es que la red no corresponda al tipo de eventos que podrían causar al evento-*explanandum*. El ejemplo *Arsénico 2*, es un claro ejemplo de inadecuación. Cualquier evento que propongamos en ese caso, no puede explicar la muerte de Juan, dados los supuestos causales del contexto teórico.

De este modo, las dos clases de inadecuación a las que me refiero lo son desde un punto de vista de la explicación que es distinto al mío. Mi modelo captura estos casos como explicaciones adecuadas aunque pudieran, en un análisis más complejo, ser calificadas de malas explicaciones. La primera clase es cuando el *explanans* falla al explicar al *explanandum*, en relación al contexto en que se exige la explicación. Esto es lo que vimos como el problema de la elección causal pero también vale para el problema de la sobredeterminación epistémica causal. Tal contexto no está incluido en mi modelo, tan sólo lo está parcialmente mediante la identificación del tipo de *explanandum* al que se refiere la explicación. El caso paradigmático puede verse con el ejemplo *Arsénico*. Podemos dar muchas explicaciones distintas, apelando a distintas redes causales en una misma trama causal. Por ejemplo podemos decir que Juan murió por haber comido arsénico pues fue esto lo que provocó que saliera descontrolado en el auto; o bien podemos decir que Juan murió por un accidente automovilístico, por ser la causa más inmediata; o bien podemos decir que la causa es lo mal diseñada que está la carretera por la que iba Juan, etc. Al decidir abandonar el ideal de buena explicación, admití que cada

una de estos candidatos a explicación tienen igual posibilidad de ser aceptados como explicaciones de la muerte de Juan. Esto se muestra claramente en una modificación del caso *Arsénico 4*.

Puede haber distintas redes propuestas por un mismo *explanans*. Algunas de ellas pueden ser más cortas que otras. Por ejemplo, en el ejemplo *Arsénico 4* podríamos omitir todos los tramos intermedios de la cadena y dejar únicamente los extremos de la cadena. Es decir, el *explanans* podría estar proponiendo que el evento de que Juan comió arsénico, C_j , causa la muerte de Juan, M_j . O que el evento de que Juan tuvo un accidente, A_j , causa la muerte de Juan, M_j . Si agregamos a $\Sigma 4$ los *defaults* d_4 y d_5 y expandemos el *explanans* agregándole la proposición A_j , podemos tener, además de la cadena causal analizada arriba, otras dos cadenas. Una de ellas $C(|C_j|, |M_j|)$ y la otra, $C(|A_j|, |M_j|)$. Estas constituyen explicaciones más simples que la que acabamos de analizar. De nuevo, bastará con que el *explanans* ayude a derivar la descripción de la causa propuesta en cada caso, en un caso la descripción de $|C_j|$ y en otro la descripción de $|A_j|$. Así, si tuviéramos un sistema complejo como $\Sigma 4$ más d_4 , d_5 , este sistema junto con el *explanans* $\{C_j, A_j\}$, podría proponer tres cadenas distintas para explicar $|M_j|$. Estas constituirían tres distintas explicaciones para $|M_j|$. Este mismo tratamiento puede darse al contraejemplo de Scriven, *Puente*, que vimos en el capítulo 2, con el que expliqué el problema de la sobredeterminación epistémica causal.

Observemos también que si agregamos conexiones causales de $|C_j|$ a $|A_j|$, podríamos tener una explicación cuya cadena causal es una fusión de dos cadenas causales que pueden ser tratadas independientemente: una que va de $|C_j|$ a $|M_j|$ y la que va de $|A_j|$ a $|M_j|$. La cadena resultante sería una que fuera de $|C_j|$ pasando por $|A_j|$ hasta terminar con $|M_j|$.

En el caso anterior la aparente inadecuación consiste en que la explicación no responde al contexto en que se exige la explicación. La segunda clase de aparente

inadecuación es cuando la explicación supone cadenas que pueden derivarse pero que no tienen relación alguna con los hechos reales. Esta segunda posibilidad de aparente inadecuación es en realidad una consecuencia interesante de mi modelo: cualquier explicación posible respecto de un contexto teórico es considerada una explicación. Más allá de si la trama causal puede interpretarse de varias maneras, puede darse el caso de que haya ramificaciones posibles de la trama central que muestren distintas posibilidades para la explicación del *explanandum* independientemente de si tienen o no algo que ver con la realidad. Esto incluye la sobredeterminación epistémica causal que expuse antes pero también otra forma particular de sobredeterminación epistémica causal que se distingue de la anterior en que el factor aparentemente no pertinente (como la bomba en el caso *Puente de Scriven*) no se establece como dato original sino que se deriva de los datos originales. Por ejemplo si agregáramos a $\Sigma 4$ en el caso *Arsénico 4*, el siguiente *default*:

$$\underline{T_x : \Diamond[G_x \wedge C(|T_x|, |G_x|)]}$$

$$G_x \wedge C(|T_x|, |G_x|)$$

donde T_x tiene el significado estipulado antes y G_x significa que x tiene una gastritis aguda.

Podemos imaginar que, naturalmente, el sistema epistémico complejo tuviera un *default* que nos diera una relación causal entre, digamos, $|T_j|$, y el evento de que Juan tenga gastritis aguda, $|G_j|$, en tal caso podríamos tener que, dado nuestros supuestos causales, podemos conectar plausiblemente $|G_j|$ con $|M_j|$ pues es posible en los límites del sistema que el tipo de evento de $|G_j|$ cause tipos de eventos como el de $|M_j|$. En tal caso tendríamos una explicación completamente ajena a los sucesos reales pero una explicación a fin de cuentas. Es interesante notar que esto no tiene que ver con el hecho de que el *explanans* sea verdadero o falso. En el caso *Arsénico 4* el *explanans* es $\{C_j\}$, C_j da lugar a T_j y T_j da lugar a G_j . Lo cual hace que de $\{C_j\}$ pueda derivarse la descripción $C(|C_j|, |T_j|) \wedge C(|T_j|, |G_j|)$ y que esto posibilite la explicación de M_j a través de $|C_j|$, $|T_j|$ y $|G_j|$.

Cj puede ser perfectamente verdadero pero Tj y Gj sólo tienen certeza transitoria. No obstante, desde mi perspectiva epistémica de la explicación esto es una explicación. Aunque pueda ser considerada una mala explicación después de haberle hecho la autopsia a Juan.

Así, el modelo de explicación que propongo es epistémico en este primer sentido, no tiene que cumplir completamente con las expectativas del contexto ni en el sentido de el tipo de explicación que se exige (primera aparente inadecuación), ni en el sentido de que todo factor propuesto en la explicación corresponda con los hechos reales (segunda aparente inadecuación). Tratemos ahora el primer caso del *explanans* como algo postulado: la falibilidad del *explanans*. Para ello seguiré basándome en el ejemplo de Achinstein.

Supongamos una situación maquiavélica. Pensemos que Roberto, el hermano de Juan, quien quería deshacerse de éste, hizo que Juan confundiera el azúcar con el arsénico y, queriendo ocultar su crimen, cuando Juan muere Roberto construye la evidencia del accidente automovilístico. Coloca el cuerpo de Juan en el auto y hace que éste caiga por un acantilado. En tal caso no es verdad Aj. En mi modelo, aún en este caso se cumple que $RE(\{Aj\}, |Mj|)_{\Sigma 4}$. Es decir, bajo el contexto teórico de $\Sigma 4$ el *explanans* {Aj} da cuenta de la muerte de Juan. Lo que bien se podría decir es que Aj no da cuenta del arsénico acumulado en los pulmones de Juan. Pero esto no hace que {Aj} deje de explicar a |Mj| respecto de $\Sigma 4$. El conjunto {Aj}, y |Mj|, están relacionados de modo que conforman una relación de explicación.

Suponiendo que Aj esté en el conjunto de fórmulas del sistema **Ex Arsénico-Aj** cumple con una relación de explicación. La continuación que hace que $\diamond_{\Sigma 4} DF(|Mj| \nrightarrow [est]_m)$ se cumpla tiene como último tramo a $C(|Aj|, |Mj|)$. $D(|Mj|)$ es la afirmación Mj. No necesitamos que pueda derivarse en el sistema $\Sigma 4$ ya el *explanans* directo {Aj}, ni que $C(|Aj|, |Mj|)$ ni que pueda derivarse que Mj. Pero tampoco es suficiente que sea posible

la conexión causal $\mathbf{C}(|Aj|, |Mj|)$. Hace falta esta posibilidad, que se cumpla la relación derivacional. La conexión es posible dado que en W de Σ_4 no tenemos restricciones para la causalidad. La derivación se cumple pues podemos derivar de $\{Aj\}$ una descripción de $|Aj|$, a saber, Aj . Esta extensión es a la que están referidas tanto la posibilidad como la derivación aludidas. Así, a pesar de que Aj es falso, se cumple que $\{Aj\} \mid_{\sim\Sigma_4(-vac)} \mathbf{D}([\mathbf{est}]_m)$. Lo que no se cumple es que sea verdadero Aj y esto hará que sea imposible derivar mediante el esquema *Arsénico-Aj*, que Mj . Todo ello a pesar de que en \mathbf{Ex} estuviera contenida la información de que efectivamente ocurre Mj . La explicación de Mj a partir de $\{Aj\}$ y Σ_4 es adecuada, hay una relación de explicación entre $\{Aj\}$ y $|Mj|$ en Σ_4 y hay una relación de pertinencia explicativa adecuada respecto de Σ_4 , entre $|Aj|$ y $|Mj|$. Sin embargo, el único elemento del *explanans* directo tomado en cuenta en tal explicación es falso. Esto significa que el modelo captura explicaciones con *explanans* falible.

Pensemos un caso distinto. Un caso trágico por cierto. Supongamos que todo ha comenzado por una broma. Que Juan realmente no tomó una cucharada de arsénico. Su amigo Roberto, quien había reñido con él el día anterior le jugó una broma haciéndole pensar que lo que había tomado era arsénico. Sin embargo, en una suerte de burla del destino, Juan sale en su coche a toda velocidad y muere en un accidente automovilístico. En tal caso, aún se cumple que $\mathbf{RE}(\{Cj\}, |Mj|)_{\Sigma_4}$. Se cumple la relación de derivación de $\{Cj\} \mid_{\sim\Sigma_4(-vac)} \mathbf{D}(\mathbf{est}_n)$ que ya hemos examinado en el ejemplo *Arsénico 4*. No se cumple, por supuesto, la verdad de Cj . Tenemos una explicación adecuada de Mj a partir de $\{Cj\}$ y Σ_4 , hay una relación de explicación entre $\{Cj\}$ y $|Mj|$ en Σ_4 y hay una relación de pertinencia explicativa adecuada respecto de Σ_4 , entre $|Cj|$ y $|Mj|$. Sin embargo, el único elemento del *explanans* directo tomado en cuenta en tal explicación es falso. De nuevo, un *explanans* falible.

Aprovechando que podemos, recompongámos la tragedia de Juan. Supongamos que todo ha sido una broma. Que Juan realmente no tomó una cucharada de arsénico. Su amigo Roberto, quien había reñido con él el día anterior le jugó una broma haciéndole

pensar que lo que había tomado era arsénico pero supongamos que lo detuvo a la salida, cuando Juan se dirigía a tomar el carro e ir a un hospital a toda velocidad. Así, no es verdad C_j ni A_j . En una situación como ésta, como bien se puede anticipar, aún se cumple que $\mathbf{RE}(\{C_j, A_j\}, |\chi|)_{\Sigma 4}$. Esto tanto en el caso de que Juan no haya muerto, como en el caso de que sepamos que Juan murió naturalmente a los 91 años de edad recordando poco antes de su muerte, con su amigo Roberto, aquella negra broma.

El modelo de explicación que propongo es epistémico en este segundo sentido, no tiene que ver con la verdad de los datos del *explanans*. Recapitulando, algo puede ser una explicación sin completa adecuación al contexto en el que se exige la explicación, sin completa correspondencia con los hechos reales por parte de los factores derivados por el *explanans* para dar cuenta del evento *explanandum* y sin completa correspondencia con los hechos reales por parte de los elementos del *explanans*.

4.2.3.2 Explanandum Falible

Al inicio de este apartado comenté dos características principales de la ambigüedad epistémica: la debilitación de la noción de derivación y, por otra parte, las conclusiones mutuamente contradictorias (inconsistencias inductivas). Como he dicho, mi concepción epistémica y no deductivista de la explicación está basada, entre otras cosas, en el abandono del supuesto de alto grado de certeza para la conclusión. Esta característica es eminentemente no deductivista. Abandonar este supuesto es debilitar la noción de derivación. Si no hay restricciones adicionales, la debilitación puede provocar las inconsistencias inductivas.

Una consecuencia de esta idea de la explicación es que podemos admitir que el *explanandum* sea falible. Esto significa que la certeza real del *explanandum* no es necesariamente alta. Así, tendríamos que modelar la explicación mediante una forma de derivación que proporcione consecuencias falibles. Otra consecuencia importante es que podemos tener inferencias con premisas insuficientes. Ahora bien, si queremos ser

racionales en esta forma de derivar y con este tipo de conclusiones, deberíamos poder imponer algún tipo de restricción a la derivación y deberíamos poder retractar nuestras conclusiones dada nueva información. Como he dicho, la restricción está dada por la relación de pertinencia explicativa exigida mediante la justificación de un argumento explicativo. Ahora trataré de ilustrar cómo mi modelo captura la retractación, las consecuencias falibles y la inferencia a partir de premisas insuficientes. Las tres características son parte de una misma forma de derivación.

En la introducción mencioné algunos ataques contra la idea de una modelación deductiva de la explicación. Uno de ellos se debe a Orayen y otro a Hempel. El de Hempel lo abordé en mi tesis de maestría, aunque todavía superficialmente. El ejemplo de Orayen es el siguiente:

(*María*)

María era golpeada frecuentemente por su esposo

María se divorció

Argumentos de este tipo tienen dos características importantes: la insuficiencia de las premisas para la conclusión y las relaciones causales implícitas que le dan fuerza a la derivación de la conclusión. Se puede anticipar una reformulación de estos argumentos en la lógica del razonamiento por *default* de Reiter. La reformulación fue la siguiente:

(*María-D*)

Gme : "María era golpeada frecuentemente por su esposo"

Dm : "María se divorció"

Gme : \diamond Dm

Dm

María-D es una forma no deductivista de modelar este tipo de casos de explicación pero adolece todavía de la ausencia de una representación de las relaciones causales implícitas en *María*. Además de la representación de relaciones causales, *María-D*, carece de la representación de otras relaciones importantes en un argumento explicativo. Una de estas relaciones, es la relación entre el vínculo causal establecido y la derivación del argumento. En mi modelo, la restricción a la inferencia puede deberse al no cumplimiento con la relación de pertinencia explicativa que, en el caso de eventos fácticos, es un vínculo causal. La restricción a la inferencia es la que nos capacitará para retractar la conclusión del argumento. Analizando la retractación en mi modelo veremos más claramente las relaciones entre ésta, el *explanandum* falible y la inferencia a partir de premisas insuficientes.

Desde mi modelo de explicación una explicación como *María-D*, podría expresarse como sigue²⁸⁴:

(*María-Ex*)

$$\frac{\{G'm\} | \{G'm\} \vdash_{\Sigma_k(-vac)} \mathbf{D}(\mathbf{est})_n : \star_{\Sigma_k} \mathbf{DF}(|Dm|) \wedge [\mathbf{est}]_n}{\mathbf{D}(|Dm|)}$$

donde:

G'm : "María era golpeada por su esposo",

Dm : "María se divorció",

Σ_k es un sistema RCS,

D(est) es la descripción del conjunto de elementos de $[\mathbf{est}]_n$,

|Dm| es el evento a explicar,

²⁸⁴Usaré en este caso "Gm" para designar que María era golpeada por su esposo. En *María-D* usé en cambio "Gme", pero esto complicaría la forma de los defaults que pondré en seguida.

{G'm} es el *explanans*,

D(|Dm|) es el *explanandum* y $D(|Dm|) = Dm$,

Así, |Dm| es un evento fáctico.

Veamos varios casos de sustitución para Σ_k en *María-Ex*. Esto nos proporcionará varios casos de explicación. Presentaré antes un conjunto de *defaults* que podrían estar involucrados en la explicación *María* y que ayudarán en la definición de cada instancia de Σ_k .

Consideremos las siguientes proposiciones:

G'x: "x es golpeada por su esposo",

M'x: "x se siente menospreciada",

A'x: "disminuye la autoestima de x",

Vx: "x desea un cambio de forma de vida",

Dx: "x se divorcia",

Px: "x tuvo un padre golpeador",

Rx: "x es incapaz de establecer relaciones no violentas con personas del sexo opuesto",

y consideremos los siguientes *defaults*:

(d6)

G'x : $\Diamond[M'x \wedge \mathbf{C}(|G'x|, |M'x|)]$

$M'x \wedge \mathbf{C}(|G'x|, |M'x|)$

(d7)

M'x : $\Diamond[A'x \wedge \mathbf{C}(|M'x|, |A'x|)]$

$A'x \wedge \mathbf{C}(|M'x|, |A'x|)$

(d8)

$$\underline{A'x : \diamond[Vx \wedge C(|A'x|, |Vx|)]}$$

$$Vx \wedge C(|A'x|, |Vx|)$$

(d9)

$$\underline{Vx : \diamond Dx}$$

$$Dx$$

(d10)

$$\underline{Px : \diamond[Rx \wedge C(|Px|, |Rx|)]}$$

$$Rx \wedge C(|Px|, |Rx|)$$

(d11)

$$\underline{Rx : \diamond \neg Vx}$$

$$\neg Vx$$

(d12)

$$\underline{\neg Vx : \diamond \neg Dx}$$

$$\neg Dx$$

Veamos ahora algunos casos particulares de Σ_k .

(*María-Ex* / Σ_5)

Supongamos que:

$$\Sigma_5 = \langle \Lambda_1, \Delta_5 \rangle,$$

$$\Delta_5 = \langle D_5, W_5 \rangle,$$

$$D_5 = \emptyset$$

$$W_5 = \emptyset$$

La única extensión de $\Sigma 5$ es $E^{\Sigma 5} = \text{Cn}^{\wedge 1}(W5)$.

La única extensión de $\Sigma 5 + \{G'm\}$ es $E^{\Sigma 5 + \{G'm\}} = \text{Cn}^{\wedge 1}(W5 \cup \{G'm\})$.

Con un sistema epistémico como $\Sigma 5$ es posible tener una cadena que sea instancia de $[\text{est}]_n$ y que cumpla con el esquema *María-Ex*. La red está constituida únicamente por el evento fáctico $|G'm|$. Designemos a la descripción de $|G'm|$ mediante " $\mathbf{D}([\text{est}]_a)$ ".

Como hemos visto en casos similares, a pesar de tener una teoría vacía, como lo es $\Delta 5$, ocurre que $\mathbf{RE}(\{G'm\}, |Dm|)_{\Sigma 5}$, es decir, ocurre que $\{G'm\}$ explica $|Dm|$ respecto de $\Sigma 5$. Así, el esquema que resulta de sustituir Σ_k por $\Sigma 5$ y $[\text{est}]_n$ por $[\text{est}]_a$ en *María-Ex*, llamémoslo "*María*($\Sigma 5, a$)", es una explicación, pues hay una relación de explicación entre $\{G'm\}$ y $|Dm|$. Por lo tanto, $\langle \{G'm\}, Dm \rangle$ es un par *explanans-explanandum* para el sistema $\Sigma 5$, pues hay una explicación que corresponde con dicho par.

Observemos que, Dm se obtendría del cumplimiento de todo el prerrequisito y toda la justificación de *María*($\Sigma 5, a$) y no sólo del hecho de que $\{G'm\}$ y $|Dm|$ mantengan una relación de explicación. Observemos también que la inferencia de Dm no necesita que pueda inferirse $\mathbf{C}(|G'm|, |Dm|)$ ni Dm en $\Sigma 5 + \{G'm\}$. De hecho no son posibles tales inferencias en dicho sistema expandido. Esto refleja que desde mi modelo un evento (por ejemplo $|G'm|$) puede funcionar como un factor explicativo básicamente si es posible que ese evento sea una causa del evento *explanandum* (en este caso $|Dm|$). En este sentido, las premisas de la explicación *María*($\Sigma 5, a$) son insuficientes para el *explanandum*. Aún más, en el caso de que en $\Sigma 5 + \{G'm\}$ se pudiera derivar no deductivamente Dm , las premisas seguirían siendo insuficientes. Por lo tanto, podemos considerar al *explanandum* como una consecuencia falible en el caso *María*. Por supuesto, la situación cambiaría si pudiese derivarse deductivamente en $\Sigma 5 + \{G'm\}$ el *explanandum* Dm , lo cual está contemplado en uno de los disyuntos de la condición usual R1.

Al par *explanans-explanandum* $\langle \{G'm\}, Dm \rangle$ respecto de $\Sigma 5$ le corresponde una sola explicación. Esto se debe a que tenemos una sola descripción de cadena que es derivable de $\Sigma 5 + \{G'm\}$. Con un sistema diferente de $\Sigma 5$ tal vez podrían derivarse varias descripciones de cadenas tales que pudieran constituir cada una diferentes explicaciones bajo un mismo *explanans* y un mismo sistema epistémico, es decir, podríamos tener varias explicaciones para un mismo par *explanans-explanandum* en relación a un mismo sistema epistémico. Desde este modelo, hay una diferencia entre el *explanans* y los factores propuestos por el *explanans* como factores de pertinencia explicativa.

Ahora bien, una explicación como *María*, es una explicación plausible pero no una explicación infalible. Supongamos que obtuviéramos nuevos datos según los cuales no podríamos seguir considerando el argumento *María*($\Sigma 5, a$) como una explicación. ¿Cómo podemos modelar el caso de que un argumento considerado antes como una explicación, deje de serlo? En mi modelo la respuesta está en basar el rechazo de la calidad explicativa de un argumento en la relación de explicación. Hay dos formas en que un argumento considerado una explicación deje de serlo: que no cumpla con la condición de derivación de la descripción del conjunto de elementos de la cadena involucrada o que no cumpla con la condición posibilidad paraconsistente de la descripción fuerte de esa cadena. Así tenemos dos formas de bloqueo de la propiedad de ser una explicación. Con ello, tenemos dos formas de bloqueo del argumento en cuestión. Por tanto, tenemos dos formas de retractación de un *explanandum*. Véamoslo con mayor detalle.

Supongamos que el argumento $\{G'm\}/D(|Dm|)$ es considerado una explicación. Lo que está de fondo es siguiente esquema robustecido:

(*María*($\Sigma 5, a$))

$$\frac{\{G'm\} \mid \{G'm\} \mid \vdash_{\Sigma 5(\sim vac)} D(est)_a : \star_{E^{\Sigma 5}} DF(|Dm|) \times (est)_a}{D(|Dm|)}$$

Pero consideremos que agregamos a $\Sigma 5$ la información de que María no es golpeada por su esposo. Esto último puede modelarse si pensamos en un nuevo sistema RDS2 que contenga todo lo que contenía $\Sigma 5$ más la proposición $\neg G'm$. Veamos entonces el siguiente caso:

(*María-Ex* / $\Sigma 6$)

Supongamos que:

$$\Sigma 6 = \Sigma 5 + \{\neg G'm\} = \langle \Lambda 1, \Delta 6 \rangle,$$

$$\Delta 6 = \langle D5, W6 \rangle,$$

$$D5 = \emptyset$$

$$W6 = \{\neg G'm\}$$

La única extensión de $\Sigma 6$ es $E^{\Sigma 6} = Cn^{\Lambda 1}(\{\neg G'm\})$.

La única extensión de $\Sigma 6 + \{G'm\}$ es $E^{\Sigma 6 + \{G'm\}} = Cn^{\Lambda 1}(W6 \cup \{G'm\})$.

Como $G'm$ no es mal comportada entonces $E^{\Sigma 6 + \{G'm\}} = \perp$.

El argumento:

(*María*($\Sigma 6, a$))

$$\frac{\{G'm\} \mid \{G'm\} \vdash_{E^{\Sigma 6}(-vac)} \mathbf{D}(\mathbf{est}_a) : \diamond_{E^{\Sigma 6}} \mathbf{DF}(\mid Dm \mid) \nexists [\mathbf{est}]_a}{D(\mid Dm \mid)}$$

no es una explicación. La primer razón es que la derivación $\{G'm\} \vdash_{E^{\Sigma 6}(-vac)} \mathbf{D}(\mathbf{est}_a)$ no es posible, dado que $\Sigma 6 + \{G'm\} = \perp$ y esto viola una de las condiciones de la definición de derivación deductiva no vacua, $\vdash_{\Sigma(-vac)}$. No sólo esa derivación no es posible sino que ninguna lo es, es decir, no hay ninguna otra red que pueda derivarse en tal situación. La segunda razón es que $\diamond_{E^{\Sigma 6}} \mathbf{DF}(\mid Dm \mid) \nexists [\mathbf{est}]_a$ tampoco se cumple dado que, por principio de cuentas, no hay ninguna derivación con la cual sea compatible y, en segundo lugar,

$\Sigma_6 + \{G'm\} = \perp$, es decir, no se cumple el requisito de posibilidad paraconsistente. Así, $María(\Sigma_6, a)$ no es una explicación.

Lo anterior modela el hecho de que una explicación, a saber, $María(\Sigma_5, a)$, puede dejar de serlo cuando aumenta la información en su contexto. Dicho de otro modo, lo que hemos modelado es cómo el aumento de información en un cierto contexto (Σ_5) ha producido la pérdida de una explicación que relacionaba a $\{G'm\}$ como *explanans* y a $|Dm|$ como *explanandum*. Consecuentemente, también hemos modelado cómo la tupla $\langle \{G'm\}, Dm \rangle$, que era considerada un par *explanans-explanandum*, dejó de serlo debido al aumento de información en el contexto. Obsérvese que el esquema $María(\Sigma_5, a)$ sigue siendo una instancia de una de las reglas especiales de **Ex**, $María(\Sigma_5, a)$ sigue siendo parte de **Ex** independientemente de que ya no sea considerado una explicación. Todo esto es un tercer sentido en el que mi modelo es un modelo epistémico de la explicación.

Por supuesto, la situación examinada para $María(\Sigma_6, a)$ es también aplicable para el caso de determinar si hay una relación de explicación entre el conjunto *explanans* directo y el evento *explanandum*. Cuando se muestra que la primera no es una explicación, esto se hace mediante la observación de que no todas las condiciones para que haya una relación de explicación entre $\{G'm\}$ y $|Dm|$ en ese contexto, se cumplen. $[est]_a$ no es derivable y no hay en la situación descrita para $María(\Sigma_6, a)$ ninguna otra posibilidad de sustitución para $[est, R]_a$ de modo que, podríamos decir, se bloquea la idea de que haya una relación de explicación entre ese conjunto y ese evento *explanandum*. Así, en el marco teórico formal de **Ex**, uno podría abandonar la idea de que haya una relación de explicación entre $\{G'm\}$ y $|Dm|$ en el contexto Σ_6 . Así, un cuarto sentido en el que mi modelo es epistémico es que podemos retractarnos de atribuir relaciones de explicación entre un conjunto de proposiciones y un evento en un cierto contexto.

El aumento de información en el contexto Σ_5 no sólo hace que el esquema $María(\Sigma_6, a)$ no sea una explicación sino que no pueda efectuarse la derivación de lo que

era el *explanandum*. La consecuencia de un esquema ha sido bloqueada y si no hubiera otra forma de obtenerla entonces podemos decir que dicha consecuencia se retractó en el marco de **Ex**. Sin necesidad de añadir información al sistema $\Sigma 5$ tenemos también otra forma de bloquear la inferencia en *María*($\Sigma 5, a$). Supongamos que la derivación se cumple y que la justificación del esquema también. Todavía faltaría, como hemos visto en algunos casos de *Arsénico*, que la verdad de G'm se cumpliera. Cambios en los sistemas epistémicos podrían producir este efecto. Estas son las formas en que mi modelo captura la falibilidad del *explanandum* y ellas constituyen un cuarto sentido en el que mi modelo epistémico.

Hasta ahora he expuesto sólo el caso de un sólo factor explicativo que puede equipararse con el *explanans*, es decir, el caso en que la descripción del conjunto de eventos involucrados en la cadena resulta en el único elemento del *explanans*. Las formas de bloqueo y de falibilidad expuestas para este caso son también aplicables al caso en el que la cadena consta de más de un sólo factor causal. Hay sin embargo, algunas características interesantes de estos casos que conviene ver con de talle.

(*María-Ex* / $\Sigma 7$)

Supongamos que:

$$\Sigma 7 = \langle \Lambda 1, \Delta 7 \rangle,$$

$$\Delta 7 = \langle D 7, W 7 \rangle,$$

$$D 7 = \{d 6, d 7, d 8, d 9, d 10, d 11, d 12\}$$

$$W 7 = \{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } |\neg V m| \text{ no puede causar } |D m|\}$$

La única extensión de $\Sigma 7$ es $E^{\Sigma 7} = Cn^{\Lambda 1}(W 7)$.

La única extensión de $\Sigma 7 + \{G'm\}$ es $E^{\Sigma 7 + \{G'm\}} = Cn^{\Lambda 1}(W 7 \cup \{G'm\}) = Cn^{\Lambda 1}\{\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con}$

personas del sexo opuesto no causa divorcio, $|\neg Vm|$ no puede causar $|Dm|$, $G'm, M'm \wedge C(|G'm|, |M'm|), A'x \wedge C(|M'm|, |A'm|), \forall x \wedge C(|A'm|, |Vm|), Dm\}$

Existe al menos una cadena $[est]_n$ que cumple con los requisitos para que $\{G'm\}/ D(|Dm|)$ sea una explicación.

De hecho existen muchas redes que los cumplen.

Cuando la trama causal propuesta por un *explanans* es mucho más detallada, la explicación puede ser considerada más interesante. En el caso *María-Ex / Σ7* el caso de red con la trama más detallada capaz de cumplir con la relación de explicación que se ha supuesto es el que se construye gracias a los defaults d6-d8 contenidos en D7 (aunque los *defaults* activados son d6-d9). La descripción de esa red es $C(|G'm|, |M'm|) \wedge C(|M'm|, |A'm|) \wedge C(|A'm|, |Vm|)$. Abreviemos esta cadena mediante $[est]_i$. Así, el siguiente esquema robustecido es una explicación:

(*María*(Σ7,i))

$$\frac{\{G'm\} | \{G'm\} \text{---}_{E^{\Sigma 7}(-vac)} D(est)_i : \text{---}_{E^{\Sigma 7}} DF(|Dm|) \wedge [est]_i}{D(|Dm|)}$$

Esto es suficiente para que el par $\langle \{G'm\}, Dm \rangle$ sea un par *explanans-explanandum* respecto de Σ7. Antes de proseguir, haré dos observaciones. Primero, es importante notar que Dm está en $E^{\Sigma 7 + \{G'm\}}$ pero que esto no es la razón por la que $\{G'm\}$ explica $|Dm|$ respecto de Σ7. Lo relevante es que haya una relación de pertinencia explicativa que vincule al *explanans* con el evento $|Dm|$. Esto se da en la justificación de *María*(Σ7,i), mediante el tramo causal $C(|Vm|, |Dm|)$. Segundo, aún en el caso de que d9 diera como resultado la expresión $C(|Vm|, |Dm|)$ esto no nos obligaría a considerar una cadena distinta de la que hemos considerado para *María*(Σ7,i), es decir, distinta de $[est]_i$, ya que la justificación de *María*(Σ7,i) exige que $|Dm|$ se pueda conectar con la cadena propuesta,

y si la cadena propuesta termina con $|Dm|$, entonces no podemos conectar causalmente $|Dm|$ con $|Dm|$. De modo que a pesar de que la cadena $[est]_i$ más $C(|Vm|, |Dm|)$ sea derivable no podríamos usarla para construir una explicación. Dm es una descripción de red derivable pero suponemos que en general los sistemas epistémicos no admiten que un evento pueda causarse a sí mismo. Así, Dm no sería una descripción de red que corresponda a una explicación.

Hay aún más explicaciones producidas por el contexto $\Sigma 7$ y $\{G'm\}$. Hay al menos 8 más definidas por las siguientes descripciones de cadena:

$$D([est]_a) = G'm$$

$$D([est]_b) = M'm$$

$$D([est]_c) = A'm$$

$$D([est]_d) = Vm$$

$$D([est]_e) = C(|G'm|, |M'm|)$$

$$D([est]_f) = C(|M'm|, |A'm|)$$

$$D([est]_g) = C(|A'm|, |Vm|)$$

$$D([est]_h) = C(|G'm|, |M'm|) \wedge C(|M'm|, |A'm|)$$

De este modo tenemos al menos 8 explicaciones más:

María($\Sigma 7, a$)

María($\Sigma 7, b$)

María($\Sigma 7, c$)

María($\Sigma 7, d$)

María($\Sigma 7, e$)

María($\Sigma 7, f$)

María($\Sigma 7, g$)

María($\Sigma 7, h$)

Si la información del sistema $\Sigma 7$ aumentara, algunos de estos 9 esquemas podrían dejar de ser explicaciones. Por ejemplo, supongamos que el sistema es expandido agregándole la proposición $\neg V_m$. Modelemos esto mediante el sistema resultante $\Sigma 8$:

(*María-Ex* / $\Sigma 8$)

Supongamos que:

$$\Sigma 8 = \Sigma 7 + \{\neg V_m\} = \langle \Lambda 1, \Delta 8 \rangle,$$

$$\Delta 8 = \langle D7, W8 \rangle,$$

$$D7 = \{d6, d7, d8, d9, d10, d11, d12\}$$

$$W8 = \{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } \neg V_m, |\neg V_m| \text{ no puede causar } |D_m|\}$$

La única extensión de $\Sigma 8$ es $E^{\Sigma 8} = C_n^{\Lambda 1}(W8)$.

La única extensión de $\Sigma 8 + \{G'm\}$ es $E^{\Sigma 8 + \{G'm\}} = C_n^{\Lambda 1}(W8 \cup \{G'm\}) = C_n^{\Lambda 1}[\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } \neg V_m, |\neg V_m| \text{ no puede causar } |D_m|, G'm, M'm \wedge C(|G'm|, |M'm|), A'x \wedge C(|M'm|, |A'm|)\}]]$. Este resultado se debe a que $d8$ se bloquea pues no es posible V_m y a que $d9$ se bloquea pues no se cumple con el prerequisite V_m . De modo que si agregamos $\neg V_m$ a $\Sigma 7$ perdemos algunas inferencias: $V_x \wedge C(|A'm|, |V_m|)$ y D_m .

La nueva situación hace que perdamos al menos dos redes: $[est]_g$ y $[est]_i$. Consecuentemente, perdemos dos explicaciones: *María*($\Sigma 7, g$) y *María*($\Sigma 7, i$). Sin embargo, el par $\langle \{G'm\}, D_m \rangle$ sigue siendo un par explicativo para $\Sigma 7$ expandido, es decir, para $\Sigma 8$. Esto es un resultado interesante por dos razones: significa que podemos tener esquemas

que dejan de ser explicaciones y, no obstante, es posible que el conjunto de proposiciones y el evento involucrados, sigan manteniendo una relación de explicación (a través de otras redes propuestas por el mismo *explanans*); significa además que puede suceder que se bloqueen ciertas explicaciones para un cierto contexto y, sin embargo, no se bloquea el prerrequisito, es decir, que hay otras redes que se siguen derivando en el mismo contexto.

Para poder bloquear el prerrequisito en una situación como la que he estado describiendo, sería necesario que ninguna de las redes adecuadas a-i pudiera derivarse. Esto haría que no tuviéramos prerrequisito cumplido en el esquema. Una consecuencia de algo así, es que tampoco tendríamos con qué conectar al evento *explanandum* y no tendríamos tampoco con qué hacer compatible esta conexión. No tener el prerrequisito en una regla robustecida de **Ex** implica no tener tampoco su justificación.

En esquemas de este tipo no es posible bloquear el prerrequisito sin bloquear al mismo tiempo la justificación. La inversa sin embargo sí es posible. Es posible que se bloquee la justificación sin bloquear el prerrequisito. Supongamos que $\Sigma 7$ se expanda agregándole la proposición: " $\neg Vm$ no puede causar Dm ". En este caso tendremos un nuevo sistema $\Sigma 8'$:

(*María-Ex* / $\Sigma 8'$)

Supongamos que:

$$\Sigma 8' = \Sigma 7 + \{ \neg Vm \text{ no puede causar } Dm \} = \langle \Lambda 1, \Delta 8' \rangle,$$

$$\Delta 8' = \langle D7, W8' \rangle,$$

$$D7 = \{d6, d7, d8, d9, d10, d11, d12\}$$

$$W8' = \{ \text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } \neg Vm \text{ no puede causar } Dm, Vm \text{ no puede causar } Dm \}$$

La única extensión de $\Sigma 8'$ es $E^{\Sigma 8'} = Cn^{\Lambda 1}(W8)$.

La única extensión de $\Sigma 8' + \{G'm\}$ es $E^{\Sigma 8' + \{G'm\}} = Cn^{\wedge 1}(W8' \cup \{G'm\}) = Cn^{\wedge 1}[\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } |Vm| \text{ no puede causar } |Dm|, G'm, M'm \wedge C(|G'm|, |M'm|), A'x \wedge C(|M'm|, |A'm|), \forall x \wedge C(|A'm|, |Vm|), Dm\}]$. La extensión $E^{\Sigma 8' + \{G'm\}}$ es muy parecida a la extensión $E^{\Sigma 7 + \{G'm\}}$. Las mismas redes a-i pueden seguir infiriéndose. Pero $E^{\Sigma 8' + \{G'm\}}$ ha ganado una proposición más: la proposición con la que se expandió $\Sigma 7$.

A pesar de que la nueva expansión de $\Sigma 7$ en este caso no produjo pérdida de inferencias en el sistema $\Sigma 7$, sí produce pérdida de explicaciones. Esto se debe a que la conexión de $|Dm|$ a las cadenas $[est]_g$ y $[est]_i$ no es posible pues, según la nueva información, no es posible que $|Vm|$ cause $|Dm|$. Nótese que la situación en $\Sigma 8'$ no hace que perdamos las dos cadenas $[est]_g$ y $[est]_i$ sino que impide que sea posible una conexión con ellas. Se bloquea la justificación pero no el prerrequisito. Estas consideraciones muestran el tipo de relación que tiene el prerrequisito y la justificación en un esquema supuesto en un Σ de **Ex**.

Hasta ahora no hemos visto qué pasa con los *defaults* d10-d12. Veamos un caso más, que muestra algunas cosas interesantes respecto de estos *defaults*. Como $\Sigma 8$ y $\Sigma 8'$, este caso también es una expansión de $\Sigma 7$. La expansión será con la proposición Pm , lo cual hará que entren en juego los *defaults* aludidos.

(*María-Ex* / $\Sigma 9$)

Supongamos que:

$$\Sigma 9 = \langle \wedge 1, \Delta 9 \rangle,$$

$$\Delta 9 = \langle D7, W9 \rangle,$$

$$D7 = \{d6, d7, d8, d9, d10, d11, d12\}$$

W9 = {Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, Pm, $\neg Vm$ no puede causar $|Dm|$, $\neg Dm$ no puede causar $|Dm|$ }

La única extensión de $\Sigma 9$ es $E^{\Sigma 9} = Cn^{\Lambda 1}(W9) = Cn^{\Lambda 1}\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, Pm, } \neg Vm \text{ no puede causar } |Dm|, \neg Dm \text{ no puede causar } |Dm|, Rm \wedge C(|Pm|, |Rm|), \neg Vm, \neg Dm\}$

Lo anterior indica que los *defaults* activados son d10-d12. No se activan d6-d9.

Tenemos dos extensiones para $\Sigma 9 + \{G'm\}$:

$E^{\Sigma 9 + \{G'm\}1} = Cn^{\Lambda 1}\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, Pm, } \neg Vm \text{ no puede causar } |Dm|, \neg Dm \text{ no puede causar } |Dm|, Rm \wedge C(|Pm|, |Rm|), \neg Vm, \neg Dm, G'm, M'm \wedge C(|G'm|, |M'm|), A'm \wedge C(|M'm|, |A'm|)\}$

Lo anterior indica que los *defaults* activados son d6, d7, d10-d12. No se activan los *defaults* d8 y d9.

$E^{\Sigma 9 + \{G'm\}2} = Cn^{\Lambda 1}\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, Pm, } \neg Vm \text{ no puede causar } |Dm|, \neg Dm \text{ no puede causar } |Dm|, Rm \wedge C(|Pm|, |Rm|), G'm, M'm \wedge C(|G'm|, |M'm|), A'm \wedge C(|M'm|, |A'm|), Vm \wedge C(|A'm|, |Vm|), Dm\}$

Lo anterior indica que los *defaults* activados son d6-d9, d10. No se activan los *defaults* d11 y d12.

La situación es conflictiva. Como es natural las dos extensiones son paraortogonales. Este es un caso típico de ambigüedad epistémica. Cada uno de los resultados posee una de al menos dos proposiciones bien comportadas que son mutuamente contradictorias. Dos psicólogos que examinaran las causas del divorcio de

María podrían tener cada uno, desde una misma teoría (sistema Σ_9) dos impresiones distintas de las capacidades de explicación del *explanans* {G'm}. Examinemos cada uno de los casos.

Desde el punto de vista de la extensión $E^{\Sigma_9+\{G'm\}1}$ es imposible que con los datos de $\Sigma_9+\{G'm\}$ pueda explicarse el divorcio de María. Uno de nuestros psicólogos podría optar por esto. De hecho, con estos datos la conclusión que se obtiene es que María no se divorcia, es decir, $\neg Dm$. De modo que jamás se podrá cumplir la justificación de una regla especial que contenga a Dm como consecuencia. Por lo tanto, Dm no puede ser *explanandum* de nada bajo estas circunstancias. Pueden derivarse algunas cadenas pero ninguna de ellas será adecuada para la justificación. Podría pensarse que podríamos considerar que Pm o Rm pueden causar Dm pero esta opción también es imposible debido a las proposiciones originales de W . Estas impiden que pensemos que hechos como Pm o Rm causen divorcios. Por otro lado, la extensión $E^{\Sigma_9+\{G'm\}1}$ representa la decisión de perder dos proposiciones respecto del sistema Σ_7 . Al expandir Σ_7 con las proposiciones Pm y " $\neg Dm$ no puede causar Dm ", se perdieron $Vm \wedge C(|A'm|, |Vm|)$ y Dm al menos para una extensión. Como veremos enseguida, no se pierden para toda extensión de la expansión Σ_9 de Σ_7 .

Desde la perspectiva de $E^{\Sigma_9+\{G'm\}2}$ sí es posible explicar el divorcio de María. El otro de nuestros psicólogos optaría por esta alternativa. Las redes que se pueden derivar para esta extensión son todas las del sistema Σ_7 :

$$D([\mathbf{est}]_a) = G'm$$

$$D([\mathbf{est}]_b) = M'm$$

$$D([\mathbf{est}]_c) = A'm$$

$$D([\mathbf{est}]_d) = Vm$$

$$D([\mathbf{est}]_e) = C(|G'm|, |M'm|)$$

$$D([\mathbf{est}]_f) = C(|M'm|, |A'm|)$$

$$D([\text{est}]_g) = C(|A'm|, |Vm|)$$

$$D([\text{est}]_h) = C(|G'm|, |M'm|) \wedge C(|M'm|, |A'm|)$$

$$D([\text{est}]_i) = C(|G'm|, |M'm|) \wedge C(|M'm|, |A'm|) \wedge C(|A'm|, |Vm|)$$

Y también, desde el punto de vista de $E^{\Sigma 9 + \{G'm\}2}$, $\Sigma 9$ posee las mismas explicaciones que $\Sigma 7$:

María($\Sigma 7, a$)

María($\Sigma 7, b$)

María($\Sigma 7, c$)

María($\Sigma 7, d$)

María($\Sigma 7, e$)

María($\Sigma 7, f$)

María($\Sigma 7, g$)

María($\Sigma 7, h$)

María($\Sigma 7, i$)

Desde la extensión $E^{\Sigma 9 + \{G'm\}2}$, la expansión $\Sigma 9$ de $\Sigma 7$ no sólo no ha perdido información sino que tampoco ha perdido explicaciones.

No obstante que bajo la extensión $E^{\Sigma 9 + \{G'm\}1}$ no es posible explicar el divorcio de María, de acuerdo a mi modelo $\{G'm\}$ explica $|Dm|$ respecto de $\Sigma 9$. La razón es que hay al menos una extensión, a saber, $E^{\Sigma 9 + \{G'm\}2}$, que es adecuada. Esto es un quinto sentido en que mi modelo es epistémico.

Como veremos más adelante, es posible construir explicaciones en relación a $\Sigma 9$ que expliquen $\neg Dm$. Por ahora lo importante es que ante la ambigüedad epistémica lo que queda es un asunto de toma de decisiones. Cada extensión es una alternativa. En la primera extensión se renuncia a una explicación del divorcio de María bajo el contexto de $\Sigma 9$ y en la segunda extensión no. En la primera extensión se abandona Vm por la razón de

Rm mientras que en la segunda extensión, por el contrario, se cree en Vm a pesar del dato de que Rm. En la primera extensión se le da prioridad a Rm sobre A'm y en la segunda extensión se le da prioridad a A'm sobre Rm. Una posición como esta, de aceptación de toma de decisiones a partir de dos alternativas, es coherente con la idea de van Fraassen de que, a fin de cuentas, las explicaciones elegidas dependen de preferencias que están a un nivel intensional. Sin embargo, no por ello estamos obligados a construir un modelo relativista de la explicación, tan solo nos quedamos con un contextualismo que facilite la representación de las interacciones de los candidatos a explicación con su entorno.

4.2.3.3 Explicación Dinámica

He mostrado cómo desde mi perspectiva epistémica el caso *María* puede ser abordado resolviendo dos de las expectativas que teníamos respecto de él: el manejo de conclusiones *explanandum* a partir de premisas insuficientes y la representación del rol de las relaciones causales (representado mediante las cadenas) en las explicaciones. El caso de *María* y sus variaciones ponen también de manifiesto el carácter pragmático de la propuesta: se trata de hacer capturar las inferencias con premisas insuficientes y la toma de decisiones que implica tener varias alternativas para la expansión de las teorías y para sus extensiones correspondientes. El caso *Magneto* propuesto por Hempel, que traté en mi investigación de maestría, puede ser tratado de la misma manera desde este marco formal general para la explicación. Lo importante en este caso es observar la diferencia entre tratarlo desde una expectativa deductivista y una no deductivista. Lo que a Hempel le parece lamentablemente irrecuperable es que por causa de la incompletud ontológica de las confirmaciones de las negaciones de las excepciones (no podemos confirmarlas todas), no podamos construir esquemas deductivos para el caso *Magneto*. Desde mi modelo partimos de una aceptación de esta incompletud ontológica y emprendemos todo desde una perspectiva distinta. La idea en este marco formal para las explicaciones es llegar a una completud epistemológica, no ontológica, pues la noción de derivación puede ser no deductiva. En este sentido, es más importante para el modelo qué esquemas son

explicaciones bajo cierto contexto y no qué explicaciones están completamente fundadas. El modelo distingue mejor el contexto de formulación del contexto de fundamentación, que proponía separar Scriven en sus críticas contra los modelos clásicos.

Es por este interés en observar las relaciones entre distintos componentes de la explicación y de su contexto, o entre explicaciones y contextos, que considero importante tratar explícitamente el tercer aspecto de la representación de la ignorancia: la explicación como algo dinámico.

Ya en los dos aspectos anteriores, el del *explanans* postulado y el del *explanandum* como algo falible, he tratado parte significativa de los puntos que intentaré dejar claros en la exposición de este tercer aspecto. De modo que me referiré a algunos de los casos anteriores para mostrar algunas características importantes del modelo en relación a este aspecto. Mi exposición entonces será tan sólo una complementación de lo que he dicho anteriormente, mediante el señalamiento de ciertos casos de relación interesantes entre los elementos que constituyen una explicación.

La primera relación que quiero enfatizar es la *influencia mutua entre lógica y teoría*. Esta relación ha permanecido estable en todos los ejemplos examinados hasta ahora. En todos ellos he supuesto, en abstracto un mismo sistema lógico (Λ_1), uno de los cálculos de la jerarquía de C_n $1 \leq n < \omega$ de lógica paraconsistente de da Costa y hemos supuesto en esta sección también algunas capacidades metalingüísticas. En mi modelo sólo es posible tratar explicaciones bajo contextos que supongan una lógica como esta, los llamados sistemas RCS. No obstante, pueden tratarse sistemas epistémicos que no supongan esta clase de lógica. Por ejemplo que supusieran lógica clásica de primer orden. Los resultados de hacer estas suposiciones se reflejarán en las extensiones de los sistemas epistémicos resultantes. Lo que a mí me parece importante es, sobre todo, la idea de que los contextos a los que se encuentran referidas las explicaciones, se componen de datos y aparatos inferenciales y que la diferencia entre aparatos inferenciales puede provocar

cambios significativos en las extensiones, y, por tanto, en sus compatibilidades con ciertas explicaciones.

Sin embargo, aún a pesar de que podamos suponer un sistema lógico específico a todo sistema epistémico RCS, los cambios a nivel de información, como hemos mostrado, pueden reflejarse también en cambios significativos. Esto es lo que podemos identificar como la relación de *compatibilidad entre explicaciones y sistemas epistémicos*. Consiste en que algunos esquemas pueden ser considerados explicaciones en relación a ciertos contextos y no en relación a otros. Como vimos, los esquemas $María(\Sigma_n, g)$ y $María(\Sigma_n, i)$ son explicaciones respecto de Σ_7 pero no lo son respecto de Σ_8 , es decir, son explicaciones si Σ_n es sustituido por Σ_7 pero no si es sustituido por la expansión Σ_8 de Σ_7 (tampoco para la expansión Σ_8').

Existe otra relación de compatibilidad que no hemos explorado detalladamente. Se trata de la relación de *compatibilidad entre pares explanans-explanandum y sistemas epistémicos*. Puede ocurrir que un mismo par sea considerado como un par *explanans-explanandum* en relación a un sistema epistémico y no en relación a otro. Observemos que si el sistema epistémico no tiene restricciones acerca de qué tipos de eventos pueden mantener una relación causal, no es posible tener una diferenciación como la que he dicho que puede ocurrir. Un caso trivial de esto es considerar si el divorcio de María es explicable respecto de una teoría distinta. Concretamente, el par $\langle \{G'm\}, D_m \rangle$ es un par *explanans-explanandum* en una teoría como Σ_2 ²⁸⁵. Σ_2 sólo tiene una restricción y es acerca de lo que puede causar la muerte. En esta situación $\{G'm\}$ explica $|D_m|$ respecto de Σ_2 . Sin embargo, los sistemas epistémicos reales, los de los contextos respecto de los cuales nuestras explicaciones son evaluadas, contienen restricciones de este tipo y son una de las razones de los problemas de asimetría examinados antes. En un caso real sucedería algo como lo que sucede en relación con el siguiente ejemplo:

²⁸⁵Véase, *supra*, ejemplo Arsénico 2.

Ejemplo *Arsénico 4'*

Supongamos que:

$$\Sigma 4' = \langle \Lambda 1, \Delta 4' \rangle,$$

$$\Delta 4 = \langle D4, W4' \rangle,$$

$$D4 = \{d1, d2, d3\}$$

$$W4' = \{(\forall x)[Cx \supset (\mathbf{C}(|Cx|, |Tx|) \wedge Tx)], \text{El que alguien coma una cucharada de arsénico no causa su muerte, la contaminación por arsénico del torrente sanguíneo no causa la muerte}\}$$

La única extensión de $\Sigma 4'$ es:

$$E^{\Sigma 4'} = Cn^{\wedge 1}(W4')$$

La única extensión de $\Sigma 4'+\{Cj\}$ es:

$$E^{\Sigma 4'+\{Cj\}} = Cn^{\wedge 1}[\{(\forall x)[Cx \supset (\mathbf{C}(|Cx|, |Tx|) \wedge Tx)], \text{El que alguien coma una cucharada de arsénico no causa su muerte, la contaminación por arsénico del torrente sanguíneo no causa la muerte, } Cj, Tj, \mathbf{C}(|Cj|, |Tj|)\}\}]$$

El ejemplo *Arsénico 4'* puede ser el caso en el que, dada la muerte de Juan, no hay razones para pensar que Juan tiene arsénico acumulado en los pulmones. Siendo así, los *defaults* de $D4$ no pueden activarse. La posibilidad que resta es intentar conectar Mj (la muerte de Juan) con Cj o con Tj . Pero dado que el hecho de que Juan haya tomado una cucharada de arsénico (Cj) no puede ser considerado como causa de su muerte en este contexto, y lo mismo pasa con la contaminación de arsénico del torrente sanguíneo de Juan (Tj), entonces no queda ningún factor que pueda ser utilizado para construir una red adecuada que explique la muerte de Juan. Consecuentemente, $\{Cj\}$ no explica $|Mj|$ respecto de $\Sigma 4'$ pero, recordemos, $\{Cj\}$ sí explica $|Mj|$ respecto del contexto $\Sigma 4$ analizado antes. Dicho de otro modo, el par $\langle \{Cj\}, Mj \rangle$ es un par *explanans-explanandum* respecto del contexto $\Sigma 4$ pero no lo es respecto del contexto $\Sigma 4'$.

He mostrado ya un caso en que un sistema epistémico, dado el candidato a *explanans*, produjo dos extensiones: el caso (*María-Ex* / Σ_9). Este tipo de situaciones reflejan una relación más que también constituye el sentido en el que he dicho que la explicación es algo dinámico: consiste en la *relación entre contextos, información y situaciones de conflicto*. Esta relación capturada en mi modelo es básicamente la forma en que mi modelo representa la ambigüedad epistémica. El caso (*María-Ex* / Σ_9) representa una situación en conflicto. En la primera extensión no es explicable D_m , en la segunda sí es explicable. Esto puede interpretarse, como señalé antes, como un conjunto de alternativas originadas por la distinción de prioridades que uno podría tener con base en la constitución del sistema epistémico en cuestión. Sin embargo, las extensiones en conflicto fueron producidas al suponer que $\{G'm\}$ podría ser un *explanans* de $|D_m|$ respecto del sistema. De modo que ambas extensiones se construyen a partir de la expectativa de esta relación de explicación. Las extensiones se producen en el sistema $\Sigma_9 + \{G'm\}$ no en el sistema Σ_9 solo. De modo que, para ser estrictos, es erróneo decir que $|D_m|$ no es explicable en $E^{\Sigma_9 + \{G'm\}1}$ pero que sí es explicable en $E^{\Sigma_9 + \{G'm\}2}$. Las explicaciones lo son en relación con sistemas epistémicos no con extensiones. Por ello decimos que $\{G'm\}$ explica $|D_m|$ respecto de Σ_9 , pues una de sus extensiones lo hace posible.

Hemos visto que Σ_9 posee una sólo extensión, a saber, $E^{\Sigma_9} = Cn^{\wedge 1}(W_9) = Cn^{\wedge 1}[\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } P_m, |\neg V_m| \text{ no puede causar } |D_m|, |\neg D_m| \text{ no puede causar } |D_m|, R_m \wedge C(|P_m|, |R_m|), \neg V_m, \neg D_m\}]$. Bajo estas condiciones no puede explicarse el evento $|D_m|$. Observemos entonces que el aumento de información que implica considerar $\{G'm\}$ como candidato a *explanans* es lo que produce el conflicto.

Vale la pena observar un caso más patente de conflicto emparentado con el caso (*María-Ex* / Σ_9). Dije que Σ_9 no puede explicar el evento $|D_m|$. Sin embargo, no hay nada

en Σ_9 que limite el tipo de factores que pueden explicar $|\neg D_m|$. Por ejemplo, podría ser que $\{P_m\}$ explicara $|\neg D_m|$ respecto de Σ_9 .

(\neg María-Ex / Σ_9)

Supongamos que:

$$\Sigma_9 = \langle \Lambda_1, \Delta_9 \rangle,$$

$$\Delta_9 = \langle D_7, W_9 \rangle,$$

$$D_7 = \{d_6, d_7, d_8, d_9, d_{10}, d_{11}, d_{12}\}$$

$$W_9 = \{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } P_m, |\neg V_m| \text{ no puede causar } |D_m|, |\neg D_m| \text{ no puede causar } |D_m|\}$$

La única extensión de Σ_9 es $E^{\Sigma_9} = C_n^{\Lambda_1}(W_9) = C_n^{\Lambda_1}[\{\text{Tener un padre golpeador no causa divorcio, ser incapaz de establecer relaciones no violentas con personas del sexo opuesto no causa divorcio, } P_m, |\neg V_m| \text{ no puede causar } |D_m|, |\neg D_m| \text{ no puede causar } |D_m|, R_m \wedge C(|P_m|, |R_m|), \neg V_m, \neg D_m\}]$

Lo anterior indica que los *defaults* activados son d10-d12. No se activan d6-d9.

Tenemos una extensión para $\Sigma_9 + \{P_m\}$:

$$E^{\Sigma_9 + \{P_m\}} = E^{\Sigma_9}$$

En esta situación, $\{P_m\}$ explica $|\neg D_m|$ respecto de Σ_9 . Tenemos para $\Sigma_9 + \{P_m\}$ tres cadenas derivables que son adecuadas para explicar $|\neg D_m|$:

$$D([\text{est}]_j) = \neg V_m$$

$$D([\text{est}]_k) = R_m$$

$$D([\text{est}]_l) = C(|P_m|, |R_m|)$$

Así, tenemos las siguientes explicaciones para $|\neg D_m|$ a partir de $\{P_m\}$ respecto del sistema Σ_9 :

$\neg \text{María}(\Sigma_9, j)$

$\neg \text{María}(\Sigma_9, k)$

$\neg \text{María}(\Sigma_9, l)$

Consecuentemente, se puede explicar $|D_m|$ y $|\neg D_m|$ respecto de un mismo sistema epistémico, aunque en este caso con diferente *explanans*. Obsérvese que el sistema $\Sigma_9 + \{G'm\}$ es una expansión del sistema $\Sigma_9 + \{P_m\}$. Esto es interesante pues captura la idea de que un aumento de información en el contexto puede producir que el contexto sea capaz de explicar *explananda* contradictorios respecto de los que antes podía explicar.

El ejemplo *Rec* planteado por Hempel en su exposición del modelo I-S de explicación, es similar a la situación que acabo de plantear en relación al caso ($\neg \text{María-Ex} / \Sigma_9$). Examinemos los siguientes casos.

Consideremos las siguiente simbolización:

Ix: "x tiene una fuerte gripa"

P'x: "a x se le aplica un antibiotico"

R'x: "x se recupera"

Nx: "x tiene 91 años"

j: "Juan"

Y consideremos los siguientes *defaults*:

(d13)

$$\underline{|x \wedge P'x : \diamond[R'x \wedge \mathbf{C}(|P'x|, |R'x|)]}$$

$$R'x \wedge \mathbf{C}(|P'x|, |R'x|)$$

(d14)

$$\underline{|x \wedge Nx : \diamond[\neg R'x \wedge \mathbf{C}(|Ix|, |\neg R'x|)]}$$

$$\neg R'x \wedge \mathbf{C}(|Ix|, |\neg R'x|)$$

Ahora veamos el siguiente caso:

(*Rec-Ex* / $\Sigma 10$)

Supongamos que:

$$\Sigma 10 = \langle \Lambda 1, \Delta 10 \rangle,$$

$$\Delta 10 = \langle D 10, W 10 \rangle,$$

$$D 10 = \{d13, d14\}$$

$W 10 = \{lj, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación}\}$

La única extensión de $\Sigma 10$ es $E^{\Sigma 10} = Cn^{\Lambda 1}(W 10) = Cn^{\Lambda 1}(\{lj, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación}\})$

En esta extensión no se activa d13 ni d14.

La única extensión de $\Sigma 10 + \{P'j\}$ es $E^{\Sigma 10 + \{P'j\}} = Cn^{\Lambda 1}(\{lj, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación, } P'j, R'j \wedge \mathbf{C}(|P'j|, |R'j|)\})$.

En esta extensión se activa d13 pero no d14.

En este caso tenemos una cadena derivable adecuada para una explicación de $|R'j|$:

$$D([\text{est}]_m) = P'j$$

Por lo tanto tenemos la explicación de $|R'j|$:

$$Rec(\Sigma_{10,m})$$

Sin embargo podemos tener una explicación de un *explanandum* contradictorio, a saber, $\neg R'j$, a partir del mismo sistema epistémico Σ_{10} . $\{Nj\}$ explica $|\neg R'j|$ respecto de Σ_{10} . La única extensión de $\Sigma_{10} + \{Nj\}$ es $E^{\Sigma_{10} + \{Nj\}} = Cn^{\wedge 1}[\{Ij, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación, } Nj, \neg R'j \wedge C(|Nj|, |\neg R'j|)\}]$.

En esta extensión se activa d14 pero no d13.

Así, tenemos una cadena derivable adecuada para una explicación de $|\neg R'j|$:

$$D([\text{est}]_o) = Nj$$

Consecuentemente tenemos la explicación de $|R'j|$:

$$Rec(\Sigma_{10,o})$$

La situación anterior muestra de nuevo cómo un mismo sistema epistémico puede servir de contexto a dos explicaciones cuyos *explananda* son lo que Hempel llamó "inconsistencias inductivas". Ahora bien, ésta es sólo una forma en que mi modelo puede representar la ambigüedad epistémica. Una forma particularmente interesante pues es más fiel a la manera en que Hempel propone el ejemplo *Rec* es la siguiente.

(Rec-Ex / $\Sigma 10'$)

Supongamos que:

$$\Sigma 10 = \langle \Lambda 1, \Delta 10 \rangle,$$

$$\Delta 10 = \langle D10, W10' \rangle,$$

$$D10 = \{d13, d14\}$$

$W10' = \{lj, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación, } P'j\}$

La única extensión de $\Sigma 10'$ es $E^{\Sigma 10'} = Cn^{\Lambda 1}(W10') = Cn^{\Lambda 1}(\{lj, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación, } P'j, R'j \wedge C(|P'j|, |R'j|)\})$

En esta extensión se activa d13 pero no d14.

La única extensión de $\Sigma 10' + \{P'j\}$ es $E^{\Sigma 10' + \{P'j\}} = E^{\Sigma 10'}$.

Así es como Hempel comienza a plantearnos el problema de la ambigüedad epistémica. Hasta este punto el resultado es que $\{P'j\}$ explica $|R'j|$ respecto de $\Sigma 10'$. Después Hempel nos informa que Juan tiene otra propiedad relevante: que tiene 91 años. Esto significa que el sistema $\Sigma 10'$ ha recibido nueva evidencia, y esta evidencia cambiará la situación. La expansión de $\Sigma 10'$ puede verse como un sistema $\Sigma 11$ tal que:

$$\Sigma 11 = \Sigma 10' + \{Nj\} = \langle \Lambda 1, \Delta 11 \rangle,$$

$$\Delta 11 = \langle D10, W11 \rangle,$$

$$D10 = \{d13, d14\}$$

$W11 = \{lj, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación, } P'j, Nj\}$

Hay dos extensiones para Σ_{11} :

La extensión $E^{\Sigma_{11},2} = Cn^{\wedge 1}(\{I_j, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación, } P'_j, N_j, R'_j \wedge \mathbf{C}(|P'_j|, |R'_j|)\})$.

En esta extensión se activa d13 pero no d14.

La extensión $E^{\Sigma_{11},1} = Cn^{\wedge 1}(\{I_j, \text{tener una fuerte gripa no causa recuperación, tener una fuerte gripa no causa no recuperación, tener 91 años no causa recuperación, la aplicación de penicilina no causa no recuperación, } P'_j, N_j, \neg R'_j \wedge \mathbf{C}(|N_j|, |\neg R'_j|)\})$.

En esta extensión se activa d14 pero no d13.

Esta situación es otra forma de modelar la ambigüedad epistémica que presenta Hempel. Hay inconsistencias inductivas en el sistema Σ_{11} . R'_j y $\neg R'_j$ son ambas derivables de Σ_{11} . No obstante, en el marco formal Ex las inconsistencias inductivas no causan estragos. R'_j y $\neg R'_j$ se derivan en diferentes extensiones. Además, una consecuencia de mi modelación es que las proposiciones contradictorias, ambas, son *explananda* del mismo sistema. $\{P'_j\}$ explica $|R'_j|$ respecto de Σ_{11} y $\{N_j\}$ explica $|\neg R'_j|$ respecto de Σ_{11} . Nuevamente, la ambigüedad epistémica se traduce en un problema de decisión acerca de cuál de las dos explicaciones es mejor que la otra, pero mi modelo considera a ambas como legítimas explicaciones en el contexto Σ_{11} . Este puede ser considerado un sexto sentido en el que mi modelo es epistémico.

Además de las relaciones importantes que he mencionado hay al menos otras dos relaciones que también tienen que ver con el sentido en que digo que la explicación debe ser entendida como algo dinámico. La primera de ellas es la relación entre pares *explanans-explanandum* y redes, que permite la sobredeterminación epistémica causal. Bajo un mismo sistema epistémico, podemos tener varias redes distintas que justifiquen cada una de ellas que un mismo par sea un par *explanans-explanandum*. Esto fue ejemplificado ya en varios de los casos presentados a lo largo de esta sección. La segunda relación es la relación entre extensiones y cadenas, que permite que conjuntos distintos

de explicaciones puedan pertenecer a un mismo sistema epistémico. Algunas de estas explicaciones pueden repetirse en algunas de las extensiones. Esto representa la capacidad de un mismo sistema epistémico para responder de distintas maneras en relación a la expectativa sobre si un esquema es o no una explicación en relación con ese sistema. Los casos presentados en los que tuvimos que examinar más de una extensión sirven como ilustración de esta relación.

Simplificadamente, concebir la explicación como una entidad dinámica significa concebirla como algo complejo, es decir, como algo que se compone de varios elementos mutuamente relacionados que, en sus diversas combinaciones, concibiendo a los elementos que conforman una explicación en un sentido muy “modular”, nos muestran diversas formas en las que puede mantenerse una relación de explicación.

4.2.4 Paraconsistencia

Hemos visto algunos casos que muestran cómo el marco formal Ex captura algunos de los ejemplos paradigmáticos de los problemas para la construcción de un modelo de explicación. Los primeros casos atendieron problemas vinculados con la relación de pertinencia explicativa. Los segundos, se refirieron a problemas relacionados con la no monotonicidad y la ambigüedad epistémica. Presentaré ahora algunos casos que se encuentren vinculados con problemas relacionados con la inconsistencia, aunque ya hemos adelantado algo al respecto, cuando hemos tratado el problema de la ambigüedad epistémica.

En el capítulo III, expuse un caso problemático de modelación de la explicación. La dificultad para modelar el caso que presenté, a saber, el de la explicación de Frege de la noción de número, consiste en que el *explanans* de la explicación fregeana es inconsistente. A pesar de esta inconsistencia, la estrategia de explicación de Frege fue preservada y, por otra parte, el *explanans* inconsistente se usó en otras explicaciones y

como fundamento de otras teorías. Esto último hace deseable la modelación de una explicación como ésta que supone un *explanans* inconsistente. Como vimos en el capítulo que acabo de mencionar, aunque con otras palabras, el problema para la modelación de este tipo de *explanans* es que si suponemos la lógica clásica como lógica subyacente a la teoría inconsistente automáticamente el conjunto de consecuencias de tal teoría respecto de esa lógica es trivial, es decir, contiene toda fórmula del lenguaje.

Sin embargo, normalmente una teoría que contenga alguna contradicción no es considerada una teoría lógicamente trivial y, como dije, puede incluso ser un caso de teoría valiosa para fundamentar muchas explicaciones. Así, siguiendo nuestro enfoque pragmático, lo que se necesita es la posibilidad de admitir algunas contradicciones bloqueando la trivialidad. Si la trivialidad no se bloqueara, no habría forma de distinguir la línea argumentativa que justifica la propuesta de la cadena a partir del *explanans*. Lo cual iría en contra de una buena parte de la filosofía que hemos tratado de sugerir en la sección donde hicimos algunas reflexiones sobre la Lógica del Razonamiento por *default* de Raymond Reiter. Modelar explicaciones cuyo *explanans* es trivial dificultaría entender la estructura lógica atribuida a tal explicación.

Una dificultad adicional, apunté antes también, tenía que ver con la pregunta de cómo hacer compatible el mecanismo inferencial de los *default* de Reiter con un sistema de lógica paraconsistente, pues, como vimos, la inferencia por *default* se encuentra basada sobre un tipo de posibilidad lógica que supone consistencia clásica.

El modelo para la explicación que presenté en la sección anterior resuelve ambos problemas. La solución para ambos está apoyada fundamentalmente en una interpretación más general y flexible de lo que entendemos por regla y por teoría *default*. En el marco formal **Ex**, se pueden construir "teorías" con un nivel mayor de generalidad. Les llamamos "sistemas epistémicos" y son tuplas compuestas de reglas de inferencia y proposiciones. La idea está basada en la misma propuesta de Reiter respecto de la forma

en que concibe la inferencia mediante razonamientos no monotónicos. Reiter construye teorías *default* compuestas de un conjunto de reglas y un conjunto de proposiciones y supone la lógica clásica como instrumento primario para la inferencia. Así, en el marco formal **Ex** se parte de sistemas epistémicos en un sentido más general y, con la ayuda de una caracterización también más general de la noción de regla, caracterizamos tanto lógicas como teorías. Las lógicas no necesariamente son clásicas y las teorías no necesariamente son teorías *default*. Podemos así, con las herramientas expuestas en la sección anterior, construir sistemas que combinan lógicas con teorías, llamados "sistemas epistémicos complejos".

En el caso del modelo de explicación que propongo, la lógica supuesta es paraconsistente y las teorías son teorías *default* modificadas (paraconsistentes). La modificación principal en las teorías *default*, es que el tipo de posibilidad lógica se encuentra basado no en lógica clásica sino en lógica paraconsistente, en la no trivialidad. Esto hace que la compatibilidad entre el mecanismo inferencial de los *default* y la lógica paraconsistente se cumpla en este ambiente formal. Consecuentemente, con ello podemos modelar sin trivialidad los casos del tipo de la explicación de Frege, casos de inconsistencia en el *explanans*.

La posibilidad de modelación de *explanantes* inconsistentes produce también algunas consecuencias respecto de la clase de cosas que debería racionalmente suponer el marco formal **Ex**. Esas consecuencias pueden restringirse en el marco formula **Ex** mediante cláusulas adicionales. En este apartado expondré cómo se capturan los casos de *explanantes* inconsistentes y explicaré con ejemplos la intención de algunas cláusulas que pueden estipularse para el marco formal **Ex**.

4.2.4.1 Inconsistencia en el Sistema Epistémico

En un sistema RCS $\Sigma = \langle \Lambda, T \rangle$, los sistemas exigidos para las explicaciones, la lógica subyacente Λ es una lógica paraconsistente.

En un sistema epistémico RCS puede haber inferencias deductivas e inferencias no deductivas. Podemos distinguir esto en Ex estipulando que en una derivación deductiva usamos sólo reglas deductivas, en una derivación no deductiva usamos al menos una regla no deductiva²⁸⁶. En ambos casos, la paraconsistencia está presente. Cuando la inferencia es deductiva los axiomas de las jerarquías C_n $1 \leq n \leq \omega$ del sistema impiden que se deduzcan conclusiones que causen trivialidad. Pero hay que recordar que en un sistema paraconsistente sólo causan trivialidad las contradicciones entre fórmulas bien comportadas y no las contradicciones entre fórmulas mal comportadas. En el caso de las inferencias no deductivas, la noción de posibilidad paraconsistente que se ha caracterizado en la sección anterior, checa que no puedan deducirse conjeturas que provoquen trivialidad en el sistema. Se bloquearán las inferencias que den conjeturas tales que, junto con las demás consecuencias en una extensión dada, produzcan trivialidad.

Las fórmulas mal comportadas pueden interpretarse como fórmulas que aparecen como partes de una contradicción que ya estaba originalmente en la teoría. La propiedad de los sistemas de da Costa de ser finitamente trivializables posibilita una estrategia clave para compatibilizar la paraconsistencia y la inferencia *default*. La noción de trivialidad sustituye en los *default* de las teorías RCS la noción de contradicción. En este sentido, el espíritu de la distinción de Newton da Costa está presente en las teorías *default* paraconsistentes que se han construido en la sección anterior.

²⁸⁶ Varias de las distinciones que haré en adelante son relativamente fáciles de construir en Ex pero no las expondré en esta investigación.

En realidad hay dos tipos de posibilidad lógica en el sistema **Ex**. Una es el que se aplica a los *default* contenidos por la teoría en los sistemas RCS (simbolizado mediante " \diamond ") y otro que se aplica a las condiciones para que un esquema en un sistema epistémico pueda ser considerado una explicación (simbolizado mediante " \diamond ").

En el primer caso, se trata de una modificación de la posibilidad que se maneja normalmente en los *default* de Reiter. La modificación en mi propuesta tiene dos aspectos importantes: el primero es la suposición de una lógica paraconsistente y no una lógica clásica en el contexto del sistema epistémico RCS. El segundo aspecto es la definición de derivación que involucra defaults, basándola en un chequeo de no trivialidad y no en un chequeo de no contradicción. Este último aspecto se definió aprovechando nuestra definición más general de regla, que permite usar, "modularmente" las condiciones de presencia y de ausencia por un lado, y los contextos de derivación y de asunción, por otro lado. Además, hay otro rasgo que se añade en los *default* de mi modelo, y que tiene que ver con la importancia que reviste, a partir de las reflexiones hechas acerca de la no monotonicidad, la idea de una inferencia bien justificada. Este rasgo es el chequeo de no trivialidad entre las justificaciones de los *defaults* usados en una estrategia heurística asociada a una extensión del sistema epistémico que se esté usando. Se checan las justificaciones respecto de la extensión final pero no enfocándose en la verificación de no contradicciones sino en si la justificación produce o no trivialidad en la extensión. En mi modelo las justificaciones no sólo se checan respecto de la extensión final sino también respecto del conjunto de justificaciones y prerequisites en una misma estrategia heurística. Cuando se checan respecto de la extensión final, el cumplimiento consiste en verificar que la extensión no se trivialice. A esto lo llamaré "paraconsistencia con la extensión y la estrategia heurística".

Cuando se checan respecto de los otros *default* de la estrategia heurística, el cumplimiento depende de que no haya en el conjunto de consecuencias lógicas de las justificaciones y los prerequisites, fórmulas que produzcan trivialidad. Así es como está

diseñada la derivación *default* en el marco formal Ex. Sin embargo, podríamos reflexionar sobre una posible variación de esta clase de derivación.

Podríamos querer preservar una racionalidad más clásica en relación a nuestra forma de justificar al interior de un sistema epistémico. Sería defendible el deseo de preservar no contradicción en las fórmulas que usamos para la justificación de un conjunto de consecuencias. Pero si esto fuera deseable, también lo sería intentar preservar no contradicción en estas fórmulas no sólo en un contexto de derivación sino en un contexto de asunción. Es decir, en este tren de pensamiento, sería deseable que ninguna de las fórmulas que han sido usadas para justificar un conjunto de consecuencias, se trate de una fórmula derivada o solo de una fórmula asumida, sea la contradictoria de usada para el mismo propósito. En pocas palabras, bajo esta racionalidad, no deberíamos admitir justificación de un conjunto de proposiciones, **usando** proposiciones contradictorias. Esto implicaría distinguir entre tener, por un lado, en el contexto usado para la justificación (o en su conjunto de consecuencias) fórmulas contradictorias y, por otro lado, usar dos fórmulas mutuamente contradictorias para justificar un mismo conjunto de consecuencias. Dada que la distinción es sutil, pongamos un ejemplo: no es lo mismo admitir que el explanans contextual de la explicación de Frege, de la noción de número contenga un principio que nos lleve a contradicciones (a saber, el principio de abstracción nos lleva ala paradoja de Russel), que usaren nuestra justificación de algo la propia contradicción (digamos, usar la idea de que el conjunto R de Russell pertenece a sí mismo, tanto como la idea de que no pertenece a sí mismo).

Así, podríamos querer preservar la racionalidad basada en chequeo de no trivialidad sólo para las fórmulas que sean mal comportadas y, en cambio, desear que no haya justificaciones o prerequisites contradictorios en el sistema. Más formalmente, estamos hablando de que no haya un conjunto de consecuencias $Cn^{\Sigma}(Z)$ en el que se haya usado, para construirlo, un par de reglas R1 y R2 y un par de fórmulas α y $\neg\alpha$, tales que R1 use α y R2 use $\neg\alpha$, sea α bien comportada o no. Lo que se estaría exigiendo con este

aspecto es una racionalidad más clásica a la hora de construir un conjunto de consecuencias. Se trata de exigir no contradicción entre los elementos que activaron las reglas usadas para construir un conjunto de consecuencias dado. ¿Cómo podríamos lograr esto?

Una manera es redefiniendo nuestra noción de regla *default* RGS, no basada sólo en no trivialidad, sino también basada en no contradicción. Esto puede ponerse en términos formales mediante una nueva versión de las definiciones 24 y 25, anotadas en la sección anterior. Las nuevas versiones podrían ser las siguientes.

Definición 23'a. Sea un sistema-RC $\langle \Lambda, \Delta \rangle$ y sea $E^i \in X(\langle \Lambda, \Delta \rangle)$. Definimos una función PJ tal que aplicada a una estrategia heurística s_k of $\langle \Lambda, \Delta \rangle$ para E^i , esto es, $\mathbf{PJ}(s_k)$, resulte en el conjunto de todas las justificaciones y todos los prerequisites de cada *default* en s_k .

Definición 23'b. BB^Σ (bad behaved) es el **conjunto de contradicciones aceptadas** $\alpha \wedge \neg \alpha$ de las fórmulas mal comportadas **de** Σ , en símbolos, $BB^\Sigma = \{\alpha \wedge \neg \alpha \mid \alpha \wedge \neg \alpha \in Cn^\Sigma(\emptyset)\}$, bajo el supuesto de que Σ es un ceps con una lógica Λ de la jerarquía de Costa $C^*_{\omega} \ 1 \leq n < \omega$.

Definición 24'. Sea el ceps $\Sigma = \langle \Lambda, \Delta \rangle + Z$ donde Δ es una teoría *default theory* tal que $\Delta = \langle D, W \rangle$. Decimos que un *default* $d = (\alpha : \diamond \beta_1, \dots, \diamond \beta_n / \omega) \in D$, i.e., un $((A, \gamma), B) = ((\{\alpha\}, \omega), \{\beta_1, \dots, \beta_n\})$, es un **default-RGSC** (*default* de Reiter con coherencia clásica de provisos) *sii* para cada extensión E, en cada estrategia heurística s_k de Σ para E:

- A) A tiene un solo contexto de derivación: $K^e \subseteq \text{FOR}$.
- B) B tiene contexto de derivación, $K^{e,d} \subseteq \text{FOR}$, y contexto de asunción, $K^{e,a} \subseteq \text{FOR}$.
- C) SI $A \subseteq K^e$, $Cn^\Sigma(B \cup K^{e,d} \cup K^{e,a}) \neq \text{FOR}$, $BB^\Sigma \cap Cn^\Lambda(\mathbf{PJ}(s_k)) = \emptyset$, Y (si $\beta \in B$ entonces $\neg \beta \notin Cn^\Lambda(K^{e,d} \cup K^{e,a})$), ENTONCES $\gamma_i \in E$.
- D) $K^e = K^{e,d} = E$.
- E) $K^{e,a} = \mathbf{PJ}(s_k)$.

Toda E construida con estas condiciones puede llamarse una **extensión RGSC** de $\Sigma+Z$. Y podemos llamar **sistemas-RCSC** a los ceps que tienen solo *defaults* RCSC.

En suma, se trata de admitir las contradicciones originales del contexto, de no admitir las nuevas que pudieran producirse, y de no admitir conjeturas que sean derivadas (mediante *defaults*) apelando a fórmulas contradictorias con otras fórmulas usadas ya.

Veámos algunos ejemplos sencillos para mostrar cómo funciona la paraconsistencia en el primer caso, esto es, al interior de los sistemas RCS con *defaults* RGSC.

Consideremos las siguientes proposiciones:

Xx: "x es de derecha"

Kx: "x simpatiza con la globalización"

Yx: "x es nacionalista"

Zx: "s es socialista"

Ux: "x cree que existe el imperialismo económico"

Ox: "x es neoliberal"

a: "Augusto"

s: "el papá de Augusto"

Y los siguientes *defaults*:

(d15)

Zs : \diamond Us

Us

(d16)

Us : \diamond -Ka

-Ka

(d17)

Xa : \diamond -Za

Ka

(d18)

Zs : \diamond -Xa

-Ka

(d19)

Xa : \diamond -Za

Ka

(d20)

Ya : \diamond Za

-Oa

(d21)

-Xa : \diamond -Oa

-Oa

(d22)

Xa : \diamond Ya

Ya

Ahora veamos el siguiente caso:

(Paraconsistencia 1 / $\Sigma 12$)

Supongamos el sistema RCSC:

$$\Sigma 12 = \langle \Lambda 1, \Delta 12 \rangle,$$

$$\Delta 12 = \langle D 12, W 12 \rangle,$$

$$D 12 = \{d 15, d 16\}$$

$$W 12 = \{Xa, (\forall x)(Kx \supset \neg Yx), (\forall x)(\neg Yx \supset \neg Xx), Xa \supset Ka, Zs\}$$

La única extensión de $\Sigma 12$ es $E^{\Sigma 12} = Cn^{\Lambda 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), (\forall x)(\neg Yx \supset \neg Xx), Xa \supset Ka, Zs, Ka, Ka \supset \neg Ya, \neg Ya \supset \neg Xa, \neg Ya, \neg Xa, Ya, \neg Ka, Us, \})$

En esta extensión se activan d15 y d16 y la única estrategia heurística para $E^{\Sigma 12}$ es $s^{E^{\Sigma 12}} = \langle d 15, d 16 \rangle$

En este caso tenemos que en la extensión $E^{\Sigma 12}$ son derivables las contradicciones: $Xa \wedge \neg Xa$, $Ya \wedge \neg Ya$ y $Ka \wedge \neg Ka$. Sin embargo, como $\Sigma 12$ es un sistema RCSC $\Lambda 1$ es una lógica de da Costa. Por lo tanto las tres contradicciones son entonces consideradas como contradicciones originales del sistema $\Sigma 12$. Las fórmulas Xa , Ya y Ka son fórmulas mal comportadas de sistema $\Sigma 12$. La información en $W 12$ y los *default* de $\Sigma 12$, aunque potencialmente inconsistente, pueden ser los de una persona común y corriente que intenta explicarse las relaciones entre la globalización, el ser de derecha y el nacionalismo. El sistema $\Sigma 12$ es un sistema que contiene contradicciones pero no es un sistema trivial, al menos en el sentido lógico.

Está claro que la lógica de referencia en el operador de posibilidad lógica es una lógica de la jerarquía de da Costa de la jerarquía que hemos estado usando, ahora observemos los dos tipos de bloqueo de la salvedad de los *defaults* de $D 12$. Por una parte, Us y $\neg Ka$ se infieren sin ningún problema. En la extensión final, ni la justificación Us ni la justificación $\neg Ka$ producen trivialidad en relación con $E^{\Sigma 12}$. Us es una fórmula bien comportada pero no se deriva su contradictoria $\neg Us$. En cambio, respecto de $\neg Ka$ sí se

deriva Ka, pero se deriva desde el principio a partir de W: es una fórmula mal comportada. De modo que la inferencia por *default* de $\neg Ka$ no afecta nada a la extensión. Se cumple así la paraconsistencia con la extensión final. Por otra parte, la segunda condición de bloqueo de la salvedad de los *defaults* d15 y d16 también es favorable. No hay contradicción en $Cn^{\wedge 1}(\{Zs, Us, \neg Ka\} \cup E)$ de la estrategia heurística $s^{E\Sigma 12}$. De modo que se cumple la consistencia con la estrategia heurística.

(Paraconsistencia 2 / $\Sigma 13$)

Supongamos el sistema RCSC:

$$\Sigma 13 = \langle \Lambda 1, \Delta 13 \rangle,$$

$$\Delta 13 = \langle D13, W13 \rangle,$$

$$D13 = \{d17, d18\}$$

$$W13 = \{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs\}$$

$\Sigma 13$ tiene dos extensiones.

La primera es $E^{\Sigma 13-1} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, Ka, \neg Ya\})$

En esta extensión se activa d17 pero no d18 y la única estrategia heurística para $E^{\Sigma 13-1}$ es $s^{E\Sigma 13-1} = \langle d17 \rangle$.

La segunda es $E^{\Sigma 13-2} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, \neg Ka\})$

En ésta se activa d18 pero no d17 y la única estrategia heurística para $E^{\Sigma 13-2}$ es $s^{E\Sigma 13-2} = \langle d18 \rangle$.

En el caso de la aplicación de d17 como en el de la aplicación de d18 la paraconsistencia con la extensión correspondiente se cumple. En d17 la justificación es $\neg Za$, y no provoca trivialidad. En d18 la justificación es $\neg Xa$ y, aunque Xa está en la extensión, esto no produce trivialidad pues Xa es una fórmula mal comportada en $\Sigma 13$. Con respecto a la consistencia con la estrategia heurística sucede lo mismo, tanto d17 como d18 la cumplen: en d17 y en d18 no hay contradicciones internas.

Nótese que en este caso si el operador de posibilidad de d17 y d18 hubiera sido definido tal com lo hizo Reiter, Σ_{13} tendría de cualquier forma dos extensiones. La razón es que Ka y $\neg Ka$ son fórmulas contradictorias y esto hace que la justificación de cada uno de los *defaults* realmente no se cumpla en combinación con el otro. Bajo mi modelo hay dos extensiones por dos razones: la primera de ellas es la misma que la que se aplica al sistema de Reiter, es decir, que Ka y $\neg Ka$ provocarían trivialidad ya que Ka es una fórmula bien comportada; la segunda es que el prerrequisito de d17 (Xa) es contradictorio respecto de la justificación de d18 ($\neg Xa$), tener ambos en una misma extensión significaría, independientemente de Ka , no cumplir con la consistencia respecto de la estrategia heurística.

(Paraconsistencia 3 / Σ_{14})

Supongamos el sistema RCSC:

$$\Sigma_{14} = \langle \Delta_{14}, \Delta_{14} \rangle,$$

$$\Delta_{14} = \langle D_{14}, W_{14} \rangle,$$

$$D_{14} = \{d_{19}, d_{20}\}$$

$$W_{14} = \{Xa, (\forall x)(\neg Yx \supset \neg Xx), \neg Xa, \}$$

Σ_{14} tiene dos extensiones.

La primera es $E^{\Sigma_{14-1}} = Cn^{\wedge 1}(\{Xa, (\forall x)(\neg Yx \supset \neg Xx), \neg Xa, Ya, Ka\})$

En esta extensión se activa d19 pero no d20 y la única estrategia heurística para $E^{\Sigma_{14-1}}$ es $s^{E^{\Sigma_{14-1}}} = \langle d_{19} \rangle$.

La segunda es $E^{\Sigma_{14-2}} = Cn^{\wedge 1}(\{Xa, (\forall x)(\neg Yx \supset \neg Xx), \neg Xa, Ya, \neg Oa\})$

En ésta se activa d20 pero no d19 y la única estrategia heurística para $E^{\Sigma_{14-2}}$ es $s^{E^{\Sigma_{14-2}}} = \langle d_{20} \rangle$.

Ambos *defaults* cumplen con paraconsistencia respecto de sus respectivas extensiones y ambos cumplen con consistencia respecto de sus respectivas estrategias heurísticas. La proposición Ya se deriva deductivamente en ambas extensiones. Ka y $\neg Oa$,

en cambio, se derivan no monotónicamente. Nótese que si los *defaults* hubieran sido definidos como en el sistema de Reiter, habría una sólo extensión. La extensión sería $Cn^{\wedge 1}(\{Xa, (\forall x)(\neg Yx \supset \neg Xx), \neg Xa, Ya, Ka, \neg Oa\})$, no habría riesgo de trivialización. Sin embargo, en mi modelo, tomando en cuenta estas nuevas cláusulas de racionalidad paraconsistente, se producen dos extensiones dado que usar d19 y d20 para una misma extensión es considerado como irracional. Esto se debe a que las justificaciones de d19 y d20, a saber, $\neg Za$ y Za respectivamente, aunque son perfectamente posibles en el sentido de que la extensión final, no contiene su respectiva fórmula contradictoria, no cumplen sin embargo con el requisito de consistencia respecto de la estrategia heurística. Es irracional, desde mi modelo, construir un conjunto de conocimientos apelando a fórmulas contradictorias, aún en el

(Paraconsistencia 4 / $\Sigma 15$)

Supongamos el sistema RDS2:

$$\Sigma 15 = \langle \Lambda 1, \Delta 15 \rangle,$$

$$\Delta 15 = \langle D15, W13 \rangle,$$

$$D15 = \{d21, d22\}$$

$$W13 = \{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs\}$$

$\Sigma 15$ tiene dos extensiones.

La primera es $E^{\Sigma 15-1} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, \neg Oa\})$

En esta extensión se activa d21 pero no d22 y la única estrategia heurística para $E^{\Sigma 15-1}$ es $s^{E^{\Sigma 15-1}} = \langle d21 \rangle$.

La segunda es $E^{\Sigma 15-2} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, Ya, \neg Ka\})$

En ésta se activa d22 pero no d21 y la única estrategia heurística para $E^{\Sigma 15-2}$ es $s^{E^{\Sigma 15-2}} = \langle d22 \rangle$.

En este caso d21 y d22 cumplen con las dos formas de bloqueo revisadas (paraconsistencia y consistencia). De nuevo, si d21 y d22 hubieran sido definidos como

Reiter lo hace, entonces tendríamos una sólo extensión. La extensión sería $Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, \neg Oa, Ya, \neg Ka\})$. En mi modelo, con la racionalidad paraconsistente que estamos suponiendo, tenemos dos extensiones, pues los prerequisites de ambos *default* se contradicen y, por lo tanto, no pueden formar parte de una misma estrategia heurística.

Los ejemplos anteriores muestran el comportamiento de la paraconsistencia en combinación con las teorías que contienen *defaults*. Este mismo comportamiento se aplica a derivaciones de redes como las que hemos visto en apartados anteriores de esta sección. Es importante notar, finalmente, que las fórmulas mal comportadas y sus contradicciones implicadas son fórmulas que aparecen en el desarrollo deductivo del sistema epistémico original. Esto implica, en mi modelo, que son fórmulas que aparecerán en toda extensión del sistema en cuestión. Se cumplen con todo esto, las tres expectativas que expuse al principio de este subapartado: admitir las contradicciones originales del contexto, no admitir las nuevas que pudieran producirse, y no admitir conjeturas apelando a fórmulas contradictorias.

4.2.4.2 Inconsistencia y Explicación

En los últimos 4 casos hemos examinado el funcionamiento de la paraconsistencia al interior de un sistema epistémico RCSC. Desde mi perspectiva, la explicación se encuentra relativizada minimamente a un sistema epistémico RCS, o bien del tipo RCSC. De modo que las características que expuse a través de estos casos podrían considerarse también como una forma en que pueden capturarse las explicaciones.

En otros trabajos formulé una serie de cláusulas que acotan lo que podríamos llamar la "racionalidad paraconsistente" del marco formal **Ex**. Algunas de las cláusulas tratan sobre propiedades de los pares explicativos y otras sobre propiedades de los *explanantes* y los *explananda*. Todas las cláusulas intentan restringir o relajar la admisión

de contradicciones bajo ciertas condiciones, en relación a la noción de explicación. Tomemos estas cláusulas como una propuesta para complementar el tipo de racionalidad que puede requerirse suponer en una modelación de la explicación científica. Por esta razón lo que haré en esta última parte de la presente sección es tan sólo explicar cada una de las cláusulas de paraconsistencia de un modo muy breve, detallando sólo aquellos aspectos que considero más importantes en relación con la forma de capturar la paraconsistencia desde el marco formal Ex.

Cláusula 6.1: No se admiten derivaciones de elementos de una cadena, tal que la derivación contenga fórmulas mutuamente contradictorias.

La cláusula 6.1 dictaría que cualquier derivación de elementos de una cadena no usa fórmulas contradictorias. La intención de poner esta cláusula es no admitir propuestas de cadena que se encuentren justificadas mediante fórmulas contradictorias. Esto es coherente con nuestra idea de que no podemos inferir nada que haya apelado a fórmulas contradictorias y respeta una mínima racionalidad para admitir que una fórmula se encuentra o no justificada.

La cláusula 6.1 se refiere a la derivación señalada en el prerrequisito de un esquema de explicación robustecido, como hemos hecho a lo largo de esta sección. Se trata de una derivación que puede ser de dos tipos: deductiva o no deductiva. Como dije antes, en una derivación deductiva usamos sólo reglas deductivas, en una derivación no deductiva usamos al menos una regla no deductiva. Así, en este último caso podemos tener una derivación no deductiva pura, cuando no usamos ninguna regla deductiva, o una derivación no deductiva no pura, cuando usamos al menos una regla no deductiva y al menos una regla no deductiva.

En el caso de una derivación no deductiva pura la cláusula se sigue directamente de la definición de sistema RCSC. Los sistemas RCSC son sistemas epistémicos complejos

cuya lógica es una lógica de da Costa y los *defaults* supuestos en la teoría del sistema son *defaults* con coherencia clásica entre salvedades y prerequisites. He mostrado ya cómo funciona un sistema constituido así y ha quedado claro que en tales sistemas no puede haber inferencias que usen sólo *defaults* y que usen dos proposiciones mutuamente contradictorias. Un ejemplo de este caso puede basarse en *Paraconsistencia 2*.

(*Paraconsistencia 2* / $\Sigma 13$)

Supongamos el sistema RDS2:

$$\Sigma 13 = \langle \Lambda 1, \Delta 13 \rangle,$$

$$\Delta 13 = \langle D 13, W 13 \rangle,$$

$$D 13 = \{d 17, d 18\}$$

$$W 13 = \{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs\}$$

$\Sigma 13$ tiene dos extensiones.

La primera es $E^{\Sigma 13-1} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, Ka, \neg Ya\})$

En esta extensión se activa d17 pero no d18 y la única estrategia heurística para $E^{\Sigma 13-1}$ es $s^{E^{\Sigma 13-1}} = \langle d 17 \rangle$.

La segunda es $E^{\Sigma 13-2} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, \neg Ka\})$

En ésta se activa d18 pero no d17 y la única estrategia heurística para $E^{\Sigma 13-2}$ es $s^{E^{\Sigma 13-2}} = \langle d 18 \rangle$.

Pensemos que alguien quisiera explicar el evento de que Augusto está de acuerdo con la venta de Banamex ($|Ba|$) mediante el *explanans* de que Augusto es de derecha y supongamos que el evento de que Augusto simpatiza con la globalización puede conectarse causalmente con el evento *explanandum*. Así, el *explanans* es $\{Xa\}$ y la descripción de red que nos interesa es Ka . Por lo tanto, tenemos que $\{Xa\}$ explica $|Ba|$ respecto de $\Sigma 13$. El par $\langle \{Xa\}, Ba \rangle$ es un par *explanans-explanandum* respecto de $\Sigma 13$. Ahora bien, en la derivación de Ka a partir de $\Sigma 13 + \{Xa\}$ no se usa ninguna regla deductiva, lo único que usamos es d17. Este es un caso de derivación no deductiva pura. Podemos

asegurar, por lo tanto que no se usaron fórmulas contradictorias para la derivación de Ka pues d17 en la primera extensión cumple con la consistencia respecto de la estrategia heurística correspondiente.

En el caso de la derivación no deductiva no pura, la cláusula 6.1 es necesaria. La lógica de da Costa está diseñada para admitir contradicciones originales de la teoría y al mismo tiempo preservar cuanto se pueda los teoremas de la lógica clásica. Es por ello que con la lógica que suponen los sistemas RCSC pueden realizarse deducciones que usen a dos fórmulas mutuamente contradictorias. Consecuentemente, si hacemos una derivación que además de *defaults* use reglas de la lógica de da Costa, corremos el riesgo de que la derivación se haya obtenido usando dos fórmulas contradictorias. Veamos el siguiente ejemplo basado en el caso anterior.

(Paraconsistencia 5 / $\Sigma 16$)

Supongamos el sistema RCSC:

$$\Sigma 16 = \langle \Lambda 1, \Delta 16 \rangle,$$

$$\Delta 16 = \langle D 13, W 16 \rangle,$$

$$D 13 = \{d 17, d 18\}$$

$$W 16 = \{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, (\neg Ya \wedge \neg Xa) \supset \neg Oa\}$$

$\Sigma 16$ tiene dos extensiones.

La primera es $E^{\Sigma 16-1} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, (\neg Ya \wedge \neg Xa) \supset \neg Oa, Ka, \neg Ya, \neg Oa\})$

En esta extensión se activa d17 pero no d18 y la única estrategia heurística para $E^{\Sigma 16-1}$ es $s^{E^{\Sigma 16-1}} = \langle d 17 \rangle$.

La segunda es $E^{\Sigma 16-2} = Cn^{\wedge 1}(\{Xa, (\forall x)(Kx \supset \neg Yx), \neg Xa, Zs, (\neg Ya \wedge \neg Xa) \supset \neg Oa, \neg Ka\})$

En ésta se activa d18 pero no d17 y la única estrategia heurística para $E^{\Sigma 16-2}$ es $s^{E^{\Sigma 16-2}} = \langle d 18 \rangle$.

Supongamos que alguien quiere explicar el evento de que Augusto no está de acuerdo con la venta de Banamex ($\neg Ba$). Esto tal vez podría efectuarse mediante el *explanans* $\{Xa\}$. Es posible deducir $\neg Oa$ a partir del contexto $\Sigma_{16} + \{Xa\}$: $\neg Oa$ está en la primera extensión. Si se quisiera usar la descripción del conjunto elementos de una cadena $\neg Oa$ para la explicación en cuestión, podríamos basarnos en la primera extensión. Sin embargo, para obtener $\neg Oa$ usamos d17 y el condicional $(\neg Ya \wedge \neg Xa) \supset \neg Oa$. Como d17 supone tener Xa y hacer un *modus ponens* con el condicional supone tener $\neg Xa$, $\neg Oa$ se obtiene mediante dos fórmulas contradictorias. La cláusula 6.1 restringe que este tipo de derivaciones pueda usarse para cadenas en una explicación. Y esto puede cubrirse mediante las modificaciones que hemos hecho a las definiciones 24 y 25.

Como he dicho, en el capítulo III expuse un caso de explicación cuyo *explanans* es inconsistente: el caso de la explicación fregeana de los números. El *explanans* contextual de la explicación de Frege es la teoría de conjuntos de Cantor. Russell descubrió una paradoja en el seno de dicha teoría. La paradoja consiste en que si suponemos un conjunto N de todos los conjuntos que no se pertenecen a sí mismos, entonces N se pertenece a sí mismo si y sólo si N no se pertenece a sí mismo.

Frege usa principalmente dos proposiciones en su explicación de la noción de número: el axioma V de su sistema (axioma de comprensión) y la definición de pertenencia a un conjunto (\in). Frege construye una definición de "número" suponiendo que a cada número le corresponde un conjunto de conjuntos que tienen la misma cardinalidad. El evento a explicar puede verse como la existencia de esta secuencia de conjuntos de conjuntos. El evento a explicar puede entonces plantearse como el evento de que la siguiente proposición es verdadera: Existe una secuencia ordenada de conjuntos de conjuntos tal que hay una relación 1-1 entre los miembros de ésta secuencia y los miembros de la secuencia de los números naturales. Podemos formular la explicación de Frege como sigue:

AxC) El conjunto formado por las x que cumplen la propiedad X es igual al conjunto formado por las x que cumplen con la propiedad Z si y sólo si para todo individuo z sucede que z cumple con X si y sólo si cumple con Z. (Ax.

Comprensión)

Df \in) x pertenece al conjunto y si y sólo si existe al menos una propiedad X tal que y es el conjunto formado por las cosas que cumplen con X y x cumple con X. (Def. de Pertenencia)

ϕ) Existe una secuencia ordenada de conjuntos de conjuntos tal que hay una relación 1-1 entre los miembros de esta secuencia y los miembros de la secuencia de los números naturales.

En realidad, ϕ , desde mi modelo, debe interpretarse como la proposición que dicta que la proposición "Existe una secuencia ordenada de conjuntos de conjuntos tal que hay una relación 1-1 entre los miembros de esta secuencia y los miembros de la secuencia de los números naturales" es verdadera. Las versiones formales de AxC y Df \in son:

$$\text{AxC) } (\forall x)(\forall y)(\forall^2 X)(\forall^2 Z)[x^{\wedge}Xx = y^{\wedge}Zy \leftrightarrow (\forall z)(Xz \leftrightarrow Zz)]$$

$$\text{Df}\in) (\forall x)(\forall y)(\forall z)[x \in y \leftrightarrow (\exists^2 X)(y = y^{\wedge}Xz \wedge Xx)]$$

donde " $x^{\wedge}Xx$ " es el conjunto formado por todas las x que cumplen con X. De estas dos proposiciones se sigue el llamado "Principio de Abstracción":

$$\alpha) u \in x^{\wedge}Xx \leftrightarrow Xu$$

que dicta que cualquier cosa u es elemento del conjunto de las x que cumplen con X si y sólo si u cumple con X.

Con ayuda del Principio de Abstracción Frege define la pertenencia a conjuntos de conjuntos con la misma cardinalidad. Supone que u es un conjunto cualquiera y X es la propiedad de tener la misma cardinalidad que los demás elementos del conjunto. Estos conjuntos son ordenables de modo que correspondan biyectivamente a los números naturales. Así, Frege ha demostrado ϕ . Supongamos que tenemos un sistema RCSC, $\Sigma 17$, que contiene los principios de la teoría de conjuntos de Cantor, y todo lo demás que haya usado Frege para su explicación. La explicación de Frege podría entonces ajustarse al siguiente esquema:

(Frege($\Sigma 17, w$))

$$\frac{\{AxC, Df\in\} \mid \{AxC, Df\in\} \vdash_{E^{\Sigma 17}(-vac)} D(est_w) : \diamond_{E^{\Sigma 17}} (|\forall\phi|) \wedge [est]_w}{D(|\forall\phi|)}$$

donde $[est]_w$ es α (el Principio de Abstracción). Así, se cumple la derivación deductiva de α a partir del *explanans* $\{Axc, Df\in\}$ y se cumple que es compatible la conexión de $|\forall\phi|$ con α , respecto de tal derivación. De hecho lo que sucede no es sólo que es posible la conexión sino que ϕ puede derivarse en el sistema

Frege usa el Principio de Abstracción (α). Con él y la ayuda de otros principios, digamos, a partir de $\Sigma 17 + \{AxC, Df\in, \alpha\}$, se sigue ϕ . Sin embargo, también de α se sigue la contradicción de Russell (β). Mediante una instanciación de u por $x^\wedge(x \notin x)$ y de X por $x \notin x$:

$$\beta) x^\wedge(x \notin x) \in x^\wedge(x \notin x) \leftrightarrow x^\wedge(x \notin x) \notin x^\wedge(x \notin x)$$

Según la cláusula 6.1 una derivación de elementos de una cadena no debe usar fórmulas mutuamente contradictorias. Para derivar el Principio de Abstracción, α , Frege usa sólo lógica de segundo orden y el *explanans* que hemos construido, es decir, el axioma de comprensión y la definición de pertenencia. La prueba, entonces, es una deducción.

Puede hacerse mediante prueba condicional suponiendo cada lado del bicondicional de $u \in x^{\wedge}Xx \leftrightarrow Xu$ y usando ambos principios $AxC, Df \in$.

La derivación de α a partir de $\{AxC, Df \in\}$ no usa fórmulas contradictorias²⁸⁷. El problema es, más bien, que α es más general de lo que debería ser. Como es sabido, el Principio de Abstracción puede producir cualquier conjunto posible. α puede construir también el conjunto de Russell, el conjunto de los conjuntos que no se pertenecen a sí mismos: $x^{\wedge}(x \notin x)$.

No obstante, el punto importante ahora es que la derivación del Principio de Abstracción no usa fórmulas contradictorias en $\Sigma 17 + \{AxC, Df \in\}$. De modo que la derivación de α para la explicación $Frege(\Sigma 17, w)$ cumple con la cláusula 6.1. Siendo este el punto difícil de cumplir de la explicación en cuestión, tenemos que $\langle \{AxC, Df \in\}, V\phi \rangle$ es un par *explanans-explanandum* respecto del contexto $\Sigma 17$. Así, la explicación de Frege es, como he dicho antes, históricamente importante para la filosofía de las matemáticas, y es, además, una explicación racional desde el punto de vista de la perspectiva de nuestro marco formal y la adición de estas cláusulas de racionalidad paraconsistente.

Cláusula 6.2: Son admisibles dos pares *explanans-explanandum* con explananda mutuamente contradictorios siempre y cuando estén asociados a diferentes sistemas epistémicos.

En el marco formal **Ex** pueden coexistir varios sistemas epistémicos RCSC. Esta característica permite que en **Ex** puedan combinarse *explanantes*, *explananda* y contextos (Sistemas RCSC). Podría suponerse que **Ex** también puede tener datos sobre el mundo, tiene además de un conjunto **S** de sistemas epistémicos, y podría tener un conjunto **C** de datos que son considerados ciertos. Así, **Ex** podría verse como un modelo del conjunto de

²⁸⁷Para una prueba formal del principio de abstracción a partir del axioma de comprensión y de la definición de pertenencia puede consultarse Resnik, Michael D., *Frege and the Philosophy of Mathematics*, Cornell University Press, Ithaca y London, 1980, pp. 212 y 213.

creencias de una persona que está interesada por las explicaciones. Este conjunto de creencias se supone imparcial respecto de los sistemas que contiene. Cada sistema es considerado como un mundo posible que puede ser compatible o no con otro de ellos. En **Ex** se asocian ciertos pares *explanans-explanandum* a ciertos contextos, es decir, dada cierta información inamovible (el contenido de **C**, **Ex** relaciona distintos contextos con distintas explicaciones y distintos pares *explanans-explanandum*).

Puede suponerse que bajo **Ex** es posible considerar *explananda* a proposiciones mutuamente contradictorias, todo depende de a qué contexto se encuentran asociados. Sucederá lo mismo con *explanantes* y pares $\langle A, D(e) \rangle$. Particularmente, algunos de los casos en que hay proposiciones mutuamente contradictorias mientras mantenemos el mismo *explanans* o el mismo contexto o la misma extensión exigirían, si queremos ser racionales, algunas restricciones. Para admitir los *explananda* contradictorios, tendrían que provenir de modo diferente: por ejemplo, que no sea el mismo *explanans* contextual, o que no sea el mismo *explanans* directo, o que no sea la misma extensión, o que no sea la misma cadena, o que no sea la misma estrategia heurística.

En el apartado sobre no monotonicidad y ambigüedad epistémica tratamos con algunos casos de inconsistencias inductivas. Estos casos son casos de *explananda* mutuamente contradictorios²⁸⁸. Particularmente, la cláusula 6.2 es interesante para los casos que podrían cumplir con tener el mismo *explanans* directo y contextual. Estos casos son los relacionados con el ejemplo propuesto por Hempel para explicar la ambigüedad epistémica.

Un resultado básico que mostré es que es posible explicar la recuperación de Juan y la no recuperación de Juan desde un mismo *explanans* contextual: por ejemplo, $\Sigma 11$. En un caso, la aplicación de penicilina a Juan explica la recuperación de Juan, por otro, el

²⁸⁸Véase, *supra*, casos *María* $\Sigma 8$, *-María* $\Sigma 9$, *Rec* $\Sigma 10$, *Rec* $\Sigma 10'$ y *Rec* $\Sigma 11$.

hecho de que Juan tiene 91 años explica la no recuperación de Juan, ambos casos respecto de Σ_{11} . Más simbólicamente:

$\{P^j\}$ explica R^j respecto de Σ_{11} .

$\{N_j\}$ explica $\neg R^j$ respecto de Σ_{11} .

Es decir, que $\langle \{P^j\}, R^j \rangle$ y $\langle \{N_j\}, \neg R^j \rangle$ son pares *explanans-explanandum* respecto de Σ_{11} .

En el mismo contexto es posible construir explicaciones con *explananda* mutuamente contradictorios que tengan no sólo el mismo *explanans* contextual sino también el mismo *explanans* directo. Basta pensar en $\{P^j, N_j\}$ como un *explanans*. Así, tendríamos que:

$\{P^j, N_j\}$ explica R^j respecto de Σ_{11} .

$\{P^j, N_j\}$ explica $\neg R^j$ respecto de Σ_{11} .

En este caso sucede que $\langle \{P^j, N_j\}, R^j \rangle$ y $\langle \{P^j, N_j\}, \neg R^j \rangle$ son pares *explanans-explanandum* respecto de Σ_{11} .

Podríamos tener muchos otros casos de tuplas *explanans-explanandum* que tienen en común el mismo *explanans* (directo y contextual). Algunos pares de estas tuplas tendrán *explananda* compatibles, otros, como en el caso anterior tendrán *explananda* mutuamente contradictorios. Sin embargo, la cláusula 6.2, dicta que en el marco formal **Ex** se podrían restringir estos pares *explanans-explanandum* de modo tengan alguna diferencia en sus combinaciones de elementos como por ejemplo cadenas, estrategias o extensiones.

Obsérvese que mantener un par *explanans-explanandum* no implica compromisos acerca de las cosas que creemos que suceden en el mundo en el marco de **Ex**. Es decir, el

analizador de explicaciones representado mediante el sistema **Ex** puede, sin ningún problema, mantener cualquier tipo de pares, pues éstos sólo corroboran el cumplimiento de ciertas condiciones mínimas entre un conjunto de proposiciones, un evento y un sistema RCS o RCSC. Las explicaciones lo son siempre en relación a otros elementos teóricos.

Cláusula 6.3 No deberían ser aceptables *explananda* mutuamente contradictorios que provienen de exactamente los mismos elementos teóricos.

Un caso distinto es cuando tenemos varias explicaciones que refieren a un mismo *explanans* contextual y directo, tales que sus consecuencias sean mutuamente contradictorias, es decir, no sería racional mantener simultáneamente los *explananda* que fueran incompatibles en este sentido.

La cláusula 6.3 refiere a este tipo de casos. Dicta que en el contexto de los mismos elementos teóricos, sólo pueden mantenerse simultáneamente los conjuntos consistentes de *explananda*.

Obsérvese que si dos *explananda* derivados mutuamente contradictorios no fueran derivados relativizadamente al mismo *explanans* (contextual o directo), entonces podrían mantenerse simultáneamente en un mismo conjunto.

Cláusula 6.4 Son aceptables *explananda* contradictorios siempre y cuando no provengan de una justificación que suponga inconsistencias y no sean fórmulas bien comportadas.

Una vez admitidos *explanantes* inconsistentes, podría suceder que tuviéramos en **Ex** *explananda* que contengan fórmulas mutuamente contradictorias.

La cláusula 6.4. dicta dos condiciones para aceptar en **Ex** este tipo de *explananda*: que no provengan de una derivación en la que se usen dos fórmulas contradictorias (sean estas bien comportadas o no) y que las fórmulas contradictorias en cuestión no sean bien comportadas. La idea de la cláusula es evitar, como en la cláusula 6.1 derivaciones que usen dos proposiciones mutuamente contradictorias y, por otro lado, no aceptar proposiciones autocontradictorias que provoquen trivialidad.

Además, la cláusula nos indica que lo mismo se aplica para las derivaciones de elementos de cadena en una explicación, que sean proposiciones que contengan dos fórmulas mutuamente contradictorias. Dicho de otra forma, a estos casos se les aplica la cláusula 6.1 referida precisamente a redes derivadas en el prerrequisito de una explicación y se adiciona la condición de que la contradicción contenida no involucre fórmulas mutuamente contradictorias.

Un caso que tiene que ver con la cláusula 6.4. es el de la cadena derivada en el prerrequisito de la explicación de Frege analizada unos párrafos antes. En esa derivación se relacionan principalmente tres proposiciones. Éstas son el axioma de comprensión, la definición de pertenencia y el Principio de Abstracción:

$$\text{AxC) } (\forall x)(\forall y)(\forall^2 X)(\forall^2 Z)[x^{\wedge}Xx = y^{\wedge}Zy \leftrightarrow (\forall z)(Xx \leftrightarrow Zz)]$$

$$\text{Df}\in) (\forall x)(\forall y)(\forall z)[x \in y \leftrightarrow (\exists^2 X)(y = y^{\wedge}Xz \wedge Xx)]$$

$$\alpha) u \in x^{\wedge}Xx \leftrightarrow Xu$$

En la explicación *Frege*($\Sigma 17$, w) el *explanans* es {AxC, Df \in } y la descripción de elementos de cadena derivada es el Principio de Abstracción (α). Cuando examiné este caso mencioné que no hay fórmulas contradictorias usadas en la derivación de α a partir de {AxC, Df \in }. Por tanto, cumple con la primera condición impuesta en la cláusula 6.4.

Sin embargo, es importante observar que tal vez el verdadero problema de la explicación *Frege*($\Sigma 17$, w) no es la derivación de α sino el contenido de α . La pregunta es si el Principio de Abstracción es un principio que contiene fórmulas contradictorias. Antes dije que el problema con α es que es demasiado general, tanto como para producir por instanciación la paradoja de Russell:

$$\beta) x^{\wedge}(x \notin x) \in x^{\wedge}(x \notin x) \leftrightarrow x^{\wedge}(x \notin x) \notin x^{\wedge}(x \notin x)$$

La cuestión está entonces en decidir si la derivación de β a partir de α significa que α contiene fórmulas contradictorias o no. Pero la implicación de una fórmula contradictoria como la de Russell (β) no significa que la fórmula antecedente (α) es también contradictoria. Considerar α como una proposición con fórmulas contradictorias por esta razón, es posible sólo si consideramos que α es equivalente con β . Aquí el punto se dirime de distintas maneras dependiendo del tipo de lógica que estemos considerando como lógica subyacente a la teoría en cuestión, es decir, la lógica contenida en $\Sigma 17$.

Si la lógica contenida en $\Sigma 17$ fuera una lógica clásica entonces α y β serían equivalentes. La razón consta de dos partes: α implica β por instanciación y, por otra parte, si se supone β , como la lógica sería clásica, el sistema se trivializa y, por ende, de β se sigue α . Por tanto, α y β serían equivalentes. Sin embargo, en mi modelo el contexto de la explicación de Frege es interpretado como un sistema con una lógica paraconsistente. Consecuentemente, en este contexto mientras que α implica β por instanciación, no es cierto que β implique α pues suponer β no implica que el sistema se trivialice: β es una fórmula obtenida del conjunto de consecuencias deductivas del sistema y por ello es una contradicción de fórmulas mal comportadas. Por tanto, desde el marco formal Ex, α y β no son equivalentes.

Dado que desde el sistema **Ex** α y β no son equivalentes, no podemos considerar a α como un principio que contenga fórmulas contradictorias. Por ello, α cumple con la

segunda condición que dicta la cláusula 6.4 para estos casos y la explicación *Frege*($\Sigma 17, w$) cumple con estos cánones de racionalidad paraconsistente.

Cláusula 6.5: No son aceptables simultáneamente dos conjuntos de explanans, directos o contextuales, para un mismo *explanandum*, que contengan fórmulas mutuamente contradictorias.

Finalmente, en el marco formal **Ex** no sólo puede tratarse con *explanantes* inconsistentes sino que podría suceder que, en un mismo *explanans* contextual se ofrecieran explicaciones, de un mismo *explanandum*, cuyos *explanantes* directos sean incompatibles entre sí. Dicho de otro modo, podría suceder, que un par de explicaciones de un mismo *explanandum* y respecto de un mismo *explanans* contextual, contengan *explanantes* tales que la unión de ellos resulta un conjunto inconsistente.

La cláusula 6.5 dicta que en estos casos sólo son aceptables los conjuntos consistentes de *explanantes*. No podemos mantener simultáneamente conjuntos de proposiciones incompatibles para la explicación de un mismo *explanandum* en un mismo contexto.

Las cláusulas que he expuesto solo son un ensayo en el camino de estipular un complemento para la racionalidad supuesta en el marco formal Ex. Hace falta una mayor fundamentación para realizar esta complementación. Deben tomarse como una sugerencia para continuar el camino, a partir del marco propuesto en Ex, para contruir una racionalidad subyacente a un modelo mínimo, epistémico y pragmático, de la explicación científica. Esta racionalidad, desde el punto de vista formal, si ha de abordar los problemas planteados por la Filosofía de la Ciencia, nos dirige como he argumentado, a modelos paraconsistentes combinados con no monotonicidad.

CONCLUSIONES

La investigación de esta tesis

La investigación que he presentado en esta tesis puede verse como desarrollada en dos grandes partes básicas. En el primero de ellos realicé un análisis y una desarticulación de los problemas principales que tendría la tentativa de modelación argumental de la explicación. En la segunda parte realicé una propuesta formal de modelación y expliqué cómo esta propuesta trataría los casos paradigmáticos analizados a lo largo de la primera parte. La primera parte está claramente enfocada a destacar los problemas con los que tropezaría una propuesta inferencial, luego se descubren algunos supuestos de estos problemas y finalmente se propone una perspectiva diferente para tratar la modelación de la explicación. La argumentación en contra de una modelación argumental se desarticula en esa primera parte mediante un cambio de perspectiva que consiste en abandonar una perspectiva de modelación deductivista y ontológica y sostener, en cambio, una perspectiva que sea no deductivista, epistémica y pragmática. Con esta nueva perspectiva logramos la desarticulación de los problemas generados por los casos de contraejemplos paradigmáticos presentados en las críticas que se revisaron. Los problemas que se pueden desarticular directamente con esta estrategia que seguí son los relacionados con la relación de inferencia y los que involucran la relación de pertinencia explicativa.

Sin embargo, esta nueva perspectiva abrió nuevos problemas a la modelación, relacionados con la necesidad de capturar en cierta medida la interacción entre explicaciones y contexto. En esta parte de la investigación se enfatizaron algunos de los problemas que habían surgido, aunque marginalmente a una propuesta inferencialista, en la crítica contra los modelos clásicos. Los nuevos problemas son básicamente: que algunos

contextos no proporcionan información suficiente para inferir el *explanandum*; además, que la interacción entre los candidatos a explicación y diversos contextos afecta el que puedan estos candidatos ser o no calificados como explicaciones, por último, que algunos contextos asociados a explicaciones pueden resultar inconsistentes. Todo ello añadió un cúmulo de problemas concretos nuevos a la modelación inferencial de la explicación.

Más detalladamente, tres supuestos importantes que analicé en los capítulos 1 y 2 son los siguientes:

- a) El deductivista en la inferencia del modelo.
- b) El de alta expectabilidad.
- c) El de corrección.

Los supuestos estuvieron asociados a problemas y contraejemplos paradigmáticos. Los problemas se presentaron en el marco de dos postes indicadores que llevaron la discusión al terreno del tema de la racionalidad científica: la tensión Hempel-Scriven y la tensión Hempel-Coffa. Esta conexión entre el problema de la modelación de la explicación científica y la noción de racionalidad no es nueva, aunque en mi investigación, motivado por las reflexiones de Morado sobre racionalidad y lógica [Morado 2004], planteo estas discusiones siempre vinculándolas a la búsqueda de mecanismos formales de representación de los cánones que constituyen la racionalidad. El desarrollo de la perspectiva epistémica y pragmática que desarticula los problemas a partir del abandono principalmente de estos supuestos, me llevó a proponer un aparato conceptual con el cual poder montar una perspectiva como la que se necesita para esta modelación y, así, poder mantener algunas distinciones importantes que ya habían adelantado, aunque de modo más vago, algunos filósofos, por ejemplo, de manera muy importante, la distinción de Scriven entre explicación y fundamentos de la explicación, y la diversificación de explicaciones a partir de la intensionalidad, que hace van Fraassen. Algunas de las nociones importantes que distingue el aparato conceptual que propuse son: explicaciones, relación de explicación, relación de pertinencia explicativa, *explanans* directo y *explanans* contextual. Este aparato conceptual me permitió poner de relieve entre otras cosas: 1) la distinción entre la explicación como una propuesta teórica para

elucidar un fenómeno, y las relaciones reales que ocurren entre los fenómenos referidos en los enunciados de esa explicación; 2) el carácter falible que, en tanto propuesta teórica, tiene la explicación y sus elementos; y 3), la interacción de la explicación con diversos contextos (sistemas epistémicos).

Así, en la segunda parte básica de la investigación, propuse un modelo formal de explicación que, partiendo de la nueva perspectiva, pudiera servir como un ambiente formal para el análisis de la interacción entre sí de los componentes de un candidato a explicación y, también, por otro lado, entre estos componentes y diversas particularidades de los contextos. El modelo se basó en una interacción entre la Lógica del Razonamiento por *default* de Reiter y una jerarquía de cálculos paraconsistentes de da Costa. Este marco formal para el análisis de las explicaciones se construyó partiendo de 1) una generalización de la noción de regla que usa de modo fundamental la idea de regla *default*, 2) una generalización de la idea de teoría que se usa en la lógica de Reiter, 3) una modificación en la forma de construir una extensión de una teoría, que enfoca la derivación *default* a un contexto lógico paraconsistente, y 4) una modificación del mecanismo interno de las reglas, que nos centra en la verificación de no trivialización en cambio de centrarse en la de no contradicción.

En este marco formal se relacionan, entonces, las distinciones que constituyen el aparato conceptual propuesto antes, junto con las capacidades de distinción y de representación que permiten las caracterizaciones generales de regla y sistema epistémico. Estas capacidades pueden así capturar la interacción entre diversos elementos de la explicación de un modo que podemos calificar de “modular”, en el sentido de que permite combinar (o intercambiar) elementos constituyentes de la explicación e intercambiar elementos constituyentes de los contextos. Lo que permite a su vez un análisis complejo y formal básico tanto de las características a considerar de un candidato de explicación, como de los cambios explicativos que son repercusión de cambios en los contextos teóricos. Entre otras cosas combinables en el modelo tenemos: *explanans*

directo, *explanans* contextual, relaciones de explicación, lógicas, teorías, relaciones de pertinencia explicativa, explicaciones, tipos de reglas, tipos de derivación, eventos fácticos y eventos no fácticos.

De este modo en la revisión de contraejemplos paradigmáticos, que hice al final de la tesis, trato de hacer notar en términos generales la adecuación que tiene el marco formal que propuse, para el análisis de la relación entre explicación, asimetrías, ambigüedades epistémicas, no monotonicidad y, finalmente, paraconsistencia. Todo ello enfocado también en mostrar que el marco formal que propuse es adecuado para modelar argumentalmente una noción mínima, epistémica, no deductivista y pragmática de la explicación, que atiende seriamente los problemas propuestos por la Filosofía de la Ciencia contra una perspectiva estándar de esta importante función científica.

Algunas aportaciones y estudios pendientes de esta investigación

Tal como ha sido desarrollada esta investigación, pone sobre la mesa de discusión una serie de discusiones de finales del siglo XX en Filosofía de la Ciencia que, de una forma o de otra, han cuestionado la utilidad de un modelo argumental de la explicación científica e, incluso, la conveniencia de usar representaciones formales para la elucidación de aspectos de la Ciencia. La discusión y el análisis que aquí se han presentado de los problemas clásicos en torno al problema de la construcción de un modelo para la explicación científica han partido de la exposición de los supuestos subyacentes a ciertas críticas que parecían tener como consecuencia que una representación argumental de la explicación no era posible. El énfasis en esta investigación fue puesto sobre la idea de que las críticas estuvieron principalmente dirigidas contra propuestas construidas desde una racionalidad deductivista o con instrumentos formales basados en Lógica Clásica. La idea era mostrar que si incorporáramos en la mesa de discusión los grandes avances que la

Lógica No Clásica ha alcanzado en la segunda mitad del siglo XX, la situación de la discusión tendría que cambiar. Esta transformación del contexto de discusión no solo debería tener repercusiones directamente filosóficas, en este caso sobre nuestra propia concepción de la explicación, sino también sobre la revaloración del uso de una metodología fundada en instrumentos lógicos para la elucidación de algunos aspectos de la Ciencia.

De manera particular, la idea de explicación que se logró configurar utilizando estas herramientas lógicas también tiene algunos aportes que deben destacarse. Considero importantes sobre todo dos de ellos. El primero consiste en la construcción de un cuerpo teórico que avanza hacia una posición epistémica y pragmática de la explicación, cuyos antecedentes se encuentran en las ideas de Scriven y de van Fraassen. A través de sus críticas, Scriven parece dirigirnos a un retorno hacia la forma en que cotidianamente se entienden las explicaciones, lo que nos puede invitar a una representación más adecuada de ellas por vía de la recuperación de sus usos frecuentes. El vínculo entonces entre la noción de explicación y las representaciones del razonamiento del sentido común es pertinente a este respecto.

Para Scriven, un aspecto fundamental de este giro teórico es, justamente, la distinción entre explicación y fundamentos de la explicación. Otro aspecto importante en este camino, que yo añadí, es la aceptación de las condiciones epistémicas y de falibilidad en que son presentadas las explicaciones en el contexto científico real. La integración en un modelo, de aspectos que se consideraron en la discusión como problemáticos, como el efecto no monotónico de algunas nociones de inferencia o su relativización epistémica, es una consecuencia de tomar esta vía de abordaje al problema de la modelación de la explicación científica.

Van Fraassen, por su parte, estableció un énfasis particular sobre lo que consideró la dimensión pragmática del problema. Esta posición lo llevó rápidamente a un

contextualismo que creo que ha sido un gran acierto. No obstante, su posición pragmática arribó también, como lo hicieron notar lúcidamente Kitcher y Salmon, a un relativismo, no solo en lo que toca a la interpretación de la intención de una pregunta que inquiriere por una explicación, sino también y más peligrosamente sobre la propia noción de relación de pertinencia explicativa.

En el planteamiento de una noción epistémica y pragmática de explicación, intenté seguir a estos dos autores, pero con algunas restricciones. Una de las restricciones que intenté construir en el caso de Scriven, fue mantener la estrategia de elucidación de la explicación basada en un modelo argumental. La restricción más importante que intenté oponer a la línea de solución van fraasseana fue lo que considero un intento por sostener un contextualismo, sin alcanzar a un relativismo. La clave para mi propuesta de solución (contextualismo sin relativismo) fue mantener fija, en el dominio intensional, nuestra noción de relación de pertinencia explicativa, es decir, proponer tener una sola definición intensional para ésta, al mismo tiempo que permitimos la variación, en el dominio extensional, de esta misma relación. En otras palabras, mi sugerencia fue mantener sin variación intensional un conjunto relaciones de pertinencia explicativa, pero permitir variación en los pares de objetos que, desde cada sistema epistémico, sean considerados como pares *relata* (variación extensional) de esta relación de pertinencia. Esta solución, desde mi punto de vista, es una salida al problema de cómo aceptar el contextualismo (consecuencia de nuestra posición pragmática) al mismo tiempo que rechazar un relativismo respecto de lo que sea la explicación. Creo que soluciones análogas pueden establecerse en diversos problemas de la racionalidad científica, por ejemplo, creo que puede haber una aplicación de ello en el problema de aceptar la idea de sociedades del conocimiento en el contexto de una relajación de los criterios para reconocer algo como conocimiento [Olivé 2007]²⁸⁹. Aunque, en este respecto, hará falta un estudio más detallado de la racionalidad científica con todos los bemoles que han sido planteados en

²⁸⁹ [Olivé 2007] Olivé, León, *La Ciencia y la tecnología en la sociedad del conocimiento: ética, política y epistemología*. FCE, México, 2007.

las discusiones de las últimas 3 décadas alrededor de dos nociones, la de comunidad epistémica y la de escepticismo.

Uno de los costos de una posición no deductivista, relacionada con la dimensión epistémica de la explicación de la que he hablado en el párrafo anterior, fue la aceptación de argumentos con baja probabilidad. Si las razones presentadas a favor de una posición sobre los argumentos de esta naturaleza son correctas, las repercusiones en Teoría de la Argumentación pueden ser muy interesantes. La idea en el contexto de esta tesis fue intentar presentar una noción de argumento que no se convirtiera en una narración trivial, y que al mismo tiempo reconociera la importancia de lo que van Fraassen interpretaba como “razones para creer”, que es uno de los pilotes que sostienen la parte menos relativista de su posición, pero que permitiera al mismo tiempo la representación de baja probabilidad en la inferencia. Convendrá también explorar, en un contexto mayor, cómo se comporta la defensa de esa idea de argumento que aparece en esta tesis, en una discusión mucho más amplia y detallada. Vinculado a lo anterior es importante mencionar que el aspecto que he llamado “modular” del modelo que presenté en esta investigación, o para decirlo más directamente, su capacidad para combinar diversos elementos y analizar sus efectos, además de ser una de las claves para la capacidad de análisis que tiene ese marco formal, es una perspectiva general que puede ser adoptada también para el estudio y la enseñanza de la argumentación en general. Un intento de esta aplicación lo es [Gaytán 2004], y considero que una perspectiva de este tipo podría contribuir al problema que en Lógica Informal, Johnson llamó “The Network problem” en [Johnson 1996]. Una posible repercusión de la misma perspectiva de abordaje que sugiero para este problema, es decir con este aspecto modular, podría tener frutos, como en otros contextos he dicho, para una teoría unificada de las condiciones de corrección de los argumentos deductivos y no deductivos.

La idea de un argumento de baja probabilidad puede representarse muy bien mediante la idea de regla *default*, de Reiter. Pero eso nos llevó al objetivo de fortalecer los

vínculos del mecanismo interno de una regla como ésta. A este respecto las reflexiones que hice sobre esta lógica de Reiter intentaban dotar de sentido epistemológico a las construcciones lógicas de este autor. Estas reflexiones fueron importantes para definir el sentido en el que debía fortalecerse la justificación de un *default* y pueden verse como aportaciones a una idea básica de razonamiento del sentido común, por ejemplo, la idea de que las justificaciones usadas en varios *default* tienen que participar en el chequeo de consistencia o de no trivialización. El camino de esta clase de reflexiones epistemológicas acerca de estas nuevas lógicas señala que podría haber áreas muy interesantes de investigación que exigen una vinculación entre los estudios de las lógicas no monotónicas, los de Filosofía de la Ciencia y, como ya algunos lógicos de la no monotonicidad lo habían señalado, los estudios epistemológicos. A este respecto es muy valiosa la aportación de Morado al estudio de los supuestos de racionalidad de las lógicas no monotónicas [Morado 2004]. Así, la investigación que presento ahora intentó, de manera central, integrar por un lado propuestas lógicas originales, como la combinación entre *defaults* y paraconsistencia basada en una jerarquía de da Costa y en la noción de modalidad paraconsistente y, por otro lado, soluciones a problemas planteados filosóficamente, como atender el problema de la relación de pertinencia explicativa. Atender esta clase de problemas filosóficos es algo que considero casi completamente ausente en los estudios contemporáneos en explicación que se hacen desde un punto de vista lógico. Y en este sentido, considero a esta investigación doctoral como una aportación a la vinculación entre Inteligencia Artificial y Filosofía de la Ciencia, pero creo que hace falta una mayor interacción entre estas dos interesantes disciplinas de conocimiento.

Finalmente, una de las consecuencias de optar por suponer una racionalidad menos dura en la actividad científica es que también deberíamos poder integrar el problema de la inconsistencia en el desarrollo de la investigación científica. A este respecto en esta tesis intenté asumir una racionalidad que incluyera la paraconsistencia, pero en el tenor de Newton da Costa y los desarrollos brasileños que lideran pensadores como Ítala D'Ottaviano o Walter Carnielli en esta área lógica de investigación. La idea

general, según yo la interpreto, es preservar lo más posible una racionalidad constituida por cánones que nos lleven a mantener distinciones útiles para entender el progreso en el conocimiento. En otras palabras, entender la inconsistencia como algo que ocurre, que es manejable desde el punto de vista formal, o incluso que es útil para modelar ciertos efectos; pero no como un objetivo *prima facie* del conocimiento, o por lo menos no de todo conocimiento. En esta línea de pensamiento, intenté incluir en mi enfoque teórico el fenómeno de la paraconsistencia, sobre todo en el caso de lo que llamé “paraconsistencia de transición”. Así, propuse diversas formas de interacción entre No monotonicidad, explicación y paraconsistencia. Hará falta, sin embargo, una construcción más elaborada de estas interacciones y, también una mayor profundidad en la investigación de la paraconsistencia como una noción que puede tener muchos frutos a la hora de discutir sobre muy diversos aspectos de la racionalidad científica. En esta línea de investigación es importante recuperar los aportes del texto sobre el conocimiento científico, de Newton da Costa [da Costa 2000]. También de los estudios sobre la llamada “verdad pragmática” en [da Costa & Bueno & French 1998] o en [D’Ottaviano & Hifume 2007], y las investigaciones sobre la lógica discusiva de Jaśkowski en [da Costa & D’Ottaviano 1970].

Referencias Bibliográficas

1. **[Achinstein 1981]** Achinstein, Peter, "Can there be a modelo f explanation?" en *Theory and Decision*, 13, pp. 201-27, 1981.
2. **[Achinstein 1985]** Achinstein, P., *The Nature of Explanation*, Oxford University Press, 1985.
3. **[Arruda & Chuaqui & Da Costa 1980]** Arruda, A. I., Chuaqui, R., Da Costa, N. C. A. "Mathematical Logic in Latin America, en: Barwise, J., Kaplan, D., *Studies in Logic an the foundations of Mathematics (Volume 99)*, North-Holland Publishing Company. Netherlands: 1980.
4. **[Arruda 1989]** Arruda, Ayda, "Aspects of the Historical Development of Paraconsistent Logic", en Priest, Graham; Routley (ahora Sylvan), Richard; Norman, Jean; *Paraconsistent Logic. Essays on the Inconsistent*. Philosophia Verlag, Alemania, 1989.
5. **[Arruda, 1988]** Arruda, Ayda, "Panorama de la Lógica Paraconsistente", en *CELIJS, Antología de la lógica en America Latina*, Universidad de Carabobo, Valencia, Venezuela, 1988.
6. **[Bell 1985]** (1940) Bell, E. T., *Historia de las Matemáticas*, Fondo de Cultura Económica, México, 1985, p. 497. Traducción del original: *The Development of Mathematics*, McGraw-Hill Book Co., Nueva York, 1940.
7. **[Brewka & Dix & Konolige, 1977]**, Brewka, Gehard & Dix, Jürgen & Konolige, Kurt, *Nonmonotonic Reasoning*, CSLI Publications, E. U. A., 1997.
8. **[Brody, 1972]** Brody B, "Towards Aristotelian theory of Scientific Explanation", en *Philosophy of Science*, 39, pp. 20-31, 1972.
9. **[Brody, 1993]** Brody B, " Towards An Aristotelian Model of Scientific Explanation" en David Hillel Ruben *Explanation* Oxford University Press, 1993, pp 13-127.
10. **[Carnap, 1950]** Carnap, Rudolf, *Logical Foundations of Probability*, University of Chicago Press, Chicago, 1950.
11. **[Carnielli & Coniglio & Marcos 2000]**, Carnielli, Walter A. & Coniglio, Marcelo E. & Marcos, Joao, "A taxonomy of C-Systems", en Walter A. Carnielli & Ítala M. L. D'Ottaviano & Marcelo E. Coniglio, *The Logical Way to the Inconsistent. Proceedings of the World Congress of Paraconsistency*. State University of Campinas, Edit. Marcel Dekker Inc., Sao Paulo, Brasil, 2000.
12. **[Carnielli & D'Ottaviano & Coniglio 2000]** Carnielli, Walter A. & D'Ottaviano, Ítala M. L. & Coniglio, Marcelo E., *The Logical Way to the Inconsistent. Proceedings of the World Congress of Paraconsistency*. State University of Campinas, Edit. Marcel Dekker Inc., Sao Paulo, Brasil, 2000.
13. **[Coffa 1974]** Coffa, Alberto J., "Hempel's Ambiguity", en *Synthese*, 28, 1974, pp. 141-63.

14. **[Coffa 1993]** Coffa, Alberto J., "Hempel's Ambiguity", en David-Hillel Ruben (ed.), *Explanation*, Oxford University Press, E.U.A. 1993. Reimpresión de "Hempel's Ambiguity", en *Synthese*, 28, 1974, pp. 141-63.
15. **[Da Costa, 1974]** Da Costa N.C.A., "On the theory of Inconsistent Formal Systems", en *Notre Dame Journal of Formal logic, NDFL*, V.15, N.4, 1974.
16. **[Da Costa & Alves 1977]** Da Costa N.C.A., & Alves, E.H., "A Semantical Analysis of the Calculi Cn", *Notre Dame Journal of Formal logic, NDJFL*, V.18, N.4, 1977.
17. **[Da Costa, 2000]** Da Costa, N. C. A., *El Conocimiento Científico*, Paidós-IF UNAM, México, 2000.
18. **[Da Costa & Marconi, 1989]** Da Costa, Newton and Marconi, Diego, "An overview of Paraconsistent Logic in the 80s", en *The Journal of Non-classical Logic (TJNCL)*, V. 6, n. 1, 1989.
19. **[Da Costa & Dubikajtis, 1968]** Da Costa, Newton C. A., y Dubikajtis, L., "Sur la logique discursive de Jaśkowski" en *Bull Acad. Polonaise des Sciences*, XVI, 1968, pp. 551-557.
20. **[Da Costa & Bueno & French 1998]** Da Costa, N.C.A. & Bueno, Otávio & French, Steven, "The Logic of Pragmatic Truth", *Journal of Philosophical Logic*, 27(6), pp. 603-620, 1998.
21. **[Dieudonné, 1989]** Dieudonné, Jean, *En Honor del Espíritu Humano. Las Matemáticas Hoy*, Alianza Universidad, Madrid, 1989.
22. **[D'Ottaviano & da Costa 1970]** D'Ottaviano, Ítala M. Loffredo, & da Costa, N.C.A., "Sur un probleme de Jaśkowski", en *Comptes Rendus Acad. Sci.*, 270, 1970.
23. **[D'Ottaviano 1990]** D'Ottaviano, Ítala M. L., "On the Development of Paraconsistent Logic and da Costa's Work", en *The Journal of Non-Classical Logic (TJNCL)*, Vol.7, No. 1-2, 1990.
24. **[D'Ottaviano & Feitosa 2000]** D'Ottaviano, Ítala, M. Loffredo & Feitosa, Hércules, "Paraconsistent Logics and Translations", *Synthese*, Vol. 125, No. 1/2, Festschrift in Honor of Newton C. A. Da Costa on the Occasion of His Seventieth Birthday (Oct. - Nov., 2000), pp. 77-95.
25. **[D'Ottaviano & Hifume 2007]** D'Ottaviano Ítala M. Loffredo & Hifume, Carlos, "Percean Pragmatic Truth and da Costa's Quasi-Truth", *Model Based Reasoning in Science, Technology and Medicine*, 2007, pp. 383-398.
26. **[Fidel, 1977]** Fidel, Manuel M., "The decidability of the calculi Cn", en *Reports on Mathematical Logic*, 8, 1977, pp. 31-40.
27. **[Frege, 1884]** Frege, Gottlob, *Die Grundlagen der Arithmetik*, 1884. Una traducción parcial al castellano está en *Conceptografía - Fundamentos de la Aritmética - Otros Estudios*, Trad. de Hugo Padilla, Colección Filosofía Contemporánea, Instituto de Investigaciones Filosóficas, UNAM, México, 1972.
28. **[Frege, 1893]** Frege, Gottlob, *Grundgesetze der Arithmetik*, vol. 1, H. Pohle, Jena, 1893. Una traducción parcial al inglés es *The Basic Laws of Arithmetic: Exposition of the System*, Trad. de M. Furth. University of California Press, Berkeley y Los Angeles, 1964.
29. **[Frege 1964]** Frege, Gottlob, *The Basic Laws of Arithmetic: Exposition of the System*, Trad. de M. Furth, University of California Press, Berkeley y Los Angeles, 1964.

30. **[Frege 1972]** Frege, Gottlob, *Conceptografía - Fundamentos de la Aritmética - Otros Estudios*, Trad. de Hugo Padilla, Colección Filosofía Contemporánea, Instituto de Investigaciones Filosóficas, UNAM, México, 1972.
31. **[Garciadiego 1992]** Garciadiego Dantan, Alejandro R., *Bertrand Russell y los Orígenes de las paradojas de la Teoría de Conjuntos*, Alianza Universidad, Madrid, 1992.
32. **[Gärdenfors, 1980]** Gärdenfors, Peter, "A Pragmatic Approach to Explanations", *Philosophy of Science*, Vol. 47, No. 3 (Sep., 1980), pp. 404-423.
33. **[Gaytán, 2005]** Gaytán Cabrera David "Una Solución al problema Hempeliano de las salvedades" en Ana Rosa Pérez Ransanz, *Perspectivas y horizontes de la filosofía de la ciencia a la vuelta del tercer milenio*, Vol II UNAM/Filosofía de la Ciencia, México, 2005, pp 95-120.
34. **[Gaytán, 2007]** Gaytán Cabrera David "Notas para un modelo formal de explicación" en Jorge Martínez C. & Aura Ponce de León *El Saber Filosófico. Sociedad y Ciencia*, Asociación Filosófica de México siglo XXI, México, 2007, pp 291-312.
35. **[Ginsberg, 1987]** Ginsberg, Matthew L., (comp.), *Nonmonotonic Reasoning*, Morgan Kaufmann Publishers, E.U.A., 1987.
36. **[Glymour, 1998]** Glymour, Bruce, "Contrastive, Non Probabilistic Statistical Explanation", *Philosophy of Science*, 65, 1998, pp. 448-471.
37. **[Gomes 2013]** Gomes, Evandro Luis, *Sobre a História da Paraconsistência e a obra de da Costa: a instauração da Lógica paraconsistente*, tesis doctoral defendida el 12 de diciembre de 2013, CLE- Universidade Estadual de Campinas, Brasil.
38. **[Grana 1983]** Grana, Nicola, *Lógica Paraconsistente. Una Introducción*, Loffredo Editore Napoli, Napoli, 1983.
39. **[Greeno 1971]** Greeno, "Theoretical Entities in Statistical Explanation" en Roger C. Buck y Robert S. Cohen (eds.), *PSA 1970*, D. Reidel Publishing Co., Dordrecht, 1971.
40. **[Heijenoort 1967]**, Heijenoort, Jean van, *From Frege to Gödel*, Harvard University Press, Cambridge, Massachusetts, 1967.
41. **[Hempel & Oppenheim 1948]** Hempel, Carl G. & Oppenheim, Paul, "Studies in the Logic of Explanation", *Philosophy of Science*, vol. 15, 1948, pp. 567-579.
42. **[Hempel & Oppenheim 1988]** Hempel, Carl G. y Oppenheim, Paul, "Studies in the Logic of Explanation", en Joseph C. Pitt, *Theories of Explanation*, Oxford University Press, Nueva York y Oxford, 1988, pp. 9-50.
43. **[Hempel 1965]** Hempel, Carl, G., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press y Collier Macmillan, Nueva York y Londres, 1965.
44. **[Hempel 1977]** Hempel, Carl G., *Filosofía de la Ciencia Natural*, Alianza Universidad, Madrid, 1977.
45. **[Hempel 1979]** Hempel, Carl G., *La Explicación Científica*, Edit. Paidós-Argentina, 1ª edición, Buenos Aires, 1979.
46. **[Hempel 1988]** Hempel, Carl G., "Provisoes: A Problem Concerning the Inferential Function of Scientific Theories", en *Erkenntnis* 28, 1988, pp. 147-164.

47. **[Hintikka & Halonen, 1999]** Hintikka, Jaako & Halonen, Ilpo, "Interpolation as Explanation", *Philosophy of Science*, 66, 1999, pp. S414-S423.
48. **[Jaśkowski 1948]** Jaśkowski, S., Rachunek zda dla systemów dedukcyjnych spreecznych, publicado bajo el título "Un Calcul des Propositions pour les Systemes Déductifs Contradictaires" en *Studia Societatis Scientiarum Torunensis*, vol. 1 (5), sectio A, pp. 55-77.
49. **[Jaśkowski 1949]** Jaśkowski, S., O konjuncji dyskusynej w rachunku zdán dla systemów dedukcyjnych spreecznych, publicado bajo el título "Sur la conjonction Discursive das le Calculu des propositions pour les Systemes Contradictaires" en *Studia Societatis Scientiarum Torunensis*, vol. 1 (5), sectio A, pp. 171-2.
50. **[Kitcher & Salmon 1989]** Kitcher, P. & Salmon, W. C. (eds.), en *Scientific Explanation*, Minnesota Studies in the Philosophy of Science, Vol. XIII, 1989.
51. **[Kitcher & Salmon, 1993]** Kitcher, Phillip & Salmon Wesley C., "Van Fraassen on Explanation" en David-Hillel Ruben (ed.) *Explanation*, Oxford University Press, EEUU, pp. 314,1993.
52. **[Kitcher 1981]** Kitcher, Phillip, "Explanatory unification", en *Philosophy of Science*, 48, 1981, pp. 507-531.
53. **[Kline, 1972]** Kline, Morris, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, Oxford, 1972.
54. **[Kline 1992]** Kline, Morris, *El Pensamiento Matemático de la Antigüedad a nuestros Días*, tomo III, Alianza Universidad, Madrid, 1992.
55. **[Kuhn 1962]** Kuhn, Thomas S., *The Structure of Scientific Explanations*, University of Chicago Press, Chicago, 1962.
56. **[Kuhn 1982]** Kuhn, Thomas S., *La Estructura de las Revoluciones Científicas*, F.C.E., México-Buenos Aires, 1982.
57. **[Kyburg 1965]** Kyburg, Henry E. Jr., "Comment", en *Philosophy of Science*, 32, pp. 147-51.
58. **[Lewin & Mikenberg 1991]** Lewin, R.A. & Mikenberg, I. "C1 is not algebraizable", *NDJFL*, V.32, N.4, 1991.
59. **[Lewis 1986]** Lewis, David, *Philosophical Papers II*, Oxford University Press, Oxford y Nueva York, 1986.
60. **[Lipton 1994]** Lipton, Peter, (1991), *Inference to the Best Explanation*, Routledge, 1994.
61. **[Marek & Truszczyński 1993]** Marek, V. Wiktor & Truszczyński, Mirosław, *Nonmonotonic Logic. Context-Dependent Reasoning*, Springer-Verlag, Berlin, 1993.
62. **[McCarthy & Hayes 1969]** McCarthy, John & Hayes, P. "Some Philosophical Problems from the Standpoint of Artificial Intelligence", en B. Meltzer & D. Michie (eds.), *It Machine Intelligence* Vol. 4, Edingburgh University Press, Edinburgh, Nueva York, 1969.
63. **[McCarthy 1977]** McCarthy, John, "Epistemological Problems of Artificial Intelligence", en *International Joint Conference on Artificial Intelligence (IJCAI)*, 1977, pp. 1038-1044.
64. **[McDermott & Doyle 1980]** McDermott, Drew & Doyle, Jon, "Non-monotonic logic I", en *Artificial Intelligence*, vol. 13 num. 1-2, 1980, pp. 41-72.

65. **[Minker 1991]** Minker, Jack, "An Overview of Nonmonotonic Reasoning and Logic Programming", en *First International Workshop on Logic programming and Nonmonotonic Reasoning*, Washington, D. C., Julio 23, 1991.
66. **[Miró 1988]** Miró Quesada, Franciso, "La lógica paraconsistente y el problema de la racionalidad de la lógica", en *CELIJS, Antología de la lógica en América Latina*, Universidad de Carabobo, Valencia, Venezuela, 1988.
67. **[Morado 2004]** Morado, Raymundo E., "Filosofía de las Lógicas No Monotónicas", en Raúl Orayen & Alberto Moretti (eds.) *Filosofía de la Lógica*, Enciclopedia Iberoamericana de Filosofía, Vol. 27, Trotta, España, 2004.
68. **[Morado 2003]** Morado, Raymundo E., "Rationality, Logic and Heuristics", en *International Conference in Artificial Intelligence*, Vol. 2, CSREA Press, 2003, pp. 791-797.
69. **[Olivé 2007]** Olivé, León, *La Ciencia y la tecnología en la sociedad del conocimiento: ética, política y epistemología*, FCE, México, 2007.
70. **[Orayen 1976]** Orayen, Raul, "Acerca de la adecuación de los modelos formales de la explicación científica", Instituto Torcuato di Tella. Reimpreso en *Antología de la lógica en América Latina*, C.E.L.I.J.S., Universidad de Carabobo, Valencia, Venezuela, 1988.
71. **[Orayen 1988]** Orayen, Raul, "Acerca de la adecuación de los modelos formales de la explicación científica", en *Antología de la lógica en América Latina*, C.E.L.I.J.S., Universidad de Carabobo, Valencia, Venezuela, 1988.
72. **[Orayen 1989]** Orayen, Raúl, *Lógica, significado y ontología*, UNAM, México, 1989.
73. **[Peña 1993]** Peña, Lorenzo, *Introducción a las Lógicas No Clásicas*, Colección Cuadernos (60), IIF-UNAM, México, 1993.
74. **[Pitt 1988]** Pitt, Joseph C., *Theories of Explanation*, Oxford University Press, Nueva York y Oxford, 1988.
75. **[Priest & Routley 1989]** Priest y Sylvan (antes Routley), "First Historical Introduction, A Preliminary History of Paraconsistent on Dialectic Approaches", en Priest, Graham & Routley, Richard & Norman, Jean (eds.), *Paraconsistent Logic. Essays on the inconsistent*, Philosophia Verlag, Alemania, 1989.
76. **[Priest & Routley 1989a]** Priest, G. & Routley (ahora Sylvan), R.; "Systems of Paraconsistent Logic" en Priest, Graham; Routley & Richard & Norman, Jean (eds.), *Paraconsistent Logic. Essays on the inconsistent*. Philosophia Verlag, Alemania, 1989a.
77. **[Priest & Routley, 1989b]** Priest, G. y Routley , R.; "An Outline of the History of (logical) Dialectic" en Priest, Graham & Routley, Richard; Norman & Jean; *Paraconsistent Logic. Essays on the inconsistent*. Philosophia Verlag Germany: 1989.
78. **[Priest & Routley, 1989c]** Priest, G. y Routley , R., "Applications of Paraconsistent Logic", en Priest, Graham & Routley, Richard & Norman, Jean; *Paraconsistent Logic. Essays on the inconsistent*, Philosophia Verlag, Alemania, 1989.
79. **[Priest & Routley 1989d]** Priest, Graham, y Routley, Richard, "The Philosophical Significance and Inevitability of Paraconsistency", en Priest, Graham; Routley,

- Richard; Norman, Jean; *Paraconsistent Logic. Essays on the inconsistent*. Philosophia Verlag Alemania, 1989, pp. 517 et s.
80. **[Priest & Routley & Norman, 1989]** Priest, Graham & Routley, Richard & Norman, Jean (eds.), *Paraconsistent Logic. Essays on the inconsistent*. Philosophia Verlag, Alemania, 1989, p. 158
 81. **[Priest, 1989]** Priest, Graham, "Reductio ad Absurdum et Modus Tollendo Ponens", en Priest, Graham & Routley & Richard (ahora Sylvan) & Norman, Jean (eds.), *Paraconsistent Logic. Essays on the Inconsistent*. Philosophia Verlag, Alemania, 1989.
 82. **[Railton, 1988]** Railton, Peter, "A Deductive-Nomological Model of Probabilistic Explanation", en Joseph Pitt, *Theories of Explanation*, Oxford University Press, E.U.A., 1988.
 83. **[Railton 1978]** Railton, Peter, "A Deductive-Nomological Model of Probabilistic Explanation", en *Philosophy of Science*, 45, 1978, pp. 206-226.
 84. **[Reiter, 1980]** Reiter, Raymond, "A Logic for Default Reasoning", en *Artificial Intelligence*, Vol. 13, 1980, pp. 81-123.
 85. **[Reiter & Criscuolo, 1981]** Reiter, Raymond & Criscuolo, Giovanni, "On Interacting Defaults", en Ann Drinan (ed.), *Proceedings of the seventh IJCAI Conference*, Vancouver, Canada, Agosto 1981, pp. 270-276.
 86. **[Reiter 1987]** Reiter, Raymond, "A Logic for Default Reasoning" en Matthew L., en Ginsberg, (comp.), *Nonmonotonic Reasoning*, Morgan Kauffmann Publishers, E.U.A. 1987.
 87. **[Resnik 1980]** Resnik, Michael D., *Frege and The Philosophy of Mathematics*, Cornell University Press, Reino Unido, 1980.
 88. **[Robles 1982]** Robles, José Antonio, *Bertand Russell: Antología I*, SEP - Setentas / Diana, México, 1982.
 89. **[Routley & Meyer 1976]** Routley (ahora Sylvan), Richard & Meyer, K., "Dialectical logic, Classical Logic and the Consistency of the World", en *Studies in Soviet Thought*, 16, 1976, pp. 1-25.
 90. **[Ruben, 1993]** Ruben, David-Hillel (ed.), *Explanation*, Oxford University Press, E.U.A. 1993.
 91. **[Russell 1944]** Russell, Bertrand, *My Philosophical Development*, Londres, George Allen y Unwin 1944.
 92. **[Russell 1964]** Russell, Bertrand, *La Evolución de mi Pensamiento Filosófico*, Aguilar, Madrid, 1964, pp. 80-81.
 93. **[Russell, 1975]** Russell, Bertrand, *Introduction to Mathematical Philosophy*, George Allen & Unwin Ltd, Londres, 1975.
 94. **[Salmon & Kitcher, 1987]** Salmon W. C. & Kitcher, P., "Van Fraassen on Explanation", en *Journal of Philosophy*, 84, 1987, pp. 315-30.
 95. **[Salmon 1965]** Salmon, "The Status of Prior Probabilities in Statistical Explanation", en *Philosophy of Science* 32, 1965, pp. 137-146.
 96. **[Salmon 1971]** Salmon, Wesley C., *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, 1971. Reimpreso parcialmente en "Statistical

- Explanation and Causality", de Joseph Pitt (ed.), *Theories of Explanation*, Oxford University Press, E. U. A., 1988.
97. **[Salmon 1984a]** Salmon, W. C., "Scientific Explanation: Three Basic Conceptions", en *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol.2: *Symposia and Invited Papers*, 1984, pp. 293-305.
 98. **[Salmon 1984b]** Salmon, W. C., *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, 1984.
 99. **[Salmon 1984c]** Salmon, W. C., *Four decades of Scientific explanation*, University of Pittsburgh, 1984.
 100. **[Salmon 1988]** Salmon, Wesley C., "Statistical Explanation and Causality", en Joseph Pitt, (ed.), *Theories of Explanation*, Oxford University Press, E. U. A., 1988.
 101. **[Salmon 1993]** Salmon, Wesley C., "Scientific Explanation and the Causal Structure of the World", en Ruben, David-Hillel (ed.), *Explanation*, Oxford University Press, E.U.A. 1993.
 102. **[Salmon 1994]** Salmon, Wesley C., "Causality without counterfactuals", en *Philosophy of Science*, 61, 1994, pp. 297-312.
 103. **[Sandewall 1972]** Sandewall, E., "An approach to the frame problem and its implementation", en Meltzer, B. y Michie, D. (eds.), *Machine Intelligence 7*, Edinburgh University Press, Edinburgo, 1972, pp. 195-204.
 104. **[Scriven 1959]** Scriven, Michael, "Explanation and prediction in Evolutionary Theory", en *Science*, No. 30, 1959, pp. 447-82.
 105. **[Scriven 1962]** Scriven, Michael, *Explanations, Predictions, and Laws*, University of Minnesota press, Minneapolis, 1962.
 106. **[Scriven 1988]** Scriven, Michael, *Explanations, Predictions, and Laws*, en Joseph Pitt, *Theories of Explanation*, Oxford University Press, E.U.A., 1988.
 107. **[Urbas,1988]** Urbas, Igor, "Paraconsistency and Subsystems of the system J1 of Arruda and da Costa", *TJNCL*, V.5, N.1, 1988.
 108. **[Urbas,1989]** Urbas, Igor, "Paracaonsistency and the C-Systems of da Costa", *Notre Dame Journal of Formal Logic*, (NDJFL) V.30, N.4, 1989.
 109. **[Van Fraassen 1996]** Van Fraassen, Bas C., *La Imagen Científica*, Trad. de Sergio Martínez, Col. Problemas Científicos y Filosóficos Paidós/IIF-UNAM, No. 2, Paidós, México, 1996.
 110. **[Van Fraassen 1980]** Van Fraassen, Bas C., *The Scientific Image*, Oxford University Press, 1980.
 111. **[Von Wright, 1971]** Von Wright, G. H., *Explanation and Understanding*, Cornell University Press, Ithaca, Nueva York, 1971.
 112. **[Weyl 1949]** Weyl, Hermann, *Philosophy of Mathematics and Natural Science*, Princeton University Press, Princeton, 1949.
 113. **[Weyl 1965]** Weyl, Hermann, *Filosofía de las Matemáticas y de la Ciencia Natural*, Centro de Estudios Filosóficos, UNAM, México, 1965.
-