INTRODUCCION A LOS FUNDAMENTOS DE LA TEORIA CIBERNETICA

Tesis Profesional

JESUS MIGUEL CHANFON OLMOS

MEXICO D. F. 1974

Universidad Nacional Autónoma de México
UNAM
Biblioteca Central
Dirección General de Bibliotecas de la UNAM

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.

1578

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE INGENIERIA

INTRODUCCION A LOS FUNDAMENTOS DE LA TEORIA CIBERNETICA

UNICO

Tesis

QUE PARA OBTENER EL TITULO DE
INGENIERO MECANICO ELECTRICISTA
P R E S E N T A

JESUS MIGUEL CHANFON OLMOS

MEXICO, D. F. 1970

F A C U L T A D D E I N G E N I E R I A

U. N. A. M.

I N T R O D U C C I O N

A L O S

F U N D A M E N T O S

D E L A

T E O R I A

C I B E R N E T I C A

J. M. C H A N F O N

1 9 7 0

FACULTAD DE INGENIERIA
Depto. de Exams.Profs.
Núm.40-
Exp.Núm.40/214.2/1.-

Universidad Nacional Autónoma de México

Al Pasante señor Jesús M. CHANFON OLMOS
P r e s e n t e .

En atención a su solici - tud relativa, me es grato transcribir a usted a continuación el tema que aprobado por esta Dirección propuso el señor profesor Ing. Daniel Barrios-Morales Frías, para que lo desarrolle como tesis en su examen profesional de Ingeniero MECANICO ELECTRICISTA.

INTRODUCCION A LOS FUNDAMENTOS DE LA TEORIA CIBERNETICA

"a) EL SISTEMA DINAMICO

1. La actividad de un sistema y su manejo simbólico
2. Los sistemas determinados y el transductor

b) LA INFORMACION

3. El concepto de probabilidad y los sistemas de Markoff
4. Naturaleza y medida de la información.

c) LA REGULACION

5. El sistema de comunicación y el concepto de entropía
6. El flujo de información y los procesos de regulación."

Ruego a usted tomar debida nota de que en cumplimiento de lo especificado por la Ley de Profesiones, deberá prestar Servicio Social durante un tiempo mínimo de seis meses como requisito indispensable para sustentar examen profesional; así como de la disposición de la Dirección General de Servicios Escolares, en el sentido de que se imprima en lugar visible de los ejemplares de la tesis, el título del trabajo realizado.

Muy atentamente,
"POR MI RAZA HABLARA EL ESPIRITU"
México,D.F. 26 de Enero de 1970
EL DIRECTOR

Ing. Manuel Paulín Ortiz

MPO'LMO'eag

No sabemos cuáles teorías son verdaderas y cuáles falsas, pero sí podemos saber en qué medida la aplicación de una hipótesis da resultados satisfactorios; la función de la ciencia debería agotarse en la previsión utilitaria del curso de los fenómenos, de modo de posibilitar una acción de eficacia óptima.

"Con frecuencia digo que cuando se puede medir aquello de lo que se habla, y expresarlo en números, se sabe algo respecto a ello; pero cuando no se puede expresar en números, el conocimiento es de una clase insuficiente e insatisfactoria; puede tratarse de los albores del conocimiento, pero escasamente se ha avanzado, en el pensamiento, al nivel de la Ciencia, cualquiera que sea la materia de que se trate".

LORD KELVIN.

"Es nuestro deber entonces, considerar el estado actual del Universo como un efecto de su estado anterior y como la causa de uno que lo sucederá. Una inteligencia que, para un instante dado, conociera todas las fuerzas de que está animada la Naturaleza y la situación respectiva de los entes que la componen, si además fuese suficientemente amplia como para someter estos datos al análisis, abarcaría en una misma fórmula los movimientos de los más grandes cuerpos del Universo y los del átomo más ligero; nada sería incierto para ella, y el porvenir como el pasado, estarían presentes a sus ojos".

PIERRE SIMON DE LAPLACE.

I N D I C E

Pág.

INTRODUCCION AL TEMA

PRIMERA PARTE: EL SISTEMA DINAMICO

I. LA ACTIVIDAD DE UN SISTEMA Y SU MANEJO SIMBOLICO 1

II. LOS SISTEMAS DETERMINADOS Y EL TRANSDUCTOR 12

SEGUNDA PARTE: LA INFORMACION

III. EL CONCEPTO DE PROBABILIDAD Y LOS SISTEMAS DE MARKOFF 31

IV. NATURALEZA Y MEDIDA DE LA INFORMACION 50

TERCERA PARTE: LA REGULACION

V. EL SISTEMA DE COMUNICACION Y EL CONCEPTO DE ENTROPIA 61

VI. EL FLUJO DE INFORMACION Y LOS PROCESOS DE REGULACION 83

CONCLUSIONES

REFERENCIA BIBLIOGRAFICA

INTRODUCCION AL TEMA.

1. Puede afirmarse que detrás de todo logro importante de la ciencia moderna, se encuentra presente la aplicación de un cuerpo de teorías que gradualmente han ido ganando terreno en los diversos campos del conocimiento. Esta nueva ciencia, o si se prefiere, este nuevo enfoque de las disciplinas científicas en su conjunto, cristaliza bajo la presión de la última Guerra Mundial, aunque tiene remotos antecedentes históricos que sin ir demasiado atrás se originan en ciertas ideas de Descartes y de Leibnitz, siguen un desarrollo más o menos irregular y culmina alrededor de los años 40's en dos de sus realizaciones concretas más espectaculares y representativas: las máquinas electrónicas de cálculo de alta velocidad y precisión, y diversos tipos de dispositivos de control automático.

2. Bautizada en 1947 por N. Wiener con el nombre de CIBERNETICA, es frecuente encontrar este término en obras científicas y técnicas de autores de la más diversa orientación intelectual, para aplicar sus principios, para elogiarla o para criticarla. Es oportuno aclarar aquí, que el término en cuestión no ha sido universalmente aceptado, y hay quienes prefieran emplear en su lugar nombres tales como: "teoría de la información", "teoría del control y la retroalimentación", "teoría de la comunicación", etc., por considerarlos más descriptivos, sin embargo, el término es etimológicamente adecuado ya que significa en griego "piloto" o "timonel", idea que es perfectamente adecuada al tema que trata. El término en sí tampoco es nuevo dentro de la ciencia, pues fue empleado por Ampere para designar a la ciencia del gobierno civil, según lo expresa en la parte II de su "Essai sur la philosophie des sciences" publicado en París en 1845.

3. Decíamos que con frecuencia se le nombra en la literatura científica actual, pero generalmente esto se hace sin mayores aclaraciones acerca de su verdadera naturaleza y alcances; y pudiéndose afirmar que se encuentra rodeada de una aureola de confusión y de malentendidos, lo que por lo demás es perfectamente explicable en función de su misma esencia y origen, pues siendo en parte una reacción contra las deficiencias de la

contínuamente creciente especialización, se entremezcló con otras ciencias y fue desarrollada simultáneamente desde diferentes frentes, por lo que no resulta extraño que se tengan nociones totalmente distintas acerca de ella según el punto de vista e intereses del autor que se consulte. Con todo, tras las múltiples formas en que la cibernética se pueda presentar, existen ciertos temas, ciertas líneas cardinales que constituyen la armazón de su estructura y en cuya comprensión radica la clave para emprender estudios más avanzados.

4. La CIBERNETICA tiene un triple aspecto, es primeramente la ciencia que estudia los sistemas dinámicos organizados, o sistemas simplemente, entendiendo por sistema a todo conjunto de elementos interactuantes que tengan las funciones de un todo. A su vez, cada elemento de un sistema se puede considerar como un subsistema, y llevar el análisis hasta la profundidad que se quiera. Ahora bien, prácticamente todo, o al menos todo aquello cuyo estudio resulte de interés, puede considerarse un sistema; ya sea un átomo, una célula viviente, un reloj, un animal, un grupo social humano, o el sistema solar; es un enfoque en verdad tan vasto, que en principio constituye una concepción del Universo como en su tiempo lo fueron las ideas de Galileo, de Descartes, de Newton o de Laplace.

5. Hay que recalcar que un sistema solamente es tal en la medida en que sus elementos tienen relaciones definidas entre sí, ya que por ejemplo, un reloj cuyas piezas se han desmontado no pasa de ser un inútil montón de objetos metálicos; además, cabe hacer aquí una aclaración importante: no se trata de antropomorfizar a las máquinas ni de mecanizar a los seres vivos, si la cibernética trata indistintamente con los mismos métodos a ambos, es simple y llanamente porque ambos son sistemas organizados. Un automóvil y una piedra, pongamos por caso, son claramente distintos y sin embargo el movimiento de ambos queda descrito por las leyes del movimiento de Newton, porque ambos poseen masa. Difícilmente podría imitarse la vida y sus procesos con los conocimientos tan deficientes que se tienen acerca de su naturaleza esencial y de sus condiciones, en efecto, dada la carencia de un concepto operante que defina y precise la naturaleza esencial de los fenómenos biológicos, no tiene sentido el afirmar o negar que mediante tal o cual recurso se ha producido o siquiera

imitado tal tipo de actividad. Destacamos este punto, porque ha sido el origen de una constante, violenta y estéril polémica entre especialistas en diversas áreas, referente a lo que se supone que puede o no puede lograr la cibernética. Creemos que en el fondo de todo se encuentra una lamentable confusión debido precisamente a la ambigüedad de las nociones empleadas, particularmente, en lo que respecta a la distinción entre lo que es la función de un sistema, o sea, aquello para lo que sirve y que desarrolla como actividad característica; la estructura de que se vale para realizar esta función; y la naturaleza esencial de las diversas funciones y estructuras. El tema es amplio, pero sólo queremos llamar la atención hacia el hecho de que en primer lugar el objetivo central de la cibernética no es la imitación de la vida en cuanto tal, ya que como se expuso, no sabemos exactamente lo que es ni en lo que consiste. En segundo lugar, que lo que sí se ha podido reproducir, y esto con gran eficiencia, han sido algunas funciones de ciertos sistemas que existen en la naturaleza como partes integrantes de organismos vivientes, tales como el vuelo, la detección de obstáculos a distancia, la navegación submarina, etc. En tercer lugar, que puede lograrse el mismo tipo de función mediante sistemas que tengan estructuras diferentes, por ejemplo, las cuatro patas de un caballo y las dos ruedas de una motocicleta tienen estructuras y apariencias e incluso modos de funcionar radicalmente distintos, pero no obstante, tienen la misma función primaria: el desplazamiento en el espacio. Igualmente, una neurona y un elemento biestable fabricado por el hombre no tienen absolutamente nada en común, excepto su función. Y naturalmente que en lo que se puede tener interés es en la reproducción de la función y no de la estructura. Recuérdense los esfuerzos infructuosos en los inicios de la aviación por reproducir el vuelo de las aves en todos sus detalles, incluyendo el batir de las alas, y el triunfo de la misma al renunciar a este tipo de imitación y separar el plano de sustentación, o sea las alas, del elemento motriz, la hélice.

6. Con la comprensión y la aplicación del proceso de retroalimentación, y con criterios biológicos de ascendencia aristotélica tales como la actividad teleológica y la auto-regulación, surgió toda una colección de perros, tortugas y demás animales sintéticos que buscan tenazmente una

finalidad que puede consistir, por ejemplo, en llegar hasta una fuente luminosa que actúa como estímulo y a la cual se aproximan, incluso teniendo que rodear toda clase de obstáculos, imitando los tropismos de ciertos organismos biológicos elementales. Por una parte han sido estos animales sintéticos, así como cierto éxito en la reproducción del mecanismo de los reflejos condicionados, o sea, de la sustitución de estímulos por simple asociación de los mismos en el espacio y en el tiempo, que les ha conferido a estos "animales" cierta capacidad de "aprendizaje"; y por otra parte, el desarrollo de circuitos "lógicos" y unidades de "memoria", dentro de la computación electrónica, que permiten la realización de complejos razonamientos, lo que ha llevado a algunos a poner seriamente en duda la adecuada clasificación de estos sistemas artificiales. Dejaremos de lado esta cuestión que todavía deja demasiado margen a toda clase de opiniones que dependen de las nociones que se tangan acerca de muchos temas, y puntualicemos finalmente algo que tiene mayor importancia que el asombrar a los visitantes de una exposición mundial con juguetes electrónicos. La observación de la Naturaleza y particularmente del modo de funcionar de los sistemas integrados a los seres vivos, es una de las fuentes de inspiración más sugestivas, y bajo esta idea se ha fundado desde hace ya varios años, la BIONICA, ciencia que se ocupa de explotar sistemáticamente estas posibilidades; y es que en realidad no interesa tanto el saber si una cámara de televisión puede "ver" efectivamente una escena, pongamos por caso, o si puede en qué sentido lo hace, cuanto saber que reproduce en muchos aspectos el mecanismo visual de los animales superiores y del hombre, lo que sí se puede constatar por el hecho de que la comprensión de tal mecanismo, ha permitido la compensación precisa de defectos visuales mediante el empleo de lentes y demás recursos que proporciona la óptica y otras ciencias. Y para terminar con este aspecto, no podemos dejar de mencionar la posibilidad de diseñar toda clase de prótesis, no de finalidad estética sino funcional y verdaderamente integradas al organismo, terreno en el cual ya se ha adelantado bastante, para luego explorar las posibilidades de diseñar órganos internos artificiales que eventualmente suplanten a las recientes y no poco tenebrosas técnicas de trasplante con su secuela de complicaciones.

7. El segundo aspecto que tiene la cibernética, es su carácter de un movimiento de síntesis dentro de las ciencias, que como decíamos, tienden a fraccionarse y a ganar en profundidad en detrimento de su eficacia. No combate a la especialización que por lo demás es necesaria, sino suple sus deficiencias, no se trata tampoco de saber un poco de todo, sino de abarcar con una misma metodología principios fundamentales análogos y comunes a todas las ciencias, tal como la tendencia universal a alcanzar un estado de equilibrio presente en la ley electromagnética de Lenz, en la ley de van't Hoff-Le Chatelier vigente en las reacciones químicas, en el concepto fisiológico de homeostásis introducido por W. Cannon, en la tercera ley de Newton de la mecánica, en la ley de la oferta y la demanda estudiada en economía, en las formaciones reactivas estudiadas por Adler en su sistema de sicología, en el equilibrio biológico de las especies, etc. etc.. O bien, la unidad esencial de principios que rigen la elección de la actitud óptima en situaciones tan dispares como la mesa de juegos de habilidad, la actividad de un diplomático en la guerra fría, o de un estratega militar, para nombrar sólo los ejemplos más comunes y evidentes.

8. Por último, la cibernética es un método de análisis científico extraordinariamente poderoso y especialmente apto para estudiar el comportamiento de sistemas de complejidad extrema, en donde los métodos tradicionales son punto menos que inútiles; y es aquí en donde aparece la fusión con los puntos de vista de la mecánica estadística y el cálculo probabilístico.

9. Las ideas de la cibernética se están difundiendo amplia y rápidamente, y además de su interés intrínseco dentro de la ciencia y la tecnología en general, es indiscutible que sus aplicaciones más características se encuentran vinculadas con la actividad profesional del ingeniero, a quien además le proporciona un excelente terreno para realizar el ideal del ingeniero como un profesionista tan alejado del tecnicismo rutinario que lleva al "obrerismo" intelectual, como de la ciencia pura que supuestamente halla sus fines en el conocimiento mismo, actitud que regularmente desemboca en la esterilidad de investigaciones lamentablemente inútiles. Pero por otra parte, es también un hecho que la gran mayoría del material bibliográfico disponible a la fecha puede catalogarse a grosso modo en

dos categorías: por un lado las obras de divulgación que divierten, pero que contienen demasiadas concesiones y deformaciones simplificadoras que impiden extraer algún material utilizable, y por otro lado, las obras especializadas de nivel tan elevado que resultan inaccesibles para la preparación de un recién egresado promedio de una escuela de ingeniería o de ciencias, pues dan por conocidas una gran cantidad de nociones y conceptos de importancia fundamental; y fue precisamente al tratar de entender el último tipo de libros citado, cuando nos encontramos en la necesidad de investigar esos conocimientos presupuestos y que los autores no se detienen a aclarar a pesar de ser imprescindibles para una comprensión cabal del tema; el resultado está contenido en este trabajo, que constituye una amalgama formada con el material extraído de una amplia bibliogafía cuyos autores más importantes se presentan en una referencia someramente comentada que damos al final. En particular, la mayor parte de las ideas originales expuestas son debidas al pensamiento de N. Wiener, C. Shannon, R. Ashby y J. von Neumann, tomadas de obras que ya son clásicas.

10. Este trabajo tiene entonces un objetivo bien definido: hacer una presentación a nivel absolutamente elemental pero tan rigurosa como sea posible, de los temas principales sobre los que descansa la cibernética, de manera que quede bien clara la naturaleza de cada concepto, su justificación, y su relación con el resto de ideas; pudiendo constituir así un primer escalón en su posterior exploración a fondo. La labor personal está presente en la selección y coordinación del material investigado, y en la aportación de algunos puntos de vista y ejemplos ilustrativos simples, con la intención de presentar un cuadro congruente y utilizable y con especial cuidado en la claridad expositiva, que desgraciadamente no siempre caracteriza a las obras mencionadas; pero sin que este trabajo pretenda ser ni más ni menos de lo que su título indica: una introducción a los fundamentos de la cibernética, manejada y presentada en el nivel y con los instrumentos conceptuales de un pasante de nuestra Facultad, realizada con el deseo de que sea de alguna utilidad para todos aquellos que tengan interés o al menos curiosidad por esta moderna visión de la ciencia de extraordinario alcance y extensión, y que tengo la seguridad ha de matizar todos los aspectos del mundo del futuro.

J. M. Chanfón.
30/IV/69.

P R I M E R A P A R T E

E L S I S T E M A D I N A M I C O

I

LA ACTIVIDAD DE UN SISTEMA Y SU MANEJO SIMBOLICO.

1. La actividad profesional del ingeniero se encuentra relacionada con el diseño, mantenimiento y operación de sistemas físicos cuyo funcionamiento reporta algún efecto útil. Ahora bien, para que este funcionamiento sea efectivamente útil, se requiere que sea controlado, es decir, que la magnitud y características de la salida, efecto o respuesta se mantenga dentro de ciertos límites prefijados, límites que pueden ser constantes, o variables en el tiempo. Por otra parte, esta idea de regulación y de control se encuentra inseparablemente unida a la de flujo de información, ya que este es el factor que permite que una actividad en que intervienen varias entidades, resulte coordinada.

2. Y son precisamente los sistemas, la comunicación entre ellos y su control los temas centrales de la cibernética, sin embargo, hay que tener presente que al hacer este estudio nos desentendemos completamente de tal o cual dispositivo particular, para considerar únicamente a los sistemas en su abstracta generalidad. Asímismo, al analizar el proceso de comunicación, no nos interesa el equipo físico mediante el cual se realiza tal proceso, sino el proceso mismo de información, su naturaleza y los principios que lo rigen.

Además, no es menos importante el puntualizar el preciso papel de las matemáticas, que aquí consiste estrictamente en ser fino instrumento de análisis y adecuado medio de expresión; de modo que puede ser alentador saber que la teoría de conjuntos, el cálculo de variaciones, la teoría de matrices, el cálculo de probabilidades, etc., no son en sí mismos partes constituyentes de la cibernética, lo que además nos justifica el empleo de versiones muy diluidas de los mismos, ya que en esta discusión de principios, lo más importante es la claridad del concepto. Igualmente, dada la dificultad de establecer y sostener generalizaciones absolutas que pudieran trascender los límites habituales de la ciencia, cada afirmación que se haga queda entendido que se refiere a su aplicación dentro de los límites del objetivo de este trabajo.

3. Todo aquello que constituye tema interesante de estudio se manifiesta como fenómeno, como variación o movimiento en un sistema, ya sea este natural o artificial.

Entendemos por sistema a un conjunto de objetos de cualquier naturaleza, que tienen relaciones definidas entre sí y que dan un todo funcionante que llamamos SISTEMA DINAMICO. El ESTADO DE UN SISTEMA es el valor que tienen sus características en un momento dado, y es en general una entidad compuesta o vector, que puede consistir simplemente de una lista, tal como: posición, presión, temperatura, humedad, velocidad, etc. La sucesión de estados se percibe y define como ACTIVIDAD del sistema que la presenta; y es el registro de esta actividad la materia prima de la ciencia, y toda la complicación matemática de que se reviste la misma en las etapas avanzadas de su desarrollo tiene por finalidad única el describir, predecir y calcular esta actividad.

4. En efecto, toda ciencia natural es inicialmente descriptiva; la acumulación de datos experimentales hace posteriormente necesaria una ordenación de esos conocimientos, hasta llegar a una formulación simbólica del sistema estudiado, que permite un adecuado manejo conceptual del mismo. Esta representación simbólica del funcionamiento del sistema real constituye su modelo matemático; y es a partir de esta etapa cuando la ciencia física, y podemos suponer que potencialmente cualquier otra ciencia, queda en posibilidad de rendir sus máximos resultados. Precisamente en esta primera parte vamos a tratar de presentar, en la forma más simple y general posible, la manera en que se formula el modelo matemático de un sistema actuante cualquiera.

5. Consideremos un sistema que tenga la posibilidad de presentar cuatro estados diferentes, estados que identificaremos con los símbolos literales a, b, c, d; y por observación directa de su actividad formamos el siguiente registro de cambios de estado, en donde la flecha indica la transición correspondiente a cada etapa del ciclo.

$$a \Rightarrow c$$
$$c \Rightarrow d$$
$$d \Rightarrow b$$
$$b \Rightarrow a.$$

Lo anterior podemos escribirlo en forma más compacta como

$$\begin{array}{c|cccc} & a & b & c & d \\ \downarrow & c & a & d & b. \end{array}$$

Sistemas reales cuyo funcionamiento podría ser descrito de este modo son por ejemplo las diferentes señales luminosas de un semáforo que regula el tránsito, y que recorre constantemente un ciclo de cambios a intervalos prefijados de tiempo; o la sucesión de las cuatro estaciones del año, etc.

6. La tabla de transiciones descrita es por supuesto una forma de relación funcional. Al examinarla, distinguimos en ella dos conjuntos de elementos que podemos llamar X y Y respectivamente, y que son:

$$X \ (a, \ b, \ c, \ d),$$

$$Y \ (c, \ a, \ d, \ b).$$

Estos elementos están relacionados de tal modo, que a cada elemento del conjunto X corresponde un único elemento del conjunto Y. Esta relación de dependencia se indica

$$X \longrightarrow Y,$$

y se lee: "Y depende de X".

A los estados posibles, que constituyen los elementos del conjunto X, se les llama **OPERANDOS**.

Los estados asociados, que forman los elementos del conjunto Y, se les conoce como **TRANSFORMADAS**.

El agente que induce el cambio correspondiente a cada estado es el **OPERADOR**.

El conjunto de transiciones de cada operando a cada transformada es la **TRANSFORMACION**.

7. Una ley funcional puede representarse de varias formas:

a) mediante palabras,

b) como fórmula o ecuación,

c) por medio de una tabla,

d) a través de cualquier tipo de representación gráfica.

El caso más frecuente es el de representación por medio de una ecuación, en donde se tiene una sola expresión que contiene en sí todos los datos que daría una tabla de longitud que podría ser teóricamente infinita; esto se logra representando por medio de un solo símbolo a cualquiera de los elementos de un conjunto. En estas condiciones, "x" es el símbolo para cualquier elemento del conjunto X de operandos, y "y" el símbolo para cualquier elemento del conjunto Y de transformadas, es decir, "x" es la variable independiente, que en los casos que nos interesan va a ser casi siempre el tiempo, y "y" la variable dependiente o función; y los conjuntos X y Y constituyen el DOMINIO y el RANGO respectivamente.

La forma más común de representación funcional es sin duda la fórmula o ecuación, y es también la más útil y precisa, particularmente dentro de la física y ciencias afines, a quienes por esta razón se les llama a veces ciencias "exactas"; pero tiene la desventaja de que sólamente es aplicable en un número relativamente reducido de casos que admiten la cuantificación directa y su expresión numérica, y la técnica de su manejo es generalmente complicada.

Una forma más primitiva pero más versátil de representación es la forma tabular, que aunque no puede mostrar variaciones continuas, sí permite representar con mayor o menor aproximación un fenómeno observado de cualquier naturaleza.

8. Puesto que trataremos primordialmente con la forma tabular de representación funcional, veamos con más detalle cuales son algunas de sus características más importantes, así como las ventajas y limitaciones de su empleo.

Hay situaciones que pueden ser descritas completamente mediante una tabla, por ejemplo, sea A el conjunto de todos los enteros positivos cuyas terceras potencias no pasen de 200,

$$A \quad (1, 2, 3, 4, 5);$$

y hagamos corresponder a cada elemento de A con un elemento del conjunto B,

$$B \quad (1, 4, 9, 16, 25),$$

formado por el cuadrado de los elementos del conjunto A, de este modo:

1	2	3	4	5
1	4	9	16	25.

La tabla anterior contiene toda la información que se pueda pedir acerca de la función en cuestión.

Por el contrario, hay casos en donde una tabla sólo constituye una aproximación más o menos buena de la ley de dependencia funcional que describe, y hay que interpolar o extrapolar para situaciones no registradas en la tabla; tal sucede en los casos de variación continua, en donde dominio y rango son conjuntos infinitos y es imposible registrar en una tabla de extensión limitada todos los casos posibles. Los modelos o descripciones de fenómenos reales son todos de este tipo, ya que toda variación real se efectúa dentro del marco del continuo físico espacio-tiempo, y cuya descripción exacta, cuando ésta es posible, lleva a ecuaciones integro-diferenciales que, como sabemos, son en general difíciles de resolver, a pesar de ingeniosas técnicas de simplificación como la transformación de Laplace y demás métodos operacionales.

Por la razón expuesta, permaneceremos fundamentalmente en el nivel tabular de representación funcional, y aunque con estos recursos no es posible explorar en detalle todas las consecuencias de un hecho dado, como lo permite la aplicación del cálculo, ni es posible el diseño de ingeniería muchas veces crítico de equipo real, no obstante, no sólo es suficiente, sino también especialmente apto para una presentación de principios, e incluso, a pesar de su sencillez, para la investigación efectiva de una gran cantidad de fenómenos de diversa naturaleza.

Uno de los conceptos más importantes, dentro del tema que tratamos, es el de FUNCION o correspondencia, que puede expresarse con diferentes notaciones. Quizá la más apropiada sea:

$$f: X \longrightarrow Y$$

que se lee: "el conjunto X está aplicado según la ley f en el conjunto Y", lo que sugiere que el operador, dado por la ley f, aplica una transformación que lleva a cada elemento del conjunto X a un elemento correspondiente del conjunto Y; y es bajo esta idea de transformación que emplearemos el concepto de función en todo lo que sigue.

Cuando se manejan distintas transformaciones o funciones que tienen igual rango Y e igual dominio X, se distinguen entre sí empleando distintos símbolos para designar al operador o ley de dependencia, por ejemplo:

$$T: X \longrightarrow Y$$

$$g: X \longrightarrow Y$$

$$p: X \longrightarrow Y$$

$$\phi: X \longrightarrow Y$$

etc.

Consideremos una función cuya ley esté expresada por una ecuación tal como

$$y = 3x - 2;$$

podemos escribirla simplemente en la forma usual de

$$f(x) = 3x - 2,$$

o bien en las siguientes formas más extensas pero más descriptivas:

A) $f: X \longrightarrow Y$ es la función cuyos valores están dados por

$$f(x) = 3x - 2.$$

B) $f: (x, y)$ es la función cuyos pares ordenados son

$$(x, 3x - 2).$$

C) $\left[(x, y) \mid y = 3x - 2\right]$ que se lee: "el conjunto de todos los pares ordenados (x, y) tales que $y = 3x - 2$".

10. Si disponemos de tres conjuntos X, Y y Z; y de dos relaciones funcionales tales que

$$f: X \longrightarrow Y,$$

$$g: Y \longrightarrow Z,$$

observamos la posibilidad de una influencia de las "x" sobre las "z" a través de las "y"; esto es lo que llamamos una función de función, que puede expresarse como

$$z = g\left[f(x)\right],$$

o bien

$$g(f): X \longrightarrow Z.$$

Esta forma de aplicar una transformación a otra, da por resultado una tercera transformación que es la composición de las dos primeras. El caso particular en que la función operador es la misma que la función operando, da lugar a la formación de potencias sucesivas de la transformación dada.

Si $T(x)$ es la transformación dada, el resultado de operar reiteradamente sobre sí misma, genera sucesivamente las potencias

$$T\left[T(x)\right] = T^2(x),$$

$$T\left[T^2(x)\right] = T^3(x),$$

.....

$$T\left[T^{n-1}(x)\right] = T^n(x).$$

Ahora bien, la actividad continuada de un sistema aislado equivale, en nuestro modelo conceptual de naturaleza discreta (no continua), a la aplicación reiterada del mismo operador, correspondiendo cada aplicación a una ETAPA de funcionamiento.

11. Hay que notar que al vincular la idea de función con la de transformación, y ambas con el concepto de funcionamiento de un sistema, tenemos una descripción del mismo y una explicación (al menos en cierto modo) de su actividad, lo que es muy conveniente para rodear cierto tipo de problemas, pues el operador o "causa" de las transiciones queda definido a través de la descripción de la misma transformación que induce, que es cuestión de hecho o bien una hipótesis útil; lo que elimina especulaciones vacías sobre el sistema y mantiene las ventajas de los métodos experimental y axiomático.

La idea, aunque sencilla, tiene su trascendencia y merece que se clarifique con ejemplos. Podría suponerse que la causa del funcionamiento de un radio-receptor, pongamos por caso, es el suministro de energía eléctrica que recibe de una fuente, porque si esta se desconecta el receptor deja de funcionar; pero también dejaría de funcionar si todas las emisoras dejaran de transmitir, o si se eliminara el tubo conversor, o el emplificador de frecuencia intermedia, o si se sumergiera en agua al aparato, o si se interrumpiera su circuito en cualquiera de una infinidad de puntos, o si se invirtiera una polaridad, o si se trabara la bobina de sonido, etc., etc. Esta multiplicidad de la "causa" del funcionamiento, o sea del operador responsable del efecto útil, se hace extraordinariamente confuso en sistemas complicados, por lo cual,-

en lugar de hacer referencia directa a él, decimos que es todo aquello que hace que cada parte del sistema funcione exactamente en la forma en que lo hace, para dar un resultado concreto; esto es lo que dice, sintéticamente en símbolos, una transformación. Es como si se definiera al fuego, como "algo" que dilata los cuerpos, vaporiza a los líquidos, produce quemaduras en la piel, etc., etc. Análogamente, un operador es "algo" que produce en un sistema, la serie de cambios característicos que describe la transformación, y que, a través de ella, permite identificar al operador, que directamente es indefinible o difícil de definir.

Aunque todo esto puede parecer poco satisfactorio, es de hecho la manera en que, a veces algo disimuladamente, se procede en la "explicación" de la naturaleza. Considérese el muy importante caso de los fenómenos eléctricos, habitualmente atribuidos al exceso, o al defecto, o al flujo de electrones; pero la naturaleza concreta de estas cargas, o partículas, o energía vibrante, o lo que sean, es harto difícil de captar, pues se reduce a una colección de propiedades y leyes; y téngase en cuenta que este tipo de entidades son los elementos constitutivos de toda la materia que, se supone, es el fundamento de lo más concreto y tangible que existe.

12. Una manera de visualizar gráficamente, de un solo golpe de vista, las etapas sucesivas que experimenta un sistema, o lo que es igual, las potencias sucesivas de la transformación que la representa simbólicamente, consiste en escribir el símbolo representativo del estado inicial, luego una flecha que representa la transición, y a continuación el símbolo del estado correspondiente a la primera etapa, producida por la aplicación del operador al estado inicial; en seguida otra flecha, y luego el símbolo para el estado correspondiente a la segunda etapa, o sea, el resultado de aplicar el mismo operador al estado final de la primera etapa; y así se forman los demás eslabones de una cadena que puede estar formada por varias partes separadas, pudiendo ser cada parte un circuito cerrado, o una cadena abierta con, o sin derivaciones. Se le llama GRAFICA CINEMATICA.

Por ejemplo, sea la transformación T dada por:

$$T: \begin{array}{c} \downarrow \\ \\ \end{array} \begin{array}{cccccccccccccccc} A & B & C & D & E & F & G & H & I & J & K & L & M & N & O \\ J & E & F & H & E & C & I & G & D & O & L & K & L & J & N. \end{array}$$

Su gráfica cinemática es:

$$\begin{tikzcd} A \arrow[r] & J \arrow[dr] & & N \arrow[ll] \\ & & O \arrow[ur] & \end{tikzcd} \qquad \begin{tikzcd} D \arrow[r] & H \arrow[d] \\ I \arrow[u] & G \arrow[l] \end{tikzcd} \qquad \begin{tikzcd} M \arrow[r] & L \arrow[r, shift left] & K \arrow[l, shift left] \end{tikzcd} \qquad \begin{tikzcd} B \arrow[r] & E \end{tikzcd} \qquad \begin{tikzcd} F \arrow[d, shift right] \\ C \arrow[u, shift right] \end{tikzcd}$$

13. De las propiedades que poseen las transformaciones, nos interesan especialmente las cinco siguientes:

A) CLAUSURA.- Una transformación es <u>cerrada</u>, si en el conjunto Y de transformadas no hay ningún elemento que no exista ya en el conjunto X de operandos; o en términos de la teoría de conjuntos, si Y es un subconjunto de X, lo que simbólicamente se indica como

$$Y \subseteq X.$$

En relación con esta propiedad, se dice con frecuencia que "la transformación no crea elementos nuevos", o que "el conjunto de operandos es ESTABLE bajo la transformación". La transformación T del ejemplo anterior tiene esta propiedad, ya que el renglón inferior está formado por literales que se encuentran presentes en la fila superior.

B) UNIFORMIDAD.- Una transformación es <u>uniforme</u>, si cada operando se transforma en una transformada y sólo en una.

C) CORRESPONDENCIA BIUNIVOCA.- Se llama <u>biunívoca</u> la transformación, o se dice que se encuentra en <u>correspondencia de uno a uno</u>, si, además de ser uniforme, le corresponde también a cada transformada solamente un operando y sólo uno. Esto implica que los elementos, que forman el conjunto de transformadas, son todos diferentes entre sí.

D) MULTIFORMIDAD.- Cuando a cada operando le pueden corresponder alguna de dos o más transformadas posibles, la transformación se llama <u>multiforme</u>. Esta propiedad es fundamental en la descripción de sistemas no determinados.

E) CORRESPONDENCIA UNIVOCA.- <u>Unívoca</u> es la transformación en la que dos o más operandos distintos se transforman en la misma transfor—

mada; lo que implica que algunos elementos del conjunto de transformadas se repitan. Por ejemplo, la transformación T del No. 12 muestra esta propiedad.

14. En base a todo lo anterior se puede afirmar que la actividad de cualquier sistema puede ser descrita en forma conveniente mediante un registro tabular de su secuencia de estados, en el cual se encuentre implícita su ley de dependencia funcional. Este registro descriptivo del sistema real constituye su MODELO MATEMATICO.

II

LOS SISTEMAS DETERMINADOS Y EL TRANSDUCTOR.

1. Un MECANISMO ABSOLUTO es un sistema aislado cuya secuencia de estados está totalmente determinada; esto significa que es posible saber el estado en que se encuentre el sistema en cualquier instante t de tiempo, o en nuestro caso, el estado correspondiente a cualquier etapa pasada o futura a partir del conocimiento del modelo matemático y de su estado inicial.

El sistema se considera aislado, si su actividad no está condicionada a factores externos al mismo y depende sólo de la naturaleza y disposición de sus elementos internos. Algunos sistemas naturales y muchos de los artificios mecánicos comunes pueden considerarse sistemas de este tipo. Un buen ejemplo de mecanismo natural absoluto es el sistema solar, cuyo estado se puede predecir con mucha anticipación; y un ejemplo de uno artificial podría ser un reloj en marcha, o un tocadiscos automático, portátil, funcionando; etc.

Toda transformación, que describe a un mecanismo absoluto, es cerrada y uniforme.

2. El funcionamiento de un mecanismo absoluto no implica necesariamente ciclicidad, su transformación puede ser cerrada y uniforme, y sin embargo, infinita. Así la transformación

$$\begin{array}{ccccccccc} 1 & 2 & 3 & 4 & . & . & . & n_o & . \; . \; . \\ 2 & 4 & 6 & 8 & . & . & . & n & . \; . \; ., \end{array}$$

que tiene por forma algebraica

$$n = 2n_o,$$
$$n_o = (1, 2, 3, \ldots),$$

es uniforme cerrada e infinita.

No obstante, todo sistema real es en la práctica finito en cualquiera de sus características, aunque teóricamente sea de longitud indefinida la transformación que lo describe. Así por ejemplo, consideremos el caso de un cuerpo que se deja caer desde cierto punto estático a gran altura (altura que en la práctica siempre será limitada), y que desciende hacia la superficie terrestre. Si no se considera la amortiguación de la atmósfera, se supone la aceleración debida a la gravedad constante e igual a 9.8 m/s^2, y se dispone de los recursos adecuados para medir y registrar a intervalos fijos de tiempo la distancia descendida y la velocidad de descenso, obtendremos, sin recurrir a las leyes de la cinemática, el siguiente registro de datos correspondientes a intervalos de tiempo de un segundo; transformación que describe la actividad del sistema determinado en cuestión.

$$\downarrow \quad \begin{matrix} (0,\ 0) & (4.9,\ 9.8) & (19.6,\ 19.6) & (44.1,\ 29.4) & \ldots \\ (4.9,\ 9.8) & (19.6,\ 19.6) & (44.1,\ 29.4) & (78.4,\ 39.2) & \ldots \end{matrix}$$

Aquí cada elemento de la transformación es un vector con dos componentes; el primero representa la distancia vertical descendida, en metros, y el segundo componente representa la velocidad, en metros/seg., del objeto en ese punto y momento de su caída. Por supuesto que el modelo matemático preciso de este sistema está dado, para cada instante de tiempo y de acuerdo con la notación (9-C) Cap. I, por:

$$\left[(y,\ v) \ \middle|\ \begin{array}{l} y = y_0 + v_0(t - t_0) + \frac{1}{2}g(t - t_0)^2, \\ v = v_0 + g(t - t_0). \end{array} \right]$$

Supongamos ahora que sobre la base de la observación de los cambios trimestrales del aspecto del paisaje natural, formamos el siguiente registro:

$$\downarrow \quad \begin{matrix} P & V & O & I \\ V & O & I & P. \end{matrix}$$

Esta transformación describe vagamente el movimiento de la Tierra alrededor del Sol, quienes evidentemente forman un sistema determinado y cíclico.

3. Todo sistema absoluto puede considerarse como formado por el acoplamiento de subsistemas más simples, y también es posible sintetizar un mecanismo absoluto a partir de submecanismos que no solamente presenten una salida, sino también una entrada. Para desarrollar estas ideas es fundamental el concepto de transductor.

Entendemos por TRANSDUCTOR a un sistema dinámico, que además de presentar una secuencia espontánea y característica de estados, admite la posibilidad de modificar el tipo de secuencia que presenta su salida mediante influencia externa ejercida sobre él por la salida de otro sistema. Esto permite analizar aisladamente los componentes de un sistema, describiéndolos en términos de entradas (causas, estímulos, antecedentes, etc.) y salidas (efectos, respuestas, consecuentes, etc.); y recíprocamente, permite el acoplamiento de sistemas para sintetizar sistemas más complejos. El concepto es además muy importante, porque como veremos, el acoplamiento de sistemas implica la existencia de un canal de comunicación entre ellos que puede ser simple o mutuo; y también tiene considerable importancia la generalización misma de la noción de estímulo y respuesta porque, en general, no se trata de sucesos aislados en el tiempo, como a veces parece suponerse (conductismo), sino de modificaciones de actividades ya existentes y que presentan una continuidad.

En los casos en que puede emplearse la representación algebraica, la entrada del transductor, que cambia la línea de actividad del sistema, corresponde a la presencia de un PARAMETRO o constante relativa; término que por extensión podemos emplear en todos los casos.

Tratemos de ejemplificar con un caso tomado de la Geometría analítica. La expresión

$$x^2 - 2xy + Ry^2 + 2x - y + 1 = 0,$$

es la representación de una cónica que será parábola, elipse o hipérbo-

la, según que el valor de su parámetro R sea igual, mayor o menor que la unidad. Puede interpretarse lo anterior imaginando un osciloscopio que pudiera presentar en su pantalla este tipo de curvas, y diciendo que el comportamiento del aparato es de parábola, elipse o hipérbola, de acuerdo con el valor del parámetro fijado en sus controles.

Vemos entonces, que para describir a un transductor se requiere no una, sino tantas transformaciones como valores distintos pueda tomar la constante arbitraria que llamamos parámetro.

Si convenimos en considerar como un solo sistema el formado por tres hipotéticas emisoras de T. V., que transmiten por los canales 30, 40 y 50, y un receptor que podemos manipular; y además suponemos que únicamente se transmiten cuatro tipos distintos de programas, tales como noticieros, documentales, etc., que designaremos con las literales n, d, p, m, entonces, la descripción del funcionamiento diario de cada canal podría estar dado por las siguientes transformaciones:

$$C_{30}: \begin{array}{cccc} n & m & p & d \\ \downarrow & & & \\ m & p & d & n \end{array}$$

$$C_{40}: \begin{array}{cccc} n & m & p & d \\ \downarrow & & & \\ p & d & m & n \end{array}$$

$$C_{50}: \begin{array}{cccc} n & m & p & d \\ \downarrow & & & \\ d & n & m & p. \end{array}$$

Este sistema funciona como transductor con respecto al espectador, siendo la salida la secuencia del tipo de programas observada en la pantalla del cinescopio, y la entrada estaría representada por el dial que selecciona los diferentes canales mediante modificaciones en los valores de algunos parámetros eléctricos en el circuito del receptor, y que definen el tipo de secuencia de estados entregado en la salida; considerando, por supuesto, a cada tipo de programa como un estado característico.

Todo lo anterior podemos representarlo en forma más compacta de este modo:

↓	n	m	p	d
C_{30}	m	p	d	n
C_{40}	p	d	m	n
C_{50}	d	n	m	p

En donde C es el parámetro.

Otro ejemplo familiar de parámetro, lo constituye la posición de la palanca de velocidades respecto al régimen de funcionamiento de un vehículo motorizado; y en general, todos los dispositivos presentes en el tablero de control de maquinaria y equipo representan los parámetros del funcionamiento de los mismos.

4. Todo sistema que es o se considere absoluto está formado por el acoplamiento de transductores; y todo transductor es o puede ser parte componente de un sistema absoluto, con sólo tomar en cuenta a un mayor número de elementos.

Sean E y A dos sistemas con parámetros A' y B respectivamente, que deseamos acoplar para formar una sola unidad.

E: ↓	a	b	c
A'_1	b	c	c
A'_2	c	a	a
A'_3	c	a	b

A: ↓	d	e	f
B_1	f	d	e
B_2	d	f	e
B_3	e	f	f

El funcionamiento del sistema acoplado depende de la forma de acoplamiento, que es necesario especificar. Supongamos que la influencia de E sobre A a través del parámetro B, esté definido por la transformación T.

$$T: \begin{array}{cccc} E) & a & b & c \\ \downarrow & & & \\ A) & B_2 & B_1 & B_3 \end{array}$$

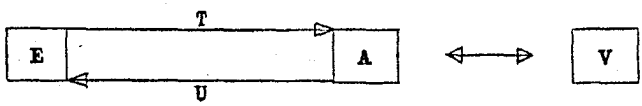
Y la acción de A sobre E por intermedio del parámetro A', esté dada por la transformación U.

$$U: \begin{array}{cccc} A) & d & e & f \\ \downarrow & & & \\ E) & A'_3 & A'_1 & A'_3 \end{array}$$

El sistema absoluto resultante V,

$$V: \begin{array}{cccccccccc} & (a, d) & (a, e) & (a, f) & (b, d) & (b, e) & (b, f) & (c, d) & (c, e) & (c, f) \\ \downarrow & & & & & & & & & \\ & (c, d) & (b, f) & (c, e) & (a, f) & (c, d) & (a, e) & (b, e) & (c, f) & (b, f) \end{array}$$

consta de nueve estados posibles con dos componentes cada uno.



5. Cuando el sistema es lineal y admite la aplicación del Cálculo infinitesimal, es usual dentro de la Ingeniería el empleo de la transformación de Laplace para pasar del dominio del tiempo al de la variable algebraica s, y entonces se facilita mucho la determinación de la función respuesta y particularmente el proceso de acoplamiento de sistemas mediante el concepto de FUNCION DE TRANSFERENCIA, llamada también TRANSMITANCIA, y que es la relación entre la transformada de Laplace de la función respuesta y la transformada de Laplace de la función excitación. En estas condiciones, el acoplamiento se reduce a multiplicar las funciones de transferencia de los sistemas componentes para obtener la función de transferencia del conjunto, y la determinación de la función

respuesta requiere únicamente la multiplicación de la función estímulo, que usualmente es dato, y la función de transferencia. Sólo con el fin de señalar el paralelismo entre este procedimiento que podríamos llamar ortodoxo, y el método que venimos presentando, vamos a hacer una breve exposición y a ilustrar con un ejemplo simple.

6. Dado un sistema físico que supondremos lineal, y cuyo modelo matemático deseamos obtener, la aplicación de leyes físicas conocidas nos permitirá establecer una ecuación diferencial lineal. Para fijar ideas supongamos que la ecuación es:

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = f(t). \qquad (5\text{-}1)\text{II}$$

En esta ecuación que ya es en sí el modelo matemático, f(t) es la entrada o función excitación, que generalmente se conoce, y "y" como función del tiempo es la salida o función respuesta, que usualmente se desconoce y es lo que interesa determinar a partir del modelo matemático inicial que es la ecuación diferencial.

Cuando f(t) es nula, se dice que la ecuación es HOMOGENEA y se tiene el caso de un sistema aislado o absoluto cuyo funcionamiento sólo depende de las características del sistema y no de la acción de un estímulo presente, y por supuesto, que la solución general de esta ecuación homogénea llamada FUNCION COMPLEMENTARIA de la ecuación completa o no homogénea (5-1)II, es, en el orden físico la RESPUESTA TRANSITORIA del sistema, o sea, la secuencia de valores que toma "y" después de la acción de un disturbio que lo puso en actividad pero que ya no está presente.

Como sabemos, la solución general clásica de una ecuación diferencial no homogénea es la suma de la función complementaria, que a su vez consta de una suma de términos con coeficientes constantes arbitrarios iguales en número al orden de la ecuación, y de una solución particular de la ecuación no homogénea, llamada INTEGRAL PARTICULAR, y que corresponde físicamente a la RESPUESTA DE ESTADO ESTABLE, que es la secuencia

de valores de "y" que se presentan como salida del sistema una vez que se ha establecido su régimen de funcionamiento.

6. Según lo anterior un transductor estará representado, para un buen número de casos prácticos de ingeniería, por una ecuación diferencial ordinaria lineal con coeficientes constantes, de orden n y que en general podemos escribir como

$$\frac{d^n y}{dt^n} + k_1 \frac{d^{n-1} y}{dt^{n-1}} + k_2 \frac{d^{n-2} y}{dt^{n-2}} + \cdots + k_{n-1} \frac{dy}{dt} + k_n y = f(t), \qquad (1\text{-}6)\text{II}$$

en donde f(t) es la entrada o función excitación. Entonces, si

$$y_c(t) = C_1 y_1(t) + C_2 y_2(t) + \cdots + C_n y_n(t) \qquad (2\text{-}6)\text{II}$$

es la solución general de la ecuación homogénea correspondiente, y

$$y_p(t) \qquad (3\text{-}6)\text{II}$$

es una solución particular de la ecuación completa (1-6)II, la función respuesta del sistema estará dada por

$$y(t) = y_c(t) + y_p(t). \qquad (4\text{-}6)\text{II}$$

Para ver la analogía de todo lo anterior con la representación de un transductor que dimos en el No. 3 de este Capítulo, enfoquemos al transductor desde otro punto de vista, como un dispositivo que modifica la naturaleza y las características de una señal que pasa a través de él; entendiendo por SEÑAL a una magnitud física que varía en el tiempo.

Sea el transductor A, definido en el No. 4 como

A: ↓	d	e	f
B_1	f	d	e
B_2	d	f	e
B_3	e	f	f

y al que le aplicamos a su entrada la señal

$$B_1 B_2 B_2 B_1 B_3 B_1 B_1 B_2, \qquad (5\text{-}6)II$$

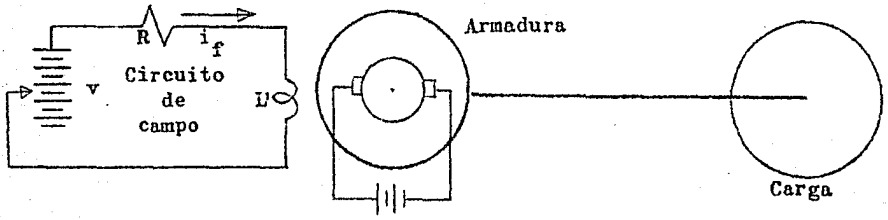
a partir del estado inicial e. El transductor transformará la señal de entrada en la siguiente señal de salida:

$$e\ d\ d\ d\ f\ f\ e\ d\ d. \qquad (6\text{-}6)II$$

De este modo, las señales de entrada y salida se presentan como una secuencia discontinua de etapas, pero se comprende que si se consideran a los estados del transductor como entidades físicas tales como posición angular de un eje, valores de diferencia de potencial, etc., y se hace tender a infinito el número de estados, y a cero el intervalo de tiempo entre etapas, tendremos en (5-6)II y (6-6)II la representación de funciones temporales continuas del tipo de f(t) y y(t), que son la función excitación y la función respuesta, respectivamente, para el caso de la ecuación diferencial (1-6)II.

7. Un ejemplo de acoplamiento en cascada, y de su solución a partir de expresiones diferenciales empleando técnicas modernas que evitan la laboriosidad de las soluciones clásicas, es presentado aquí para ilustrar desde el punto de vista que nos interesa, diferencias, analogías y ventajas.

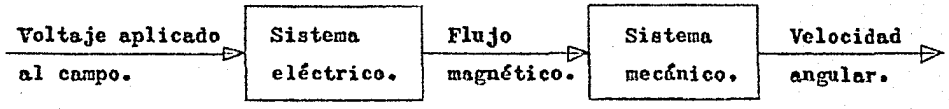
Se trata de un motor de C. D. con excitación independiente, que impulsa una carga, y en el que se puede controlar la velocidad de rotación mediante el control de la intensidad del campo.



En el diagrama, R, v, i_f y L representan respectivamente la resistencia, la diferencia de potencial aplicada, la corriente y la induc-

tancia correspondientes al circuito de campo del motor. D es el coeficiente de amortiguación mecánica, y J el momento de inercia del motor y de la carga. Todo en el sistema mks de unidades.

Se considera al circuito de campo como un sistema eléctrico que se va a acoplar a un sistema mecánico formado por la armadura, la flecha y la carga. Estos dos transductores, con sus respectivas entradas y salidas, pueden representarse mediante bloques, así:



En este caso no hay retroacción por tratarse de sistemas unilaterales, y el sistema eléctrico domina al sistema mecánico.

La aplicación de la ley de voltajes de Kirchhoff al transductor eléctrico, nos da:

$$v = \frac{R}{k_1}\phi + k_2 \frac{d\phi}{dt}, \qquad (1\text{-}7)II$$

suponiendo que $\phi = k_1 \, i_f$.

Aplicando la transformación de Laplace a (1-7)II, obtenemos:

$$L(v) = L(\phi) \left(\frac{R}{k_1} + sk_2 \right).$$

$$\text{Si} \quad a = \frac{k_1 \, k_2}{R}, \qquad (2\text{-}7)II$$

la función de transferencia de este transductor, es

$$\frac{L(\phi)}{L(v)} = \frac{k_1}{R} \; \frac{1}{1 + sa}. \qquad (3\text{-}7)II$$

Para el transductor mecánico se procede análogamente, partiendo de la siguiente ecuación cinética que es expresión del principio de D'Alembert:

$$T - Dw - J\frac{dw}{dt} = 0, \qquad (4\text{-}7)II$$

en donde T es el par mecánico ejercido magnéticamente sobre el rotor. Si se mantiene constante la corriente en la armadura, este par será proporcional a la intensidad del flujo magnético del campo, y podremos escribir

$$k\,\phi = Dw + J\frac{dw}{dt}. \qquad (5\text{-}7)II$$

Aplicando a (5-7)II la transformación de Laplace, obtenemos:

$$k\,L(\phi) = L(w)\,(D + Js). \qquad (6\text{-}7)II$$

Entonces, si hacemos

$$b = \frac{J}{D}, \qquad (7\text{-}7)II$$

la función de transferencia de este transductor mecánico, es

$$\frac{L(w)}{L(\phi)} = \frac{k}{D}\,\frac{1}{1 + sb}. \qquad (8\text{-}7)II$$

En este ejemplo el primer transductor transforma la señal que se le aplica a su entrada, en la forma de variaciones de diferencia de potencial, en variaciones de la intensidad del campo magnético; y el segundo transductor cambia esta última señal en variaciones de la velocidad angular. El acoplamiento de ambos transductores se logra multiplicando las funciones de transferencia (3-7)II y (8-7)II para obtener

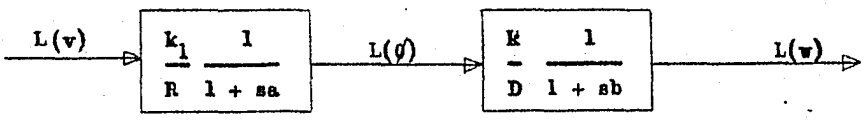
un único transductor, siendo el eslabón de enlace, en este caso, el flujo magnético. La función de transferencia del sistema integrado es entonces

$$\frac{L(w)}{L(v)} = \frac{k_1\ k}{R\ D}\ \frac{1}{(1 + sa)(1 + sb)}. \qquad (9\text{-}7)\,II$$

Esta expresión es ya el modelo matemático del sistema, del cual se puede obtener la transformada de Laplace de la función respuesta correspondiente a la transformada de Laplace de una función estímulo específica, para posteriormente aplicarle la transformación inversa de Laplace, y obtener directamente la respuesta en función del tiempo.

8. En la descripción de sistemas acoplados frecuentemente se emplean los diagramas de bloques, en el que cada bloque es un subsistema que queda totalmente definido por su función de transferencia que se escribe en él, y con flechas que indican su entrada y su salida. Este tipo de bloque llamado también a veces CAJA NEGRA ("black box"), es simplemente la representación de un transductor bajo la idea de un aparato sellado, cuya estructura interna es inaccesible a la observación directa, pero que posee ciertos medios para que se observen los resultados de su funcionamiento (salida), y controles para modificarlo (entrada).

El diagrama de bloques del ejemplo anterior sería el siguiente:



9. Los sistemas acoplados pueden presentar influencias mutuas de gran complejidad entre sus elementos. Para dar una leve idea de esto, vamos a ilustrar con un caso algebraico que permite desarrollos simbólicos más simples y compactos.

Sean P y Q los sistemas por acoplar, de tal modo, que Q afecte a P a través del acoplamiento R, quien da los valores de los parámetros de P en función de las variables de Q; y P afecte a Q a través del acoplamiento S, que indica los parámetros de Q correspondientes a las variables de P.

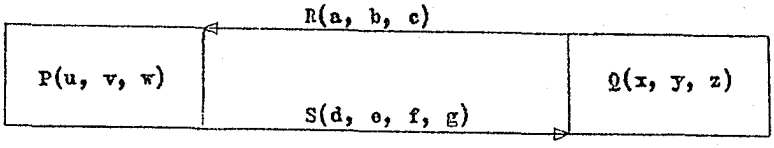
$$P: \begin{cases} u = u_0^3 - av_0 + w_0 \\ v = v_0 + bw_0 \\ w = cw_0 \end{cases} \qquad Q: \begin{cases} x = dx_0 - e \\ y = efy_0 + z_0 \\ z = 6fgz_0 \end{cases}$$

$$R: \begin{cases} a = 5x_0 - 2y_0 \\ b = -3z_0 \\ c = y_0 \end{cases} \qquad S: \begin{cases} d = v_0^3 \\ e = v_0 - w_0 \\ f = u_0^2 \\ g = 2w_0 \end{cases}$$

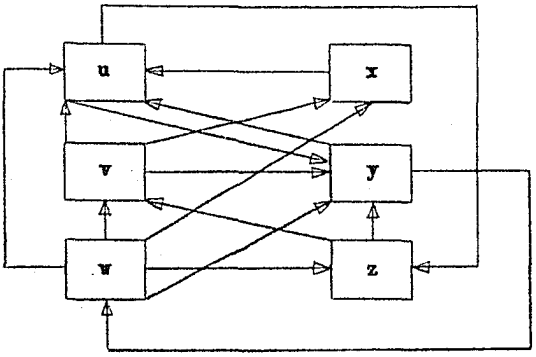
Llevando R a P y S a Q, obtenemos por sustitución algebraica, el sistema acoplado único T de seis variables

$$T: \begin{cases} u = u_0^3 - v_0(5x_0 - 2y_0) + w_0 \\ v = v_0 - 3z_0w_0 \\ w = y_0w_0 \\ x = v_0^3x_0 + w_0 - v_0 \\ y = u_0^2y_0(v_0 - w_0) + z_0 \\ z = 12u_0^2w_0z_0 \end{cases}$$

y el diagrama de bloques de la operación sería:



En T se observa que cada estado componente (o variable), depende no solamente de su propio estado inicial, sino también de los estados de los otros componentes correspondientes a la etapa anterior. Podemos trazar un DIAGRAMA DE EFECTOS INMEDIATOS, en que se muestren todas estas acciones y reacciones individuales.



10. A lo largo de la historia del pensamiento, desde Leucipo en el siglo V A. C. hasta los grandes instauradores del método científico en el siglo XVII, ha surgido bajo diferentes formas una concepción del Universo todo como un gran sistema cerrado y muy complejo, cuyas partes están trabadas entre sí por relaciones necesarias. Esta idea ha sido estimulada y apoyada principalmente por la observación de la perfecta regularidad cíclica de los movimientos astronómicos y de sus fenómenos consecuentes, como la sucesión de las estaciones, del día y de la noche, etc.; y desde luego, toda la mecánica clásica o de Newton está construida sobre esta base de un DETERMINISMO MECANICISTA, según la cual, cada estado del Universo físico es efecto o consecuencia necesaria de

su estado inmediato anterior y causa eficiente de su estado inmediato futuro, y así sucesivamente en una cadena ininterrumpida, según fue entendido y expresado insuperablemente por Laplace. Este esquema fue el fundamento de toda ciencia natural durante cerca de tres siglos, y continúa siendo el supuesto metódico obligado de toda actividad científica y tecnológica, pero trataremos de ver en la segunda parte cómo ha sido modificada radicalmente esta concepción mecanicista absoluta desde principios de este siglo, por la introducción del concepto de probabilidad y la labor de W. Gibbs y otros.

11. En relación con todo lo expuesto en esta primera parte, se puede concluir con la distinción de dos aspectos fundamentales:

A) Dado un sistema real, obtener por observación y experimentación directa su modelo conceptual.

B) Mediante el acoplamiento de sistemas de modelo matemático ya conocido, sintetizar el modelo de un sistema que cumpla con ciertos requisitos que nos convengan, y que se pueda materializar dentro de límites aceptables en un sistema real.

El primer aspecto es el de la investigación y el descubrimiento de leyes naturales. El segundo aspecto es el de invención y diseño de ingeniería; pero ambos están relacionados, ya que el diseño implica la aplicación de leyes y principios aportados previamente por el aspecto A.

El método general para establecer el modelo matemático de un sistema es en el fondo el método inductivo desarrollado por Galileo y por Bacon, y consiste esencialmente, partiendo del supuesto de una ignorancia completa acerca del sistema, en llevar un registro primario o PROTOCOLO, tan detallado y extenso como sea posible, de la actividad del sistema investigado consecuente a la aplicación al azar de una serie de estímulos. Una vez que se dispone de un registro protocolario de suficiente longitud, se procede a buscar lo sistemático, lo regular y reproducible, en suma, las relaciones constantes entre los elementos que varían, para finalmente ordenar estos datos en forma más compacta según la expresión funcional que más convenga, de acuerdo con el No. 7 del Capítulo I.

12. Vamos a terminar este Capítulo y esta Parte I referente a los sistemas y a su representación simbólica, con la solución de un problema satirizado planteado por Ashby, y al que me he permitido complicar ligeramente para destacar e ilustrar mejor el enfoque cibernético que se interesa menos por los sistemas en sí, a los que se les deja ser tan "fantasmales" como quieran, y se interesa más por los procesos mismos de regulación y de flujo de información que se manifiestan en ellos. El problema está contenido en la siguiente carta dirigida a un hipotético experto en la aplicación de los principios y métodos expuestos en esta parte.

Estimado amigo:

Hace algún tiempo que compré una vieja casa, pero me he encontrado con que está habitada por tres molestos fantasmas; un alquimista que satura el ambiente con las vaporizaciones de sus inútiles experimentos, una solterona escoptofílica, que con el murmullo de sus intrigas no deja dormir, y un aprendiz de detective, el rumor de cuyos pasos me sigue por todos lados. Consecuentemente es difícil de habitar. Sin embargo, he encontrado empíricamente que su actividad está sujeta a ciertas reglas extrañas pero infalibles, y que dependen de que encienda la luz, interprete música de órgano, o queme incienso.

En cada minuto, cada fantasma se encuentra en actividad o en reposo, y lo que hará cada uno en el minuto siguiente está estrictamente definido por lo que ha estado sucediendo en el minuto precedente.

He llevado un extenso registro, que le adjunto, tomado cuidadosamente a intervalos de un minuto, combinando los diversos estímulos de música, incienso y luz sin ningún plan preestablecido.

En el minuto presente, el murmullo y los pasos están sonando y no estoy aplicándoles excitación alguna. Indíqueme por favor qué combinaciones de luz, música e incienso debo hacer para tranquilizar a los fantasmas, y mantenerlos en reposo.

Atentamente.

Supongamos que el análisis y la ordenación del registro primario da las siguientes conclusiones:

I) Si la luz está apagada, el murmullo continúa en el estado en que estaba en el minuto anterior; pero si la luz está encendida, el murmullo hace lo opuesto de lo que estaba haciendo en la etapa precedente.

II) El alquimista por su parte, hace lo contrario de lo que estaba haciendo la solterona en la etapa anterior, a menos que se queme incienso, en cuyo caso, copia la actividad o el reposo de la solterona en el minuto precedente.

III) Finalmente, los pasos mantienen el estado de su etapa anterior si la luz está apagada, pero hacen lo contrario de lo que hacían si se enciende la luz y además se quema incienso o se toca música; pero si al mismo tiempo se quema incienso y se toca música, entonces copian el estado anterior del murmullo, y si solamente se deja encendida la luz, hacen lo mismo que el alquimista en el estado inmediato anterior.

Nótese ya la complicación de la descripción verbal en un sistema tan simple, comparada con la concisión de la representación simbólica en forma de matriz que damos en seguida.

La entrada del sistema es un parámetro con tres componentes (órgano, incienso, luz), y cada estado del sistema posee también tres componentes (murmullos, vapores, pasos). Si representamos con 1 a cada estado componente que esté activo, y con 0 si se encuentra en reposo, tendremos ordenaciones con repetición de dos elementos tomados de tres en tres, o sea,

$$^{2}OR_{3} = 2^{3} = 8 \text{ operandos diferentes}$$

e igual número de parámetros, que con los datos anteriores permiten formar la siguiente matriz:

(m,v,p)

	↓	(0,0,0)	(0,0,1)	(0,1,1)	(0,1,0)	(1,1,1)	(1,1,0)	(1,0,0)	(1,0,1)
	(0,0,0)	(0,1,0)	(0,1,1)	(0,1,1)	(0,1,0)	(1,0,1)	(1,0,0)	(1,0,0)	(1,0,1)
	(0,0,1)	(1,1,0)	(1,1,0)	(1,1,1)	(1,1,1)	(0,0,1)	(0,0,1)	(0,0,0)	(0,0,0)
	(0,1,1)	(1,0,1)	(1,0,0)	(1,0,0)	(1,0,1)	(0,1,0)	(0,1,1)	(0,1,1)	(0,1,0)
	(0,1,0)	(0,0,0)	(0,0,1)	(0,0,1)	(0,0,0)	(1,1,1)	(1,1,0)	(1,1,0)	(1,1,1)
(o,i,l)	(1,1,1)	(1,0,0)	(1,0,0)	(1,0,0)	(1,0,0)	(0,1,1)	(0,1,1)	(0,1,1)	(0,1,1)
	(1,1,0)	(0,0,0)	(0,0,1)	(0,0,1)	(0,0,0)	(1,1,1)	(1,1,0)	(1,1,0)	(1,1,1)
	(1,0,0)	(0,1,0)	(0,1,1)	(0,1,1)	(0,1,0)	(1,0,1)	(1,0,0)	(1,0,0)	(1,0,1)
	(1,0,1)	(1,1,1)	(1,1,0)	(1,1,0)	(1,1,1)	(0,0,0)	(0,0,1)	(0,0,1)	(0,0,0)

El problema tiene soluciones múltiples para cada situación inicial, según se desprende de la inspección de la matriz representativa. Si el parámetro inicial es (0,0,0), y el operando inicial (1,0,1), como especifica la carta, una solución sería pasar al parámetro (0,1,0) en la primera etapa, luego al (1,0,1) en la segunda, y al (1,1,0) en la tercera, permaneciendo ahí; otra solución consistiría en pasar en la primera etapa al parámetro (0,0,1), y en la segunda al (0,1,0) manteniéndolo, y permitiendo que el sistema siguiera en el primer caso, la trayectoria

$$(1,0,1) \longrightarrow (1,1,1) \longrightarrow (0,0,0) \rightleftarrows (0,0,0),$$

y en el segundo

$$(1,0,1) \longrightarrow (0,0,0) \rightleftarrows (0,0,0).$$

Terminando ambas trayectorias en un estado de equilibrio en cero.

S E G U N D A P A R T E

L A I N F O R M A C I O N

III

EL CONCEPTO DE PROBABILIDAD Y LOS SISTEMAS DE MARKOFF.

1. De acuerdo con la mecánica clásica, puede calcularse el estado futuro de un sistema mecánico cualquiera a partir del conocimiento preciso de su estado inicial, y la aplicación de las leyes de Newton. Inicialmente se pensó que el cálculo de probabilidades era un recurso provisional para manejar aquellas situaciones y áreas de conocimiento en las que aún no era aplicable el esquema determinista anterior. Dentro de la misma mecánica, los métodos estadísticos se desarrollaron para tratar con sistemas de un número tan grande de grados de libertad, que la aplicación de los métodos tradicionales resultaban prácticamente inútiles; no obstante, actualmente el punto de vista es completamente distinto, pues se considera que la determinación total es un caso particular y límite de un enfoque general probabilístico. En efecto, la condición clásica de un conocimiento preciso de los estados iniciales, obtenido por observación, no pasa de ser una hipótesis idealizada imposible de cumplir aun para los casos más artificialmente simples. Con todo, personalmente nos inclinamos a creer que la carencia de determinación no es una característica intrínseca de la realidad física, sino más bien una limitación impuesta a nuestro conocimiento que al parecer, y según el Principio de Indeterminación de Heisenberg, tiene el carácter de definitiva.

2. El método que aplica la mecánica estadística a los sistemas muy complicados es un tipo peculiar de análisis que da por sentado la parcialidad del conocimiento, y que en lugar de tratar de analizar la actividad de un sistema dado estudia la conducta de una colección de sistemas de estructura semejante que presenten un conjunto de estados diferentes, manejando valores representativos promedio de su actividad. Esta sustitución del estudio completo de un único sistema, por el comportamiento medio de una colección de sistemas similares, es lo que origina una notable simplificación y posibilita el establecimiento de

conclusiones sobre sistemas que requieren para su descripción de un número elevado de variables.

Es imposible calcular el momento y las circunstancias precisas en las cuales un elemento individual actúe de cierto modo, si el número de variables esenciales implicadas y mutuamente dependientes es excesivo, pero una gran cantidad de sistemas semejantes en circunstancias parecidas, presentarán con seguridad ciertas regularidades estadísticas que permitan manejar características promedio dentro de límites utilizables. Puede pensarse que estos principios son muy movedizos, y que su aplicabilidad es limitada, pero puede comprobarse su validez en los casos en que es posible verificar las conclusiones a que da lugar, por métodos más directos. Respecto a su aplicabilidad, es muy dudoso que exista algo totalmente caótico de lo cual no se pueda extraer ninguna sistematización. De hecho, se requieren precauciones especiales para garantizar una auténtica asistematicidad en la elaboración de las tablas de números de azar. Puede hacerse la prueba, por ejemplo, de tratar de escribir una serie de dígitos con la intensión expresa de que no tengan ninguna relación entre sí; y en series de suficiente longitud se podrá comprobar la efectiva imposibilidad de lograrlo, a causa entre otras cosas, de las conocidas leyes de asociación: de contraste, contiguidad, etc.

3. El desconocimiento de un número mayor o menor de factores de los que depende un fenómeno implica que las transiciones de un estado a otro no las podamos fijar con precisión, y tengamos la posibilidad de varias alternativas para cada operando, lo que nos lleva al establecimiento de una transformación multiforme que describe a un sistema no determinado, llamado SISTEMA DE MARKOFF. En este tipo más general de sistema un mismo operador, al operar sobre un mismo operando, puede originar una transformada de entre dos o más que son posibles, a causa de que nuestro grado de conocimiento sobre el sistema sólo nos permite establecer, entre antecedente y consecuente, relaciones de probabilidad y no de determinación.

Por todas las importantes consideraciones anteriores, y por ser el

eje y soporte de los procesos de información, se requiere que precisemos un concepto de probabilidad que esté acorde con la orientación que hemos llevado hasta aquí.

4. Cualquier situación en que se puedan obtener resultados diferentes como consecuencia de procesos aparentemente iguales, es un FENOMENO DE AZAR o EXPERIMENTO ALEATORIO, tal es el caso de cualquier tipo de medición en donde es difícil obtener dos lecturas iguales a pesar de proceder en cada medición en forma análoga; y es justamente por este carácter que tienen todas las mediciones, de ser fenómenos de azar, que es imposible obtener por observación datos absolutamente exactos acerca del estado inicial de un sistema, hecho que aunado a la virtual infinitud del número de variables de las que depende todo suceso real y a otros motivos teóricos de los que ya se habló, nunca es posible establecer un pronóstico científico de probabilidad verdaderamente unitaria. Y desde luego, de acuerdo con lo que se ha mencionado, cada etapa de funcionamiento de un sistema de Markoff es un fenómeno de azar.

5. Al conjunto completo de consecuencias posibles y distintas de un mismo experimento aleatorio, y que son mutuamente excluyentes en cada ensayo, se le llama ESPACIO MUESTRAL. Cada subconjunto del espacio muestral es un SUCESO, y si el subconjunto se reduce a un solo elemento, este se llama SUCESO ELEMENTAL.

Por ejemplo, si el experimento aleatorio consiste en tirar un dado homogéneo y simétrico, el espacio muestral U correspondiente, será

$$U = (1, 2, 3, 4, 5, 6),$$

en donde cada número representa la cara respectiva del dado y cuyas salidas constituyen los sucesos elementales.

Definido un cierto experimento aleatorio, y conocido su espacio muestral

$$U = (e_1, e_2, e_3, \ldots, e_n), \qquad (1\text{-}5)\text{III}$$

se sabe experimentalmente que si se repite el experimento un número N -

grande de veces, en varias series de pruebas distintas, las frecuencias relativas en cada serie de pruebas correspondientes a la ocurrencia de un suceso elemental cualquiera e_i del espacio muestral, se agrupan en la vecindad de una constante fija y desconocida $P(e_i)$ que se denomina PROBABILIDAD DEL SUCESO e_i.

O sea, si el suceso e_i ocurre n' veces en un número N de ensayos, entonces la frecuencia relativa de la ocurrencia de e_i, $f(e_i)$ es

$$f(e_i) = \frac{n'}{N}. \qquad (2\text{-}5)\text{III}$$

Si se repite el mismo proceso en m ocasiones, con valores diferentes para N, podremos obtener frecuencias absolutas n' también distintas, pero las relaciones correspondientes $\frac{n'}{N}$ tenderán a diferir más ligeramente entre sí a medida que los valores de N sean mayores. Teniendo suficiente cuidado en el empleo del concepto de límite, se puede establecer que si las frecuencias relativas $f(e_i)$ tienden a una constante $P(e_i)$ como límite, a medida que el número m de series de pruebas y el número N de ensayos en cada serie tienden a infinito, entonces la constante $P(e_i)$ es la probabilidad de la ocurrencia del suceso e_i. Esta es la idea fundamental que hay que aplicar siempre que no sea posible hacer suposiciones teóricas en problemas prácticos.

6. Se dice que los sucesos elementales de un espacio muestral son EQUIPROBABLES, si no existe ninguna razón conocida para esperar que uno de ellos ocurra con preferencia a cualquiera otro. Experimentalmente se puede verificar que, si se conoce el espacio muestral U de un experimento de azar, y los sucesos elementales e_i $(i = 1, 2, 3, \ldots, n)$ son equiprobables, entonces, en un número grande N de ensayos el número de veces que ocurre el suceso e_i es aproximadamente igual a $\frac{N}{n}$, y por lo tanto la frecuencia relativa de la ocurrencia de e_i será

$$f(e_i) \doteq \frac{N/n}{N} = \frac{1}{n} \doteq P(e_i); \qquad (1\text{-}6)\text{III}$$

y recíprocamente, si se cumple (1-6)III siendo n el número de sucesos elementales, entonces los sucesos e_i son equiprobables.

La idea de derivar directamente los conceptos de la probabilidad y sus teoremas asociados, de la teoría de conjuntos con los que naturalmente se vincula, se debe originalmente, al parecer, a C. B. Allendoerfer y C. O. Oakley. Este punto de vista unificante es muy conveniente, porque la usual aplicación directa que hacen los textos elementales de la definición de probabilidad, como una relación entre los casos favorables y los casos posibles, lleva con frecuencia en problemas algo elaborados a resultados absurdos o a paradojas; asímismo, el centrarse en las ideas de frecuencia relativa y conjuntos, deja sin ningún sentido el tratar de aplicar el cálculo de probabilidades a situaciones que por su naturaleza no lo admiten. Permítasenos enunciar, demostrar e ilustrar la aplicación de algunos de los teoremas más importantes, empleando aproximadamente la misma notación de los autores citados (a cuya obra referimos para una consulta más amplia), pero en una forma que creemos es algo más clara y explícita. Para esto, es necesario definir primero los conceptos de intersección y de unión entre conjuntos.

7. La INTERSECCION de dos conjuntos Q y R, es el conjunto formado por todos los elementos que son comunes a ambos conjuntos. Por ejemplo, si el conjunto Q es el formado por todos los domingos del año, y el conjunto R está formado por todos los días pares del mismo año, entonces la intersección de Q y R es el conjunto formado por todos los días pares del año que son domingos.

La intersección de Q y R se indica así:

$$Q \cap R \qquad (1\text{-}7)\text{III}$$

En cambio, la UNION de Q y R es el conjunto formado por todos los elementos que pertenecen al conjunto Q, o al conjunto R, o a ambos conjuntos a la vez. En el ejemplo anterior, la unión de Q y R que se indica como

$$Q \cup R, \qquad (2\text{-}7)\text{III}$$

sería el conjunto formado por todos los días del año que son pares, o que son domingos, o que son domingos y pares a la vez.

8. Como los sucesos elementales de un espacio muestral E son por definición incompatibles entre sí, todo suceso cualquiera S_j de E se puede considerar que está formado por la unión de los sucesos elementales que lo forman, tomando a cada uno de ellos como un conjunto de un solo elemento. Por tanto

$$S_j = (s_1, s_2, \ldots, s_i, \ldots, s_n) = s_1 \cup s_2 \cup \ldots \cup s_i \cup \ldots \cup s_n . \qquad (1\text{-}8)\text{III}$$

Esto significa que la ocurrencia del evento S_j es la ocurrencia de uno cualquiera de los sucesos elementales s_i.

9. El procedimiento más o menos uniforme que se sigue en la solución de problemas sobre probabilidades, consiste en definir primeramente el espacio muestral correspondiente, de tal modo que se abarquen todas las alternativas posibles y sean estas mutuamente excluyentes; esto quiere decir que en cada caso un suceso debe de ocurrir necesariamente de entre los varios que son posibles, y que la ocurrencia del mismo imposibilita la ocurrencia simultánea de cualquier otro. En segundo lugar, hay que asignar un valor numérico a la probabilidad de ocurrencia de cada suceso elemental, proceso denominado DISTRIBUCION DE PROBABILIDADES, y finalmente, la aplicación de algunos teoremas pertinentes dará la respuesta buscada.

La determinación del espacio muestral, y la distribución de las probabilidades, se realiza sobre datos de experiencia o bien basándose en hipótesis justificadas, tal como la muy frecuente y útil de equiprobabilidad; en estos casos se parte de un análisis de la naturaleza de la situación propuesta y de la constitución de los objetos involucrados. Cuando se presupone o comprueba la equiprobabilidad, la probabilidad es la misma para todos los sucesos elementales y es igual, según lo indica (1-6)III, al recíproco del mismo número de sucesos elementales que forman el espacio muestral; a este tipo de distribución de probabilidades

se le llama UNIFORME.

10. De todo lo dicho hasta aquí respecto a la probabilidad, se derivan las siguientes conclusiones como consecuencias directas de las relaciones existentes entre las frecuencias relativas, como representativas del valor numérico de la probabilidad.

A) Cada suceso elemental de un espacio muestral, tiene asociado un número real mayor o igual a cero y menor o igual a la unidad, que representa la probabilidad de su ocurrencia.

$$1 \geq P(s_i) \geq 0. \qquad \text{(A-10)III}$$

Esta doble condición es obvia, debido al carácter de porcentaje que tienen las frecuencias relativas.

B) La suma de las probabilidades asignadas a todos y cada uno de los sucesos elementales de un espacio muestral, debe ser igual a la unidad.

$$\sum_{i=1}^{n} P(s_i) = 1. \qquad \text{(B-10)III}$$

Lo anterior debido a que la suma de las frecuencias relativas de todos los sucesos elementales ocurridos en una serie de pruebas, es necesariamente la unidad, puesto que la suma de los sucesos elementales ocurridos, correspondientes a las diversas alternativas, tiene que ser igual al número total de ensayos.

C) La probabilidad de un suceso que sea un subconjunto cualquiera del espacio muestral, formado según (1-8)III por la unión de varios sucesos elementales, deberá ser igual a la suma de las probabilidades asignadas a dichos sucesos elementales, o sea, si

$$S_j = s_1 \cup s_2 \cup \ldots \cup s_i \cup \ldots \cup s_n,$$

entonces

$$P(S_j) = \sum_{i=1}^{n} P(s_i). \qquad \text{(C-10)III}$$

Esto porque como los sucesos elementales se excluyen mutuamente, la probabilidad de la unión de varios de ellos, es decir, la probabilidad de que ocurra cualquiera de ellos, es la suma de las probabilidades asociadas individualmente a cada suceso elemental, pues equivale a no hacer distinción alguna entre las diversas alternativas que abarca el suceso S_j, en cuyo caso, las frecuencias absolutas con que aparecen esos sucesos elementales, se suman para dar una frecuencia absoluta total; lo que se puede expresar así:

$$f(S_j) = \frac{N_1+N_2+\dots+N_i+\dots+N_n}{N} = \frac{N_1}{N} + \frac{N_2}{N} + \dots + \frac{N_i}{N} + \dots + \frac{N_n}{N} =$$

$$= f(s_1) + f(s_2) + \dots + f(s_i) + \dots + f(s_n),$$

y que transladado a probabilidades, nos lleva a la conclusión enunciada.

D) La probabilidad del espacio muestral tomado como un suceso formado por la unión de todos los sucesos elementales, es la unidad; es decir, certeza.

$$P(E) = 1. \qquad \text{(D-10)III}$$

Esta conclusión se desprende inmediatamente de B) y de C).

E) Todo suceso S_j que sea subconjunto del espacio muestral E, cumple por esta misma razón con la doble inecuación

$$1 \geq P(S_j) \geq 0; \qquad \text{(E-10)III}$$

puesto que se trata por A) y B) de una suma de términos que no puede sobrepasar a la unidad.

F) Si Q y R son sucesos que son subconjuntos de un espacio muestral E, y tales que su intersección sea un suceso dado por

$$Q \cap R = (s_1, s_2, \ldots, s_n),$$

entonces, la probabilidad de la ocurrencia de este suceso es la suma de las probabilidades de los sucesos elementales que forman la intersección.

$$P(Q \cap R) = \sum_{i=1}^{n} P(s_i). \qquad \text{(P-10)III}$$

Esto es consecuencia de C), y es importante por aplicarse en la solución de problemas en los que se tienen datos adicionales que restringen el espacio muestral, como se verá más adelante.

Ejemplo._ ¿Cuál es la probabilidad de que en un tiro con un dado homogéneo y simétrico, se obtenga un número par mayor de tres?

En problemas de este tipo, si se tienen condiciones múltiples, son frecuentes los errores cuando se parte de la definición habitual de probabilidad, y se trata de proceder intuitivamente; veamos como se resuelve este caso muy simple, aplicando el teorema anterior.

El espacio muestral es

$$E = (1, 2, 3, 4, 5, 6),$$

con distribución uniforme de probabilidades igual a 1/6.

Hagamos que Q sea el subconjunto formado por las caras de número par, y R el de las caras de número mayor de tres.

$$Q = (2, 4, 6);$$

$$R = (4, 5, 6).$$

Entonces, el conjunto de sucesos elementales que cumplen ambas condiciones, es

$$Q \cap R = (4, 6);$$

y por F) sabemos que su probabilidad es

$$P(Q \cap R) = P(4) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

F') Sucede a veces que es necesario tomar en cuenta dos o más experimentos aleatorios, que implican varios espacios muestrales distintos. Al considerar la ocurrencia simultánea de todos ellos, se obtiene un único espacio muestral; este ESPACIO MUESTRAL CONJUNTO, por el Principio de Multiplicación, base del Análisis Combinatorio, está formado por el producto cartesiano de los espacios muestrales componentes.

Así, para el caso de dos experimentos de azar con espacios muestrales

$$E_1 = (e_1, e_2, \ldots, e_n), \text{ y}$$

$$E_2 = (s_1, s_2, \ldots, s_m);$$

y distribución de probabilidades

$$P(e_1), P(e_2), \ldots, P(e_n), \text{ y}$$

$$P(s_1), P(s_2), \ldots, P(s_m);$$

al asociar cada evento simple de E_1 con cada uno de los de E_2, se obtienen los nuevos sucesos elementales que forman el espacio muestral conjunto

$$E = E_1 \times E_2 = (e_1 s_1, e_1 s_2, \ldots, e_1 s_m,$$

$$e_2 s_1, e_2 s_2, \ldots, e_2 s_m,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$e_n s_1, e_n s_2, \ldots, e_n s_m).$$

Y la probabilidad de los sucesos elementales combinados $e_i s_j$, con

$i = 1, 2, \ldots, n,$

$j = 1, 2, \ldots, m;$

serán, también por el Principio de Multiplicación

$$P(e_i s_j) = P(e_i).P(s_j). \qquad (F'-10)III$$

Ejemplo.- Sea el evento combinado de arrojar un cubo y un tetraedro, con sus caras numeradas.

El espacio muestral para el cubo, es

$$E_c = (e_1, e_2, e_3, e_4, e_5, e_6),$$

y para el tetraedro

$$E_t = (s_1, s_2, s_3, s_4).$$

Si suponemos distribuciones uniformes de probabilidades, entonces

$$P(e_i) = \frac{1}{6};$$

$$P(s_j) = \frac{1}{4}.$$

Como cada suceso referente al cubo puede asociarse con cada uno de los sucesos en el tetraedro, se obtiene el siguiente espacio muestral - conjunto, con veinticuatro elementos que abarcan todas las posibilida— des:

$$E = E_c \times E_t = (e_1s_1,\ e_1s_2,\ e_1s_3,\ e_1s_4,$$
$$e_2s_1,\ e_2s_2,\ e_2s_3,\ e_2s_4,$$
$$\cdots\cdots\cdots\cdots$$
$$e_6s_1,\ e_6s_2,\ e_6s_3,\ e_6s_4).$$

Las conclusiones y observaciones referentes a frecuencias relati— vas, son igualmente válidas considerando una sola prueba hecha con va— rios objetos similares, o realizando una serie de ensayos en sucesión - con un mismo espécimen en situaciones análogas; y de cualquier forma - que sea se adopta en la práctica como valor de la probabilidad, el pro- medio de las frecuencias relativas resultantes en muchos experimentos - múltiples, o en muchas series de experiencias singulares. Esta inter— cambiabilidad permite aclarar muchos aspectos, así, la aplicación del - Principio de Multiplicación, para determinar la distribución de probabi- lidades en el espacio muestral combinado, se entiende mejor si se con— sidera en el ejemplo que nos ocupa, no una serie de pruebas distintas - sino a varios especímenes (cubos y tetraedros) arrojados simultáneamen- te. Supongamos que se arrojan al mismo tiempo cien cubos y cien tetra— edros iguales, y que resultan diecisiete cubos en la cara i y veinti— cinco tetraedros en la cara j; se puede entonces asociar cada una de — las caras i con cada una de las caras j para tener una frecuencia abso- luta de caras combinadas ij igual a 425, y una frecuencia relativa de - 0.0425.

En general, suponiendo a los poliedros hechos de material homogé— neo, y simétricos, si se lanzan N cubos y M tetraedros, el número de cu- bos que saldrán en cualquier cara i serán aproximadamente $\frac{1}{6}N$, y el nú—

mero de tetraedros que resulten en cualquier cara j será aproximadamente $\frac{M}{4}$; y entonces la frecuencia relativa del suceso combinado $e_i s_j$ será:

$$f(e_i s_j) = \frac{n_i m_j}{N\,M} = \frac{N/6}{N} \cdot \frac{M/4}{M} = \frac{1}{6} \cdot \frac{1}{4},$$

y la distribución de probabilidades será uniforme e igual a:

$$P(e_i s_j) = P(e_i) \cdot P(s_j) = \frac{1}{24}.$$

Veamos ahora con otro ejemplo como se aplica el Teorema F), cuando se tiene la situación descrita en F').

En una urna hay tres fichas, una roja, una negra y una blanca. Si se arroja un dado y se extrae una ficha, determinar la probabilidad de que salga un número par en el dado y que la ficha sea roja.

El espacio muestral para el dado es

$$E_d = (1, 2, 3, 4, 5, 6),$$

con distribución uniforme de probabilidades de valor 1/6.

Para la urna, el espacio muestral es

$$E_u = (R, N, B),$$

con distribución uniforme de probabilidades igual a 1/3.

Por lo tanto, el espacio muestral conjunto, será

$$E = E_d \times E_u = (R1, R2, \ldots, R6, N1, N2, \ldots, N6, B1, B2, \ldots, B6);$$

y por (F'-10)III, su distribución uniforme de probabilidades es 1/18.

Sea A el suceso cara número par, y B el suceso ficha roja, entonces,

$$A = (R2, R4, R6, N2, N4, N6, B2, B4, B6),$$

$$B = (R1, R2, R3, R4, R5, R6);$$

y su intersección es

$$A \cap B = (R2, R4, R6);$$

y por el Teorema F),

$$P(A \cap B) = P(R2) + P(R4) + P(R6) = \frac{3}{18} = \frac{1}{6}.$$

G) La probabilidad de la ocurrencia del suceso Q, o del suceso R, o de ambos, es la suma de la probabilidad de Q y la probabilidad de R, menos la probabilidad de la ocurrencia simultánea de Q y de R.

En símbolos de la teoría de conjuntos, lo anterior se expresa así:

$$P(Q \cup R) = P(Q) + P(R) - P(Q \cap R). \qquad \text{(G-10)III}$$

Si los sucesos fueran incompatibles entre sí, la probabilidad de que ocurriera uno cualquiera de ellos sería, por la misma razón expuesta para C), únicamente la suma de las probabilidades correspondientes a cada suceso; pero en el caso general los eventos no se excluirán, pudiendo ocurrir que existan elementos comunes en ambos subconjuntos, y para no tomarlos dos veces en cuenta, hay que restar la probabilidad de la intersección de Q y R.

Ejemplo.- Se desea saber cual es la probabilidad de sacar de una urna, que contiene diez fichas marcadas del uno al diez, una ficha que sea múltiplo de dos, o una múltiplo de tres, o una que sea múltiplo de dos y de tres; pudiendo sacar sólo una ficha.

El espacio muestral es

$$E = (1, 2, \ldots, 10),$$

y se justifica la hipótesis de equiprobabilidad, con valor de 1/10.

Llamemos Q al suceso "múltiplo de dos", y R al suceso "múltiplo de tres"; entonces,

$$Q = (2, 4, 6, 8, 10); \quad P(Q) = \frac{5}{10}.$$

$$R = (3, 6, 9); \quad P(R) = \frac{3}{10},$$

por tanto

$$Q \cap R = (6); \quad P(Q \cap R) = \frac{1}{10},$$

y de acuerdo con G)

$$P(Q \cup R) = \frac{5}{10} + \frac{3}{10} - \frac{1}{10} = \frac{7}{10}.$$

H) Existen situaciones en las cuales, el número de posibilidades para un suceso, queda modificado por el conocimiento de cierta condición restrictiva. Estas PROBABILIDADES CONDICIONALES tienen importantes aplicaciones, tales como la descripción de la generación de símbolos en ciertos tipos de fuente de información.

Supongamos, teniendo en cuenta la equivalencia mencionada en F'), que se arrojan N dados iguales, simultáneamente, y que resultan N_R caras de número par y $N_{Q \cap R}$ caras de número par y mayores de dos; entonces, si se eliminan todos los dados que resultaron con número impar, las frecuencias relativas de las caras pares mayores de dos, respecto al número total de dados que ahora es de sólo N_R, es

$$\frac{N_{Q \cap R}}{N_R}.$$

Dividiendo el numerador y el denominador de la expresión anterior, por el número total original de dados N, tendremos

$$\frac{N_{Q \cap R}/N}{N_R/N} = \frac{f(Q \cap R)}{f(R)};$$

BIBLIOTECA CENTRAL
U. N. A. M.

resultado que trasladado a probabilidades, se expresa como

$$P(Q \mid R) = \frac{P(Q \cap R)}{P(R)}; \quad P(R) \neq 0, \qquad \text{(H-10)III}$$

y se lee: "la probabilidad condicional de la ocurrencia de Q dada la ocurrencia de R, es igual a la probabilidad de la ocurrencia simultánea de Q y de R, dividida entre la probabilidad de la ocurrencia de R".

En otras palabras, esto significa que el dato extra limita las posibilidades, ya que en el caso incondicionado la referencia para calcular las frecuencias relativas es el número total N de ensayos, en tanto que en el caso condicionado, la condición limita la referencia a N_R, que es un subconjunto de N.

Ejemplo.- En un librero hay tres libros de Historia, dos de Geografía y cuatro de Matemáticas. Si se saca uno al azar, ¿cuál es la probabilidad de que sea de Geografía, si por algún medio sabemos que no puede ser de Matemáticas?

$$E = (H_1, H_2, H_3, G_1, G_2, M_1, M_2, M_3, M_4);$$

$$P(H) = \frac{1}{3},$$

$$P(G) = \frac{2}{9},$$

$$P(M) = \frac{4}{9}.$$

$$Q = (G_1, G_2); \qquad P(Q) = \frac{2}{9}.$$

$$R = (H_1, H_2, H_3, G_1, G_2); \qquad P(R) = \frac{3}{9} + \frac{2}{9} = \frac{5}{9}.$$

$Q \cap R = (G_1, G_2)$; $P(Q \cap R) = \frac{2}{9}$. Por tanto aplicando el Teorema (H-10)III,

$$P(Q \mid R) = \frac{2/9}{5/9} = \frac{2}{5}.$$

11. A pesar de que el primer problema sobre probabilidades fue resuelto por B. Pascal, hace más de tres siglos, aún no constituye un cuerpo de teorías totalmente desarrollado y unificado, y todavía es susceptible de varias interpretaciones distintas; pero es indudable que muchas de sus aplicaciones importantes, tales como la delimitación de la causa de un fenómeno observado, cuando existen varias posibles causas productoras del mismo; la verificación de hipótesis; etc., se comprenden mejor partiendo de una interpretación estadística y desarrollando consecuencias con el auxilio de la teoría de conjuntos. Sin embargo, nuestro interés principal está aquí orientado a presentar y justificar la relación que tienen las probabilidades con la descripción de sistemas parcialmente conocidos, y con la cuantificación de los procesos informativos; y creemos que los principios expuestos son suficientes para lograr este objetivo.

12. Sólo puede ser objeto de la ciencia, aquellas relaciones que tienen carácter de invariantes, que mantienen su constancia y validez en todo tiempo y lugar, y a través del flujo de situaciones cambiantes; así son todas las leyes naturales. Según esto, los sistemas que se han llamado de Markoff no podrían ser manejados por la ciencia, puesto que no presentan uniformidad en sus transiciones de un estado a otro, pero es aquí en donde la probabilidad aporta el elemento de constancia. No se conoce cual será el estado siguiente de un sistema de Markoff, pero se puede saber con suficiente precisión la probabilidad de las diferentes alternativas posibles, y a su vez, las proposiciones que así lo expresen, no serán de verdad probable, sino efectiva. Se trata pues de un recurso para manejar racionalmente a la ignorancia, recortándola, pesándola, cuantificándola. Pero la probabilidad tiene además otro aspecto, no solamente permite el empleo óptimo de un conocimiento parcial, sino que también justifica la existencia de la comunicación, pues podemos afirmar, ya desde ahora, que la existencia de cualquier mensaje no tendría ninguna razón de ser, si su contenido no fuera probable sino cierto. En la medida en que la ocurrencia de un suceso se aproxima a la certeza, deja de ser objeto de información, así por ejemplo, la predicción de la salida aparente del Sol por el Este, no puede constituir una

noticia, en su sentido usual; y recíprocamente, para un ser omnisciente, la recepción de información no tendría ningún sentido.

13. En un sistema de Markoff, cada estado puede transformarse, con una probabilidad definida, en cualquier otro de un conjunto de estados posibles; esto puede representarse también en forma tabular mediante una MATRIZ DE TRANSICIONES PROBABLES, cuyo aspecto es el siguiente:

$$
\begin{array}{c|ccccccc}
 & a_1 & a_2 & a_3 & \cdots & a_r & \cdots & a_n \\
\hline
a_1 & p_{11} & p_{21} & p_{31} & \cdots & p_{r1} & \cdots & p_{n1} \\
a_2 & p_{12} & p_{22} & p_{32} & \cdots & p_{r2} & \cdots & p_{n2} \\
\cdot & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_q & p_{1q} & p_{2q} & p_{3q} & \cdots & p_{rq} & \cdots & p_{nq} \\
\cdot & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_m & p_{1m} & p_{2m} & p_{3m} & \cdots & p_{rm} & \cdots & p_{nm}.
\end{array}
\qquad (1\text{-}13)\text{III}
$$

En donde el símbolo general p_{rq}, indica la probabilidad de que el operando a_r se transforme en la transformada a_q. De este modo, cada columna está formada por todas las transiciones posibles que tienen por estado inicial al elemento que encabeza la columna, y puesto que de todos los estados que son posibles, un sistema puede únicamente encontrarse en un momento dado, en uno solo de ellos, entonces, de acuerdo con 5.III, cada columna de la matriz de transiciones probables constituye un espacio muestral, y cada elemento de la misma es un suceso elemental; por lo que, según el Teorema (B-10)III, su suma debe ser la unidad.

14. Así, si en cierta región de la Tierra las características climáticas y meteorológicas, obtenidas mediante observaciones registradas durante varios años, son tales que las probabilidades del estado del tiempo en el mes X están expresadas por la matriz

	S	N	L
S	1/4	1/3	1/8
N	5/8	1/2	1/4
L	1/8	1/6	5/8,

significando S, N y L, soleado, nublado y lluvioso respectivamente, entonces, esta matriz de transiciones probables nos dice, en forma condensada, que si hoy es un día nublado, la probabilidad de que mañana también lo sea, es de 50%; y que si hoy el cielo está despejado, la probabilidad de que mañana llueva es de 0.125; etc.

Los sistemas de Markoff no solamente pueden ser absolutos, también pueden ser transductores, en cuyo caso su descripción requiere una matriz para cada parámetro; y en general les son aplicables las mismas operaciones que se pueden realizar con los sistemas determinados, los cuales como ya dijimos, son en realidad casos particulares de los sistemas no determinados.

IV

NATURALEZA Y MEDIDA DE LA INFORMACION.

1. Al abordar el tema de la información, dos son los obstáculos iniciales con que se tropieza: los múltiples sentidos en que puede entenderse el término, y el desconcierto que produce la forma en que se le mide. El término con que con más frecuencia se le confunde es el de significado, pero aunque el tema tiene relaciones con la ciencia de la semántica, un mensaje puede contener mucha información y no tener ningún significado. Esto revela, además de que el empleo de esta palabra es en un sentido muy particular y técnico, el tipo de problemas que le dió origen a esta teoría: la necesidad de medir la capacidad de los sistemas telefónicos de mucha complejidad, y de los que hay que obtener el máximo rendimiento; tipo de problema que no se presentó mientras los sistemas fueron relativamente simples, y con bajo nivel de tráfico informativo. Se comprende que el que diseña un sistema telefónico no le interesa, en principio, lo que se diga ni lo que esto signifique, sino que su interés se centra en que lo que se transmita, sea esto lo que fuere, sea suficientemente claro y distinto, no en cuanto a sentido, sino en cuanto a fidelidad reproductiva. Por consiguiente, su preocupación primaria no es ni la gramática, ni la lógica, ni la fonética, en cuanto tales, sino únicamente el soporte sensible de que se sirven estas ciencias para comunicar el pensamiento, es decir, los símbolos y las relaciones que deben existir entre ellos, si han de ser capaces de transportar un mensaje; y lo mismo vale decir para las órdenes mecánicas dadas a una máquina, o para las instrucciones almacenadas en la cinta magnética o en la tarjeta perforada de una computadora.

2. Antes del preciso establecimiento formal de esta teoría, publicada por el matemático Claude E. Shannon en 1948, se intentaba solucionar cierto tipo de cuestiones basándose en suposiciones de "sentido común", que actualmente se han descartado o modificado radicalmente. El resultado de esta investigación teorética ha sido una sólida estructu—

ra que ha ido mucho más allá de sus objetivos iniciales, pues ilumina no sólo a la ingeniería de comunicaciones y a la lingüística, sino que también proporciona nuevas y mejores bases para atacar, quizás por primera vez con éxito, toda una amplia gama de problemas. En realidad, tiene algo importante que decir en relación con el estudio de todo aquello que implique un dinamismo organizado. Pero será mejor partir de una situación concreta, por medio de la cual se pueda establecer directamente lo que se entiende por información y por su medida.

3. Designemos con a, el estado de un individuo X antes de la recepción de un mensaje que espera, y que sabe puede ser únicamente m_1 o m_2. Sea p la probabilidad a priori de que el mensaje que reciba sea m_1, y $q = 1 - p$ la probabilidad de que sea el mensaje m_2. Supongamos que X se encuentre particularmente interesado en que el mensaje elegido sea m_1, en cuyo caso, q representará la magnitud de su incertidumbre. Si posteriormente, m_1 es el mensaje que efectivamente recibe, X habrá pasado del estado a a un nuevo estado b. ¿Qué es lo que caracteriza a esta transición?; únicamente la anulación de q; el paso de p a la unidad; el tránsito de la posibilidad al hecho, a la certeza. Pero nótese bien que la certeza a la que nos referimos aquí, es estrictamente la del suceso de selección del mensaje, y no a la verdad del posible contenido del mismo, que ya involucra otros factores.

Ahora bien, si la probabilidad de m_1 es igual a la de m_2, entonces, por el Principio (1-6)III

$$p = q = \frac{1}{2},$$

y se dice que la recepción de m_1 le ha proporcionado a X una unidad de información.

4. Hemos visto pues, que la probabilidad nos proporciona un punto de apoyo para manejar a la información, a la cual podemos ya definir como la posibilidad de elegir una alternativa de un espacio muestral, siendo su medida, la magnitud de la incertidumbre eliminada por la elección de una alternativa específica.

Pasemos ahora a una exposición más sistemática y formal de esta materia.

5. Tanto los procesos de control como los de comunicación, enfrentan problemas afines, pues ya se trate de un servomecanismo o de un sistema de telecomunicaciones, el problema fundamental consiste en reproducir con fidelidad óptima una señal de entrada.

Conviene aquí hacer notar, que a pesar de ser frecuente el uso indistinto de las palabras "información" y "comunicación", estas no son términos sinónimos, aunque sí muy relacionados, ya que la COMUNICACION es la actividad o situación que tiene como resultado la información; o en otras palabras, la INFORMACION es el producto de la comunicación. Análogamente, hay que hacer distinción entre "señal" (de la que ya se habló en 6.II), y "mensaje". El MENSAJE, que Wiener define como "una secuencia discreta o contínua de eventos mensurables distribuidos en el tiempo", es la señal antes de pasar por el elemento transmisor, o expresado en forma más precisa, la SEÑAL es la forma codificada del mensaje. De este modo queda claro también, que toda señal portadora de información es una función temporal; aun los símbolos gráficos participan de esta naturaleza temporal, puesto que tanto su escritura (codificación) como su lectura (decodificación) se realiza necesariamente en el tiempo, pero esta dependencia temporal, como ya hemos sugerido, no es determinada sino que tiene el carácter de fenómeno aleatorio, pues como se dijo, un mensaje sólo cumple su finalidad informativa cuando el elemento receptor no está cierto de su contenido por adelantado. Otra característica de los mensajes, es que usualmente cuando se reciben, llevan sobrepuesto algún tipo de disturbio, llamado genéricamente RUIDO, que altera la información y disminuye la fidelidad. Estos elementos parásitos, tienen también el carácter de funciones de azar.

6. Toda comunicación implica los siguientes elementos:

a) Una fuente de información, que genera los mensajes por selección de elementos de un espacio muestral.

b) El mensaje, que constituye la salida de la fuente de información.

c) El transmisor o codificador, que transforma en señales a los mensajes.

d) La señal, que constituye la salida del transmisor.

e) El canal de comunicación, que es el medio por el que viaja la señal desde el emisor hasta el receptor.

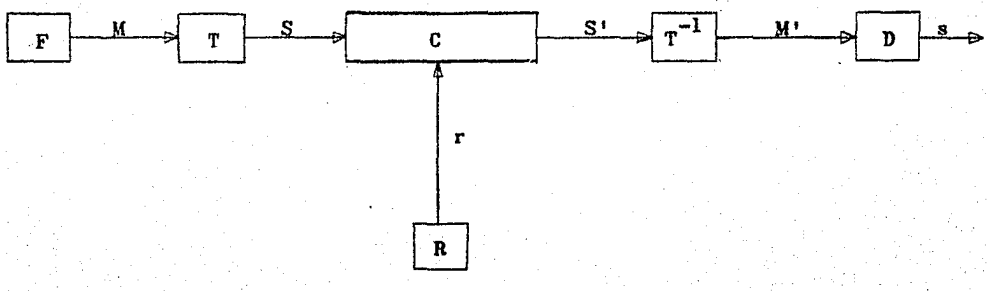
f) La fuente de ruido, elemento generalmente indeseable, que genera señales parásitas.

g) El ruido, que hace que la señal recibida difiera en mayor o menor grado de la señal original.

h) El receptor o decodificador, que reestructura el mensaje original en forma más o menos fiel.

i) El destinatario, que es el sistema que tiene por entrada el mensaje recibido, y cuya salida queda modificada por efecto del mismo.

Esquemáticamente los distintos elementos que forman un sistema de comunicación unilateral, y sus relaciones, están contenidos en el siguiente diagrama de bloques,



(1-6)IV

en donde:

F = Fuente de información.

M = Mensaje seleccionado.

T = Transmisor.

S = Señal emitida.

C = Canal de comunicación.

R = Fuente de ruido.

r = Ruido.

S' = Señal recibida.

M' = Mensaje decodificado.

D = Sistema destinatario.

s = Salida del sistema destinatario.

T^{-1} = Decodificador o receptor.

7. Si una situación es indeterminada, esto implica la existencia de más de una posibilidad, siendo el caso más simple, el de dos alternativas con iguales posibilidades cada una, como en el ejemplo inicial presentado en 3.IV. Con esta situación binaria y equiprobable, se asocia naturalmente la unidad de información, pero ¿cómo emplear esta unidad informativa para valorar situaciones más elaboradas, en que intervengan muchas alternativas con diferentes probabilidades? En principio, el número mismo de alternativas diferentes disponibles, es ya un medio para apreciar la capacidad informativa de la situación; Ashby llama simplemente a este número, VARIEDAD, así, si una situación admite cinco alternativas, se caracterizará por una variedad de cinco, y esto es una indicación de su capacidad para contener información. Pero si la situación, además de ser múltiple, no presenta iguales posibilidades para cada alternativa, entonces hay que agregar a la noción de variedad, la de probabilidad numérica. Finalmente, cabe preguntar: ¿Porqué a la medida de la información se le da forma logarítmica? R. V. Hartley fue quien primero comprendió las ventajas de esta forma de medición; dos de las principales son:

1) La expresión matemática del número de posibilidades, involucra comunmente potencias, y naturalmente, el empleo de logaritmos facilita su manejo.

2) El empleo de logaritmos permite mantener la idea espontánea de medir por comparación proporcional con la unidad adoptada; o sea, concuerda con la tendencia a creer que si una página de un libro, o una cinta magnética, o una fotografía, contienen cierta cantidad de datos o de información, N número de estos elementos, contendrán una cantidad de información N veces mayor.

Tratemos de ilustrar ambos aspectos con un ejemplo-problema. Supóngase, que un cierto tipo de semáforo tiene dos lámparas, cada una de las cuales puede estar en un momento dado, prendida o apagada. Si simbolizamos con 1 el estado de encendido y con 0 el de apagado, hay cuatro estados posibles, que son: 00, 01, 10, 11; utilizables como otras tantas señales de tránsito. Pero si se le agregan otras dos lámparas, ¿qué tanto aumenta su capacidad informativa?

El Análisis Combinatorio, nos dice que el número de estados posibles, o la variedad disponible, está dada por el número de ordenaciones con repetición de dos elementos (es decir, los estados 0 y 1) tomados de cuatro en cuatro. No es necesario saber cuáles son, aquí nos basta con saber que

$$^{2}OR_{4} = 2^{4} = 16.$$

Se dispone pues de dieciséis alternativas; la duplicación de lámparas, cuadruplicó las posibilidades.

Ahora bien, en el semáforo original se tenía variedad de cuatro; si tomamos logaritmos decimales, tendremos que con aproximación de regla de cálculo,

$$\log_{10} 4 = 0.602;$$

y para el semáforo ampliado, se tendrá

$$\log_{10} 16 = 1.204.$$

Pudiéndose decir que el primer semáforo tiene asociada una capacidad informativa de 0.602 unidades x, y el semáforo doble la tiene, en las mismas unidades, de 1.204; es decir, también doble.

Aunque la base de los logaritmos es arbitraria, como ya se dijo, la situación binaria por ser la más simple sugiere la adopción de una base binaria; si así se hace, a la unidad resultante se le denomina BIT; contracción de las palabras "BInary uniT".

En este caso, el primer semáforo tendrá una capacidad de

$$\log_{2} 4 = 2 \text{ bits},$$

y el semáforo doble, la tendrá de

$$\log_{2} 16 = 4 \text{ bits}.$$

Análogamente, para el ejemplo inicial dado en 3.IV, la recepción del mensaje m_1, le proporcionó a X

$$\log_2 2 = 1 \text{ bit de información.}$$

También se suele emplear la base 10 y la base e = 2.718..., llamándose entonces a las unidades resultantes, HARTLEY y NAT (contracción de "NAtural uniT") respectivamente.

Ahora ya estamos en condiciones de definir formalmente a la información y a su medida. Si adoptamos la base binaria, diremos que la <u>información</u> I(j), asociada con la ocurrencia del suceso S_j, es el número de veces que es necesario duplicar la probabilidad a priori del suceso S_j, para obtener certeza; y a su vez, si la probabilidad de ocurrencia del suceso S_j es P(j), su ocurrencia proporciona

$$I(j) = \log_2 \frac{1}{P(j)} \qquad (1\text{-}7)\text{IV}$$

UNIDADES BINARIAS de información, o BITS.

Puede verificarse fácilmente con ayuda de (1-6)III, que si la situación es equiprobable, como se supuso en el ejemplo de los semáforos, (1-7)IV se reduce al logaritmo de la variedad presente en la situación, o sea, al logaritmo del número de alternativas distintas.

8. Como las tablas de logaritmos disponibles, son generalmente de base decimal, puede ofrecer dificultades el cálculo de los logaritmos binarios, por lo que es necesario disponer de una expresión que relacione una base de logaritmos con otra. Dado que la deducción es fácil, los libros usualmente la omiten. Su utilidad y su uso frecuente, justifican que la demos aquí.

Se dispone de una tabla de logaritmos de base a, de modo que para cualquier número N se tiene como dato, que

$$\log_a N = x; \qquad 1)$$

y se desea conocer el logaritmo del mismo número en una base distinta, b.

Supóngase que la incógnita buscada sea "y", entonces

$$\log_b N = y. \qquad 2)$$

De 1) y 2) se obtiene

$$a^x = N, \qquad 3)$$

$$b^y = N. \qquad 4)$$

Por tanto

$$a^x = b^y. \qquad 5)$$

Tomando logaritmos de base a en ambos miembros de 5), tendremos — que

$$\log_a(a^x) = \log_a(b^y),$$

o sea

$$x = y \log_a b. \qquad 6)$$

Llevando 1) y 2) a 6),

$$\log_a N = \log_b N \log_a b;$$

y por tanto

$$\log_b N = \frac{\log_a N}{\log_a b}, \qquad (1\text{-}8)\text{IV}$$

es la relación de equivalencia deseada.

Si b = 2, y a = 10, (1-8)IV da:

$$\log_2 N = \frac{1}{\log_{10} 2} \log_{10} N,$$

en donde el factor de transformación vale aproximadamente 3.322, por lo que

$$\log_2 N = 3.322 \log_{10} N; \qquad (2\text{-}8)IV$$

y por consiguiente

$$1 \text{ Hartley} = 3.322 \text{ bits}. \qquad (3\text{-}8)IV$$

Si en el ejemplo de 3.IV, no hubiera existido equiprobabilidad, — sino que las probabilidades hubieran sido

$$P(m_1) = \frac{1}{3}; \qquad P(m_2) = \frac{2}{3},$$

entonces la recepción de m_1 le habría proporcionado a X, según (1-7)IV,

$$I(1) = \log_2 \frac{1}{1/3} = \log_2 3 \text{ bits},$$

lo que por (2-8)IV, es

$$3.322 \log_{10} 3 = 1.5849 \text{ bits}.$$

En igual forma, la aplicación de (1-8)IV para el caso de que b=2, y a=e= 2.7182..., nos dice que

$$\log_2 N = 1.442 \log_e N, \qquad (4\text{-}8)IV$$

y por lo tanto

$$1 \text{ nat} = 1.442 \text{ bits}. \qquad (5\text{-}8)IV$$

Análogamente para b=e y a=10,

$$\log_e N = 2.302 \log_{10} N, \qquad (6\text{-}8)IV$$

por lo que

$$1 \text{ Hartley} = 2.302 \text{ nats.} \qquad (7\text{-}8)\text{IV}$$

9. Dado que difícilmente se puede sobreestimar la importancia que tiene el poder cuantificar exactamente a la información, vamos a cerrar esta II Parte, solucionando el problema de la sección anterior, para el caso de la recepción del segundo mensaje, m_2, para luego comparar los resultados; y como la mera aplicación de una fórmula puede no satisfacer, justifiquémosla partiendo directamente de la definición dada en 7.IV.

Si el mensaje que recibe X es m_2, (1-7)IV nos dice que

$$I(2) = \log_2 \frac{3}{2}, \qquad (1\text{-}9)\text{IV}$$

o sea,

$$2^{I(2)} = \frac{3}{2}. \qquad (2\text{-}9)\text{IV}$$

Multiplicando ambos miembros de (2-9)IV por $\frac{2}{3}$, se tiene que

$$(2/3)\ 2^{I(2)} = 1. \qquad (3\text{-}9)\text{IV}$$

Esta última expresión nos indica que I(2) es el número de veces que se tiene que multiplicar por dos a la probabilidad inicial de m_2, para transformarla en certeza. El único valor de I(2) que cumple tal condición, es

$$I(2) = 0.5849, \qquad (4\text{-}9)\text{IV}$$

que es por definición la información en bits, que proporciona su recepción efectiva.

Puesto que $0.5849 \angle 1.5849$, X recibe menos información en este último caso, ya que la incertidumbre removida es menor por ser mayor la probabilidad de la recepción de m_2.

T E R C E R A P A R T E

L A R E G U L A C I O N

V

EL SISTEMA DE COMUNICACION Y EL CONCEPTO DE ENTROPIA.

1. Para que sea posible el análisis del flujo de información, y la integración de los elementos de que disponemos en una síntesis que haga inteligible el proceso de regulación, es preciso sustituir las partes reales de que consta un sistema de comunicación, por elementos idealizados que posibiliten su eficiente manipulación conceptual. Estas nociones abstractas a que nos referimos, son los modelos matemáticos correspondientes a los distintos bloques del diagrama (1-6)IV; de ellos, sólo trataremos a los que tienen importancia básica, que son: la fuente, el transmisor y el canal.

2. Un SISTEMA ESTOCASTICO, es un sistema dinámico cuya salida presenta estados, que aquí conviene llamar símbolos, con una probabilidad definida.

Los sistemas indeterminados que hemos definido como sistemas de Markoff, son sistemas estocásticos. Si las probabilidades de sus estados son constantes e independientes, el sistema quedará definido por el espacio muestral de su alfabeto, y la distribución fija de probabilidades correspondiente; dándosele el nombre de SISTEMA DE MARKOFF DE MEMORIA CERO.

Un caso más interesante se presenta cuando las probabilidades de los distintos símbolos son función de los símbolos emitidos anteriormente; si esta dependencia sólamente se extiende al símbolo emitido inmediatamente antes, el sistema se llama sistema o FUENTE DE MARKOFF DE PRIMER ORDEN; si abarca a los dos símbolos precedentes, será de SEGUNDO ORDEN, y en general se denominará de orden m, si la probabilidad de emitir cada símbolo está condicionada por la emisión de los m símbolos emitidos inmediatamente antes, y llamándose entonces ESTADO del sistema, en un momento dado, a esa secuencia de m símbolos emitida hasta el momento.

Aquí nos será suficiente el empleo de los sistemas de Markoff de -

primer orden, cuyo estudio ya hemos hecho en la Segunda Parte.

Dentro de los sistemas de Markoff, el CASO ERGODICO es aquel en donde la observación de la salida del sistema, durante un lapso de tiempo de suficiente duración, permitirá registrar una secuencia característica de símbolos, que se presenta con probabilidad unitaria. Por suerte para la ciencia y la tecnología, es relativamente difícil descubrir sistemas que no posean este tipo de homogeneidad estadística.

3. La Fuente de Información.- Como ya se insinuó, una FUENTE DE INFORMACION F, es un sistema estocástico absoluto, que genera símbolos por selección, y de acuerdo con cierta ley, de un espacio muestral que constituye su ALFABETO. Este alfabeto fuente, está formado por un número limitado de elementos invariables que son los SIMBOLOS de la fuente, o SIMBOLOS MENSAJE.

El caso más simple es la fuente de Markoff de memoria cero, que queda totalmente especificada mediante su alfabeto

$$F_o = (s_1, s_2, \ldots, s_i, \ldots, s_n), \qquad (1\text{-}3)V$$

y su distribución de probabilidades

$$P(s_1), P(s_2), \ldots, P(s_i), \ldots, P(s_n). \qquad (2\text{-}3)V$$

El objetivo principal de analizar mediante modelos teóricos a las fuentes de información, consiste en determinar la manera que nos permita asignarles una medida a su capacidad informativa. Ya se ha visto que es posible medir la información que da la ocurrencia de un suceso aislado, e igualmente se podría medir, mediante la expresión (1-7)IV, la información contenida en una fotografía, en un disco fonográfico, en un periódico, o en una imagen estática de televisión, pero ¿cómo valorar la información en situaciones que presentan un dinamismo, una secuencia de sucesos, tal como lo hacen las fuentes con las que estamos tratando? Esperamos poder contestar satisfactoriamente esta pregunta, con lo que sigue.

En la definición de una fuente de memoria cero, se tiene un espa-

cio muestral y un conjunto de probabilidades asociadas $P(s_i)$, con las que, a través del empleo de la expresión (1-7)IV, es posible calcular la cantidad de información asignada a la ocurrencia de cada símbolo; pero para tener idea de la capacidad de algo, se requiere un número y no una colección de ellos. Inmediatamente se piensa en un valor promedio, pero obviamente no puede tomarse el mero promedio aritmético como valor representativo de todo el conjunto, porque en el caso general cada símbolo tendrá un nivel distinto de probabilidad, y no se podrá tomar en cuenta uniformemente la información que proporciona un símbolo que aparece rara vez, y la información que da otro que se presenta con mucha frecuencia; sino que el valor medio que da una idea más fidedigna de la capacidad informativa de la fuente, tomada en conjunto, es aquel que valora la cantidad de información correspondiente a cada símbolo, proporcionalmente a la probabilidad con que es emitido. A la expresión resultante

$$H(F_o) = \sum_{i=1}^{n} P(s_i) \log_2 \frac{1}{P(s_i)}, \qquad (3\text{-}3)V$$

que también se acostumbra escribir como

$$H(F_o) = - \sum_{i=1}^{n} P(s_i) \log_2 P(s_i), \qquad (3'\text{-}3)V$$

se le llame **ENTROPIA** de la fuente de memoria cero, y el número de bits que resulten de su aplicación, representa la cantidad media de información por símbolo que suministra la fuente; o también se le puede interpretar como el valor promedio de la incertidumbre de un observador de la fuente, antes de conocer su salida.

La expresión (3-3)V es el eje de la teoría de la información y la base para hacer cálculos más complicados. El símbolo H para la entropía y la palabra misma, pertenecen a la termodinámica, en cuyo enfoque estadístico aparece una expresión semejante. Aunque estrictamente, entropía en su sentido termodinámico, e información en el sentido ya definido, son conceptos diferentes, se acostumbra emplearlos indistintamente;

en realidad son opuestos, ya que un sistema aislado tiende generalmente a aumentar su entropía a medida que transcurre el tiempo, involucionando espontáneamente hacia un estado cada vez más desorganizado, más homogéneo, más indiferenciado y consecuentemente más probable, y por este mismo aumento en la uniformidad, el sistema va perdiendo, a medida que esto sucede, su organización y su capacidad informativa. Por esta razón Wiener dice: "Just as the amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its degree of disorganization; and the one is simply the negative of the other". De manera que para Wiener, si llamamos I a la información y H a la entropía, ambas están relacionadas por la expresión

$$\frac{\Delta I}{\Delta H} = -1, \qquad (4\text{-}3)V$$

lo que ha originado sugerencias para llamar a la información en su sentido técnico, "negoentropía" o "anatropía" para indicar una medida del grado de organización que posea un sistema dinámico.

La explosión atómica en Hiroshima, popularizó la ecuación de equivalencia de A. Einstein, entre masa y energía, sin embargo, si se tuviera que elegir un símbolo representativo del adelanto científico de la época, creemos que sería más adecuado adoptar este de la entropía, a causa de su trascendencia y extensión, pues como lo expresa W. Weaver al recordar una cita de Eddington: "... not only speaks the language of arithmetic; it also speaks the language of language".

La fuente de orden cero es demasiado simple, su imagen la daría un niño pequeño que tecleara en una máquina de escribir en forma totalmente arbitraria, pero con una preferencia constante para cada letra. Pero las fuentes de primer orden ya tienen interés práctico, su modelo teórico es la matriz de transiciones probables dada en 13.III, y que describe a un sistema indeterminado. En forma compacta, la matriz (1-13)III se representa mediante la probabilidad condicional

$$P(a_q \mid a_r), \qquad (5\text{-}3)V$$

que indica la probabilidad de que la transición, estando el sistema en el estado a_r, sea al estado a_q; o equívalentemente, la probabilidad de que si el último símbolo emitido fue el a_r, el inmediato siguiente sea el a_q.

La determinación de la entropía asociada a este tipo de fuentes, requiere la aplicación de (3-3)V al espacio muestral formado por cada columna de la matriz de transiciones probables. Así, si se tiene una fuente de primer orden con un alfabeto de tres símbolos mensaje

$$F_1 = (s_1, s_2, s_3), \qquad (6\text{-}3)V$$

las entropías correspondientes a cada columna de transiciones probables, de acuerdo con (3-3)V, serán:

$$H(s_i \mid s_1) = \sum_{i=1}^{3} P(s_i \mid s_1) \log_2 \frac{1}{P(s_i \mid s_1)},$$

$$H(s_i \mid s_2) = \sum_{i=1}^{3} P(s_i \mid s_2) \log_2 \frac{1}{P(s_i \mid s_2)}, \qquad (7\text{-}3)V$$

$$H(s_i \mid s_3) = \sum_{i=1}^{3} P(s_i \mid s_3) \log_2 \frac{1}{P(s_i \mid s_3)}.$$

Estos valores de H se toman en cuenta para obtener un valor total promedio de acuerdo con la distribución estacionaria de probabilidades; si la fuente es ergódica, estas probabilidades de estado estable para cada símbolo se pueden calcular a partir de las probabilidades condicionales. Es decir, si se conoce $P(s_i \mid s_j)$, se pueden calcular los valores de $P(s_j)$; y para el caso de n=3, se tendrá:

$$H(s_i \mid s_j) = H(F_1) = P(s_1) \sum_{i=1}^{3} P(s_i \mid s_1) \log_2 \frac{1}{P(s_i \mid s_1)} +$$

$$+ P(s_2) \sum_{i=1}^{3} P(s_i \mid s_2) \log_2 \frac{1}{P(s_i \mid s_2)} +$$

$$+ P(s_3) \sum_{i=1}^{3} P(s_i \mid s_3) \log_2 \frac{1}{P(s_i \mid s_3)}. \qquad (8\text{-}3)V$$

Y en general, la entropía de una fuente de primer orden expresada en bits por símbolo, o por etapa de funcionamiento, está dada por:

$$H(F_1) = \sum_{j=1}^{n} P(s_j) \sum_{i=1}^{n} P(s_i \mid s_j) \log_2 \frac{1}{P(s_i \mid s_j)}. \qquad (9\text{-}3)V$$

Dos ejemplos ilustrarán el método de aplicación.

Sea una fuente, cuya matriz es la misma del caso concreto presentado en 14.III.

	S	N	L
S	1/4	1/3	1/8
N	5/8	1/2	1/4
L	1/8	1/6	5/8.

Las entropías correspondientes a cada columna, son:

$$H(S) = \frac{1}{4}\log_2 4 + \frac{5}{8}\log_2 \frac{8}{5} + \frac{1}{8}\log_2 8 = 1.298 \text{ bits/etapa}$$

$$H(N) = \frac{1}{3}\log_2 3 + \frac{1}{2}\log_2 2 + \frac{1}{6}\log_2 6 = 1.459 \text{ bits/etapa}$$

$$H(L) = H(S) = 1.298 \text{ bits/etapa.}$$

Supóngase ahora que se dispone de 1000 réplicas iguales de este sistema, cada una presentando uno cualquiera de los tres estados. Como son fuentes ergódicas, si se ponen a funcionar y se deja que lo hagan durante un lapso de tiempo suficiente, de manera que ocurra un gran nú-

mero de transiciones de estado, llegará un momento en que se estabilice el número de réplicas que presenten o emitan cada símbolo; cada réplica está cambiando constantemente de estado, pero en conjunto, habrá un número fijo de ellas que presenten constantemente cada uno de los estados. Si se toma el ejemplo original de 14.III, esto significa que el estado del tiempo cambia de un día a otro, pero la observación hecha durante varios años (1000 meses X), hará patente que el número relativo de días en que hay sol, nubes o lluvia, se mantiene constante.

Si llamamos s_o, n_o y l_o al número de réplicas de la fuente que se encuentren en un momento dado en los estados S, N y L respectivamente, y s, n y l es el número de las mismas en la etapa siguiente; entonces, teniendo en cuenta los datos contenidos en la matriz de transiciones probables, se puede establecer el siguiente sistema de ecuaciones:

$$s = \frac{1}{4} s_o + \frac{1}{3} n_o + \frac{1}{8} l_o$$

$$n = \frac{5}{8} s_o + \frac{1}{2} n_o + \frac{1}{4} l_o$$

$$l = \frac{1}{8} s_o + \frac{1}{6} n_o + \frac{5}{8} l_o,$$

sin embargo, en el estado de equilibrio, el número final y el número inicial se mantiene el mismo para cada estado, considerando toda la colección de réplicas; por lo tanto, en estas condiciones, el sistema de ecuaciones es:

$$\frac{3}{4} s - \frac{1}{3} n - \frac{1}{8} l = 0$$

$$-\frac{5}{8} s + \frac{1}{2} n - \frac{1}{4} l = 0$$

$$-\frac{1}{8} s - \frac{1}{6} n + \frac{3}{8} l = 0.$$

De estas ecuaciones, sólamente dos son independientes, pero la tercera ecuación la proporciona el hecho de que la suma de las réplicas que se encuentran en cada estado, debe ser 1000, o sea

$$s + n + l = 1000.$$

La solución del sistema, da:

$$s = 252.253$$

$$n = 459.459$$

$$l = 288.288,$$

que corresponde a $P(s_j)$ en la expresión (9-3)V. Con estos datos ya se puede valuar H, o sea, la información media por etapa, o por símbolo, que proporciona esta fuente, que es

$$H = \frac{s\,H(S) + n\,H(N) + l\,H(L)}{s + n + l},$$

y que numéricamente da

$$H = 1.372 \text{ bits/símbolo.}$$

Veamos ahora el caso de una población urbana de 10^6 habitantes de la que, como resultado de una encuesta realizada para el proyecto de ciertas obras públicas, se obtuvo el siguiente modelo teórico:

	T	C	E
T	7/8	0	1/4
C	0	4/5	1/2
E	1/8	1/5	1/4

Los símbolos T, C y E representan el estado de los habitantes de esa población, en su trabajo, en su casa y en la calle repectivamente; y la matriz describe las transiciones de un estado a otro. Se trata de calcular la H de este sistema.

Procediendo análogamente al caso anterior, se calcula primero la entropía correspondiente a cada columna de la matriz, lo que da

$$H(T) = \frac{7}{8} \log_2 \frac{8}{7} + \frac{1}{8} \log_2 8$$

$$= 0.542 \text{ bits/etapa.}$$

$$H(C) = \frac{4}{5} \log_2 \frac{5}{4} + \frac{1}{5} \log_2 5$$

$$= 0.721 \text{ bits/etapa.}$$

$$H(E) = \frac{1}{4} \log_2 4 + \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4$$

$$= 1.500 \text{ bits/etapa.}$$

También mediante un razonamiento semejante, se llega a la formulación del siguiente sistema de ecuaciones:

$$-\frac{1}{4} e + \frac{1}{8} t = 0$$

$$-\frac{1}{2} e + \frac{1}{5} c = 0$$

$$e + t + c = 10^6;$$

su solución

$$e = \frac{2}{11} 10^6 = 181819$$

$$t = \frac{4}{11} 10^6 = 363636$$

$$c = \frac{5}{11} 10^6 = 454545,$$

nos indica el número de habitantes (no los mismos de una etapa a otra), que constantemente se encuentran en cada lugar (calle, trabajo o casa), o bien, el tiempo que un habitante promedio pasa en cada uno de estos - lugares; y en consecuencia, la H media de todo el sistema será:

$$H = \frac{t\,H(T) + c\,H(C) + e\,H(E)}{t + c + e},$$

que numéricamente es

$$H = 0.798 \text{ bits/etapa.}$$

4. El Transmisor.- La salida de una fuente generadora de informa-- ción es recibida por el transmisor, que la transforma en una señal apta para pasar a través del canal de comunicación. La codificación que e-- fectúa el transmisor es una transformación, y puede ser especificada -- como tal, y así como se ha representado mediante un modelo teórico a -- los elementos generadores de información, también el transmisor se pue- de representar convenientemente mediante un transductor, según se le -- definió en 3.II. En efecto, un transmisor, esencialmente, no es más que - un transductor que recibe como entrada una secuencia de símbolos, y que entrega a su salida otra secuencia que es función de la primera.

En relación con este proceso de cifrado, es necesario hacer una -- advertencia para no interpretarlo en un sentido militar, y por tanto, de- masiado estrecho, que tuviera como finalidad principal el ser de difi-- cil decodificación por parte de receptores extraños; pues ya que excep- tuando quizá a la dudosa posibilidad de una transmisión directa del -- pensamiento, no hay un sólo caso de interés práctico en que lo generado por la fuente sea directamente lo que viaja por el canal hacia el ele-- mento receptor, es decir, las corrientes variables que circulan por la línea telefónica, no son los sonidos de la voz; los conceptos no son -- las palabras con que se les expresa; la imágen de televisión se parece muy poco a las alteraciones electromagnéticas que difunde la antena -- emisora; etc., aunque claro, entre señal y mensaje existen relaciones -

estrechas, que van desde la simple sustitución de símbolos, hasta complejos cambios en el patrón de estructura. En conexión con esto, es interesante señalar que la traducción de un idioma a otro es un proceso de decodificación (o de codificación, según se le vea) bastante más complicado que el deciframiento de una clave basada sólo en la transposición de términos, ya que la existencia de expresiones idiomáticas y los diferentes patrones de estructura gramatical, hace insuficiente el mero empleo de un diccionario común. Por lo demás, creemos que es inútil subrayar la importancia de este proceso, tanto en sentido directo como inverso, pues un geólogo que explora, un reportero que describe, un médico que ausculta, un instrumento que registra, una sensación que se percibe, son todas situaciones que se refieren a formas codificadas de hechos o fenómenos no directamente accesibles, pero sujetas todas, esencialmente, al mismo tipo de leyes y relaciones.

Un transmisor, al igual que cualquier transductor, establece una correspondencia entre los símbolos de un alfabeto de entrada

$$S = (s_1, s_2, \ldots, s_m),$$

llamado **ALFABETO FUENTE**, y los símbolos o secuencias de ellos de un alfabeto de salida

$$T = (t_1, t_2, \ldots, t_n),$$

llamado **ALFABETO CODIGO**.

Cada símbolo de S constituye un **SIMBOLO MENSAJE**.

Cada secuencia de símbolos de T, que corresponde a cada símbolo mensaje, se le denomina **PALABRA CODIGO**.

En el transductor

A:	d	e	f
B_1	f	d	e
B_2	d	f	e
B_3	e	f	f,

presentado en 6.II, el alfabeto fuente es

$$(B_1, B_2, B_3),$$

y el alfabeto código es

$$(d, e, f).$$

Es obvio que la condición suficiente para que la codificación no produzca pérdida de información, es que la transformación correspondiente sea biunívoca.

5. <u>El Canal</u>.- La función de un canal de comunicación no es generar información, ni codificarla, sino transferirla de un punto a otro. Los canales son elementos muy versátiles en su aspecto externo, aunque no en sus funciones; pueden ser algo tan definido y concreto como un cable telefónico o una línea telegráfica, o tan difuso como el aire atmosférico, o tan primitivo como un mensajero que lleva un recado verbal, o bien, tan inmaterial como la anchura de banda de una radioemisora; y así mismo, sus extremos pueden cubrir distancias tan cortas como las requeridas por los pequeños sistemas de T. V. de circuito cerrado, o aún menores, o tan amplias como la que separa la Luna de la Tierra, y aún mayores.

Para fines analíticos, un canal queda definido por una matriz del tipo (1-13)III, encontrándonos con que esencialmente se trata también de un transductor.

Al abstraer los elementos esenciales que toman parte en la composición y funcionamiento de un canal, se distingue un <u>alfabeto de entrada</u>

$$A = (a_1, a_2, \ldots, a_i, \ldots, a_r), \qquad (1\text{-}5)V$$

un <u>alfabeto de salida</u>

$$B = (b_1, b_2, \ldots, b_j, \ldots, b_s), \qquad (2\text{-}5)V$$

y un conjunto de probabilidades condicionales, que expresan la probabilidad de recibir un símbolo b_j cuando se ha enviado un símbolo a_i. Todo esto queda contenido en la matriz

	a_1	a_2	...	a_i	...	a_r
b_1	$P(b_1 \mid a_1)$	$P(b_1 \mid a_2)$	...	$P(b_1 \mid a_i)$	...	$P(b_1 \mid a_r)$
b_2	$P(b_2 \mid a_1)$	$P(b_2 \mid a_2)$	...	$P(b_2 \mid a_i)$	...	$P(b_2 \mid a_r)$
.	. . . .	. . . .	. . . .	. . . .	. . . .	. . . .
b_j	$P(b_j \mid a_1)$	$P(b_j \mid a_2)$	...	$P(b_j \mid a_i)$	...	$P(b_j \mid a_r)$
.	. . . .	. . . .	. . . .	. . . .	. . . .	. . . .
b_s	$P(b_s \mid a_1)$	$P(b_s \mid a_2)$	...	$P(b_s \mid a_i)$	...	$P(b_s \mid a_r)$,

(3-5)V

que se representa sintéticamente por su elemento general

$$P(b_j \mid a_i) = P_{ij}. \qquad (4\text{-}5)V$$

Conociendo $P(a_i)$ y $P(b_j \mid a_i)$, se pueden calcular las probabilidades de los símbolos de salida $P(b_j)$, mediante las expresiones

$$\begin{aligned}
P(b_1) &= P(a_1)P(b_1 \mid a_1) + P(a_2)P(b_1 \mid a_2) + \dots + P(a_r)P(b_1 \mid a_r) \\
P(b_2) &= P(a_1)P(b_2 \mid a_1) + P(a_2)P(b_2 \mid a_2) + \dots + P(a_r)P(b_2 \mid a_r) \\
&\dots\dots\dots\dots\dots\dots\dots\dots \\
P(b_s) &= P(a_1)P(b_s \mid a_1) + P(a_2)P(b_s \mid a_2) + \dots + P(a_r)P(b_s \mid a_r).
\end{aligned} \qquad (5\text{-}5)V$$

En adelante sustituiremos, por sencillez, el símbolo de interse-

ción definido en 7.III, por una coma. Entonces, teniendo en cuenta la ley de Bayes que dice que

$$P(a_i \mid b_j)P(b_j) = P(b_j \mid a_i)P(a_i) = P(a_i, b_j), \qquad (6\text{-}5)V$$

se pueden calcular las llamadas <u>probabilidades hacia atrás</u>, o sea, las probabilidades de que si se recibió un símbolo b_j, el símbolo enviado haya sido a_i. De acuerdo con (5-5)V y (6-5)V, se tiene que

$$P(a_i \mid b_j) = \frac{P(b_j \mid a_i)P(a_i)}{\sum_{i=1}^{r} P(a_i)P(b_j \mid a_i)}. \qquad (7\text{-}5)V$$

En ausencia de ruido, la entropía o información media por símbolo proporcionada a la salida del canal, sería la entropía del alfabeto de entrada

$$H(A) = \sum_{i=1}^{r} P(a_i) \log_2 \frac{1}{P(a_i)}, \qquad (8\text{-}5)V$$

pero la inevitable presencia del ruido hace que esta cantidad sea menor; vamos a determinar que tanto disminuye.

Sin ruido, la probabilidad condicional o "a posteriori" $P(a_i \mid b_j)$, valdría la unidad, significando que la observación de la salida del canal da un conocimiento cierto acerca de los símbolos de entrada; pero en presencia de ruido, $P(a_i \mid b_j) < 1$, y solamente indicará la posibilidad de que si se recibió b_j, el símbolo efectivamente enviado haya sido a_i. La entropía asociada a esta última situación, para todo el alfabeto de entrada, respecto al símbolo recibido b_j, es

$$H(A \mid b_j) = \sum_{i=1}^{r} P(a_i \mid b_j) \log_2 \frac{1}{P(a_i \mid b_j)}; \qquad (9\text{-}5)V$$

y si se extiende el cálculo, hasta abarcar a todos los símbolos de sa--

lida b_j, se obtiene

$$H(A \mid B) = \sum_{j=1}^{s} P(b_j)\, H(A \mid b_j)$$

$$= \sum_{j=1}^{s} P(b_j) \sum_{i=1}^{r} P(a_i \mid b_j) \log_2 \frac{1}{P(a_i \mid b_j)} \qquad (10\text{-}5)V$$

que por (6-5)V, da

$$H(A \mid B) = \sum_{i,j=1}^{r,s} P(a_i, b_j) \log_2 \frac{1}{P(a_i \mid b_j)}. \qquad (11\text{-}5)V$$

A $H(A \mid B)$ se le llama AMBIGÜEDAD ("equivocation") de A respecto a B, y representa la incertidumbre media por símbolo, de las señales enviadas, respecto a las señales recibidas, cuando estas últimas se conocen. De manera que la cantidad útil de información que se transmite, a pesar del efecto del ruido, es

$$I(A;B) = H(A) - H(A \mid B), \qquad (12\text{-}5)V$$

que se le conoce con el nombre de INFORMACION MUTUA del canal.

Llevando (8-5)V y (11-5)V a (12-5)V, se tiene que

$$I(A;B) = \sum_{i=1}^{r} P(a_i) \log_2 \frac{1}{P(a_i)} - \sum_{i,j=1}^{r,s} P(a_i, b_j) \log_2 \frac{1}{P(a_i \mid b_j)}. \qquad (13\text{-}5)V$$

Es posible demostrar que

$$\sum_{i=1}^{r} P(a_i) = \sum_{i,j=1}^{r,s} P(a_i, b_j), \qquad (14\text{-}5)V$$

puesto que por (6-5)V

$$\sum_{i,j=1}^{r,s} P(a_i,b_j) = \sum_{i,j=1}^{r,s} P(b_j \mid a_i)\, P(a_i), \qquad (15\text{-}5)V$$

y al desarrollar el segundo miembro de (15-5)V, se obtiene, como coeficiente de cada $P(a_i)$, una probabilidad condicional que es un elemento de la columna a_i en la matriz (3-5)V, y cuya suma es, por lo expuesto en 13.III, la unidad. Por lo tanto

$$I(A;B) = \sum_{i,j=1}^{r,s} P(a_i,b_j) \log_2 \frac{1}{P(a_i)} - \sum_{i,j=1}^{r,s} P(a_i,b_j) \log_2 \frac{1}{P(a_i \mid b_j)}$$

$$= \sum_{i,j=1}^{r,s} P(a_i,b_j) \log_2 \frac{1}{P(a_i)} + \sum_{i,j=1}^{r,s} P(a_i,b_j) \log_2 P(a_i \mid b_j),$$

lo que por las leyes de los logaritmos, es equivalente a

$$I(A;B) = \sum_{i,j=1}^{r,s} P(a_i,b_j) \log_2 \frac{P(a_i \mid b_j)}{P(a_i)}. \qquad (16\text{-}5)V$$

La información mutua, si aplicamos de nuevo la ley de Bayes, se puede expresar también como:

$$I(A;B) = \sum_{i,j=1}^{r,s} P(a_i \mid b_j) P(b_j) \log_2 \frac{P(a_i \mid b_j)}{P(a_i)}. \qquad (17\text{-}5)V$$

A la información mutua también se le puede interpretar, simplemente, como la cantidad media de información que proporciona el conocimiento de la ocurrencia de un suceso b_j, respecto a la ocurrencia de otro suceso a_i no presenciado directamente; o sea, es la cantidad por la cual el conocimiento de B, disminuye o reduce la incertidumbre sobre A.

Si estos sucesos son independientes,

$$P(a_i \mid b_j) = P(a_i),$$

y (17-5)V hace ver que la información recibida es nula, tal como intuitivamente es de suponerse.

6. Otra de las relaciones que se cita con frecuencia, sin aclarar suficientemente su origen ni sus implicaciones, es la propiedad que tiene la información mutua de no ser nunca negativa. La demostración con todos sus pasos, que no es más que una versión detallada de la propuesta por Abramson, es la siguiente:

Si se trazan en un mismo sistema coordenado las gráficas correspondientes a las ecuaciones

$$y = \log_e x,$$

$$y = x - 1;$$

se hará evidente que

$$\log_e x \leq x - 1. \qquad (1\text{-}6)V$$

Por otra parte, teniendo en cuenta (1-8)IV, sabemos que

$$\sum_{i=1}^{n} P(x_i) \log_2 \frac{P(y_i)}{P(x_i)} = \frac{1}{\log_e 2} \sum_{i=1}^{q} P(x_i) \log_e \frac{P(y_i)}{P(x_i)}, \qquad (2\text{-}6)V$$

y tomando a $\frac{P(y_i)}{P(x_i)}$ de (2-6)V como la x de (1-6)V, se tiene

$$\sum_{i=1}^{n} P(x_i) \log_2 \frac{P(y_i)}{P(x_i)} \leq \frac{1}{\log_e 2} \sum_{i=1}^{n} P(x_i)\left(\frac{P(y_i)}{P(x_i)} - 1\right),$$

$$\sum_{i=1}^{n} P(x_i) \log_2 \frac{P(y_i)}{P(x_i)} \leq \frac{1}{\log_e 2} \left[\sum_{i=1}^{n} P(y_i) - \sum_{i=1}^{n} P(x_i) \right];$$

además, por (B-10)III, se infiere que

$$\sum_{i=1}^{n} P(x_i) \log_2 \frac{P(y_i)}{P(x_i)} \leq 0. \qquad (3\text{-}6)V$$

Si ahora se aplica (H-10)III a (16-5)V, se obtiene

$$I(A;B) = \sum_{i,j=1}^{r,s} P(a_i,b_j) \log_2 \frac{P(a_i,b_j)}{P(a_i)P(b_j)}, \qquad (4\text{-}6)V$$

que por las propiedades de los logaritmos, se puede escribir también como

$$I(A;B) = - \sum_{i,j=1}^{r,s} P(a_i,b_j) \log_2 \frac{P(a_i)P(b_j)}{P(a_i,b_j)}; \qquad (5\text{-}6)V$$

en esta última expresión, $P(a_i,b_j)$ equivale a $P(x_i)$ de (3-6)V, y $P(a_i)P(b_j)$ a $P(y_i)$, por lo que se puede establecer que

$$I(A;B) \geq 0. \qquad (6\text{-}6)V$$

La importante conclusión que se infiere de la inecuación anterior, es que la información promedio relativa a un suceso A, nunca puede ser reducida por el conocimiento de otro suceso B, sino que, por el contrario, generalmente será posible obtener alguna cantidad positiva de información de uno respecto del otro; salvo en el caso, ya citado, de inde-

pendencia estadística de sucesos.

7. Según ya se señaló en 11.III, el concepto de probabilidad, y los resultados que se derivan de su empleo, son susceptibles de interpretarse de diferentes formas; afortunadamente, esta diversidad de opiniones interpretativas se refiere únicamente al concepto mismo y a los resultados, sin afectar en nada al cálculo a que da lugar. Cuando se adopta el punto de vista de la llamada "interpretación subjetiva", y se considera a la probabilidad de la ocurrencia de un suceso s, como una medida de la firmeza de la opinión sustentada por un observador, sobre la base de ciertos datos, se llegan a obtener conclusiones muy interesantes; en cierto modo ya hemos empleado este punto de vista en la sección 3.IV, al relacionar el estado de conocimiento del sujeto X, con la magnitud de la probabilidad de los mensajes.

En conexión con la información mutua, y con el tipo citado de interpretación, J. L. Massey hace un análisis que por su interés y capacidad ilustrativa quisieramos comentar brevemente aquí.

Si se extiende el punto de vista subjetivo al concepto de información, es más conveniente partir de la noción de información mutua, y definir a la cantidad de información como la información que proporciona la observación o conocimiento de un acontecimiento k, respecto de la ocurrencia de otro suceso j; y entonces, si el conocimiento de k duplica la probabilidad de j, se dice que la ocurrencia de k ha proporcionado UN BIT de información respecto al acontecimiento j, y en estas condiciones la expresión (1-7)IV, de la que originalmente partimos, queda reducida al caso particular de información mutua

$$I(j;k) = \log_2 \frac{P(j \mid k)}{P(j)}, \qquad (1\text{-}7)V$$

en el que la ocurrencia de k proporciona certeza respecto a la ocurrencia de j.

Cuando se tiene el caso de una situación cuyo espacio muestral se reduce a la ocurrencia de un suceso a_1, y a su no ocurrencia a_2, su distribución de probabilidades será:

pendencia estadística de sucesos.

7. Según ya se señaló en 11.III, el concepto de probabilidad, y los resultados que se derivan de su empleo, son susceptibles de interpretarse de diferentes formas; afortunadamente, esta diversidad de opiniones interpretativas se refiere únicamente al concepto mismo y a los resultados, sin afectar en nada al cálculo a que da lugar. Cuando se adopta el punto de vista de la llamada "interpretación subjetiva", y se considera a la probabilidad de la ocurrencia de un suceso s, como una medida de la firmeza de la opinión sustentada por un observador, sobre la base de ciertos datos, se llegan a obtener conclusiones muy interesantes; en cierto modo ya hemos empleado este punto de vista en la sección 3.IV, al relacionar el estado de conocimiento del sujeto X, con la magnitud de la probabilidad de los mensajes.

En conexión con la información mutua, y con el tipo citado de interpretación, J. L. Massey hace un análisis que por su interés y capacidad ilustrativa quisieramos comentar brevemente aquí.

Si se extiende el punto de vista subjetivo al concepto de información, es más conveniente partir de la noción de información mutua, y definir a la cantidad de información como la información que proporciona la observación o conocimiento de un acontecimiento k, respecto de la ocurrencia de otro suceso j; y entonces, si el conocimiento de k duplica la probabilidad de j, se dice que la ocurrencia de k ha proporcionado UN BIT de información respecto al acontecimiento j, y en estas condiciones la expresión (1-7)IV, de la que originalmente partimos, queda reducida al caso particular de información mutua

$$I(j;k) = \log_2 \frac{P(j \mid k)}{P(j)}, \qquad (1\text{-}7)V$$

en el que la ocurrencia de k proporciona certeza respecto a la ocurrencia de j.

Cuando se tiene el caso de una situación cuyo espacio muestral se reduce a la ocurrencia de un suceso a_1, y a su no ocurrencia a_2, su distribución de probabilidades será:

$$P(a_1) = e,$$

$$P(a_2) = 1 - e; \qquad (2\text{-}7)V$$

$$e \angle\angle 1.$$

Como $e \angle\angle 1$, esto significa que se tiene la casi certeza de que a_1 no ocurrirá, o equivalentemente, de que la proposición que afirma la — ocurrencia de a_1 es falsa. Es posible hacer el cálculo mediante la fórmula (17-5)V, para determinar la cantidad de información que debería — proporcionar un suceso b, a alguien que sustentara originalmente la — fuerte opinión expresada por las condiciones (2-7)V, de modo de modificarla hasta el punto de llevarla exactamente al extremo contrario.

Massey obtiene la expresión aproximada

$$I(A;b) = -\log_2 e, \qquad (3\text{-}7)V$$

que muestra que tal cantidad necesaria, es bastante pequeña.

Los detalles de la demostración son los siguientes:

Por hipótesis

$$P(a_1) = e; \quad P(a_2) = 1 - e.$$

$$P(a_1 \mid b) = 1 - e; \quad P(a_2 \mid b) = e. \qquad (4\text{-}7)V$$

Si en (17-5)V, B se reduce a un solo elemento b, entonces

$$I(A;b) = (1 - e) \log_2 \frac{1 - e}{e} + e \log_2 \frac{e}{1 - e}, \qquad (5\text{-}7)V$$

o sea,

$$I(A;b) = (1-e) \log_2(1-e) - (1-e) \log_2 e + e\log_2 e - e\log_2(1-e),$$

$$= (1-e) \log_2(1-e) + (2e-1)\log_2 e - e\log_2(1-e); \qquad (6\text{-}7)V$$

pero, puesto que $e \doteq 0$, entonces $1-e$ es casi la unidad, y los términos en que aparece el $\log_2(1-e)$ se pueden despreciar; asímismo, $(2e-1)$ es prácticamente igual a -1, y $(6-7)V$ se reduce a la expresión aproximada $(3-7)V$.

8. En las aplicaciones de la interpretación subjetiva, el calificativo de "subjetiva" no debe entenderse en el sentido emocional de una creencia, que como tal, no constituye base firme para ningún cálculo, sino en el sentido de que el conocimiento efectivo de ciertos hechos dan el fundamento para que un sujeto sustente una opinión con mayor o menor intensidad; aunque comunmente, claro está, el grado de conocimiento que se tiene sobre un hecho, y el sentimiento de seguridad o de duda que suscita, están relacionados.

Otra situación a la que es aplicable la misma fórmula, se refiere al cálculo de la cantidad de información que se requiere para reforzar o debilitar ligeramente una opinión muy categórica. Suponiendo que e_1 y e_2 son magnitudes positivas y diferentes, pero muy próximas a cero, la situación original sería descrita como

$$\begin{aligned} P(a_1) &= e_1, \\ P(a_2) &= 1 - e_1; \end{aligned} \qquad (1-8)V$$

y después del conocimiento del suceso b, la situación cambia a

$$\begin{aligned} P(a_1 \mid b) &= e_2, \\ P(a_2 \mid b) &= 1 - e_2. \end{aligned} \qquad (2-8)V$$

El cálculo de $I(A;b)$, revela que la cantidad de información que debe proporcionar la observación del suceso b, para debilitar la opinión, de modo que $e_2 > e_1$, es muchas veces mayor que la cantidad necesaria para realizar el proceso contrario, en que $e_2 < e_1$.

Como el cálculo numérico de expresiones que involucran magnitudes muy pequeñas es muy laborioso, especialmente si contienen logaritmos, - ya que no son comunes las tablas con mantisas de más de seis cifras, — Massey propone la siguiente expresión, también aceptablemente aproximada:

$$I(A;b) = 1.45\ e_1 + e_2 \log_2 \frac{1}{e_1}. \qquad (3\text{-}8)V$$

Aunque por lo demás, no tiene caso hacer el cálculo para casos — concretos, ya que lo interesante es la conclusión general a que da lugar: la facilidad que existe para reforzar las convicciones pre-existentes, y la imposibilidad teórica para descartar totalmente una certeza subjetiva ya establecida, puesto que esto requeriría de b el suministro de una cantidad infinita de información, según se desprende de - (3-8)V. Esto proporciona un fuerte factor explicativo de porqué es perder el tiempo el tratar de argumentar con cierta clase de personas, o - bajo ciertas circunstancias, y también hace patente la conveniencia, — dada la esencial irreversibilidad de este proceso, de limitar las auténticas convicciones a aspectos fundamentales, si se ha de mantener la flexibilidad intelectual frente a nuevos hechos. Massey cita además, y con mucho acierto, la intuición por parte de Platón de este principio, según lo revela la forma en que expone sus puntos de vista doctrinarios, en "La República".

VI

EL FLUJO DE INFORMACION Y LOS PROCESOS DE REGULACION.

1. Se suele hacer en ingeniería electromecánica, la distinción entre sistemas de potencia y sistemas de comunicación. Los que diseñan sistemas del primer tipo, norman su criterio sobre la base de una máxima eficiencia y un mínimo de fluctuaciones; mientras que por el contrario, los sistemas de comunicación cumplen su función tanto mejor, cuanta de mayor variabilidad o carencia de uniformidad se pueda disponer, por las razones ya expuestas. La calidad de un sistema de potencia se puede apreciar por su eficiencia en términos energéticos, y por su regulación; la de un sistema de comunicación, por su capacidad informativa en términos de entropía. Los sistemas de distribución de potencia manejan energía de ejecución, y de un volumen generalmente muy grande; los sistemas de distribución de información manejan energía de mando, de magnitud relativamente pequeña. Así por ejemplo, en un avión de propulsión a reacción, la energía de ejecución está representada por la capacidad de sus motores, que llega a varios miles de caballos de potencia, mientras que la energía primaria de mando es aportada por la energía muscular del piloto, cuyo rendimiento energético total, podría apreciarse en aproximadamente $\frac{1}{20}$ H.P. Pero a pesar de sus diferencias, ambos tipos de sistemas, y sus respectivos procesos característicos de regulación y de comunicación, mantienen una estrecha relación.

Vamos a presentar muy sucintamente un último tema, que constituye uno de los principales objetivos de todos los desarrollos anteriores: la regulación, pero bajo un aspecto que no es muy frecuente, y que exhibe magníficamente la capacidad de síntesis de la cibernética; se trata de hecho de una de las ideas más brillantes de Ashby, quien a su vez se inspiró en los trabajos de Shannon y de von Neumann. Como la naturaleza y finalidad de la obra en la que Ashby la presenta no le permite dar una exposición detallada y rigurosa, vamos a tratar de hacerlo aquí, apoyándonos en el material ya presentado.

2. La REGULACION, núcleo de los procesos de control, no es más que la búsqueda de la constancia en relación con una referencia de mando — fija o variable, frente a factores disturbio. Pero la estabilidad que - da la regulación no es estatismo, no es inacción, sino un estado de — equilibrio dinámico; se podría expresar diciendo que se trata de una — anulación del movimiento por medio del movimiento, y así es precisamente como lo comprende Ashby cuando dice: "Only variety can destroy variety", en lo que él llama la Ley de la Variedad Necesaria ("the Law of - Requisite Variety"), referente a este tema.

El caso extremo de una situación variable que hay que regular, no se presenta en relación con situaciones determinadas, como las examinadas en la Primera Parte (Cap. II), ni en relación con fenómenos de la - naturaleza, estadísticamente previsibles, como los que se vieron en la Segunda Parte (Cap. III), y que por lo tanto son solubles de una vez — por todas, sino en aquellas situaciones en las que interviene un factor inteligente de oposición, cuya finalidad expresa consiste en generar y mantener sistemáticamente un disturbio en oposición a los intentos de - regulación. Estas situaciones conflictivas con intereses cruzados que - alientan una actitud finalista y más o menos racional, representan por esto mismo una situación eminentemente humana, cuyo modelo teórico ha - sido desarrollado satisfactoriamente por J. von Neumann, bajo el título de Teoría de los Juegos de Estrategia.

La situación política de las naciones en las que surgió la ciber— nética y en las que actualmente se la desarrolla e investiga, le ha — tendido a dar a la misma una orientación bélica a causa de sus posibi— lidades en la planificación de sistemas de defensa, y porque algunos de sus temas se prestan a una aplicación directa a ciencias militares, tales como la codificación a la criptografía, y este de los juegos de habilidad, a la táctica; pero de hecho, la cibernética rebasa con mucho - el terreno militar, en el que sólo tiene las aplicaciones que se derivan de todo recurso poderoso, tal y como ocurrió con la aviación y, más re— cientemente, con las investigaciones en el campo de la física nuclear, - pero a diferencia de estos casos, las aplicaciones de la cibernética no son de orden práctico sino normativo; no sirven para hacer, sino para -

saber qué hacer. Por otra parte, su capacidad para optimizar resultados y empleo de recursos, hará que tenga un creciente papel en la solución de los intrincados problemas que plantea la economía mundial, en la que se presentan las situaciones competitivas más complejas, más permanentes y más comunes, y en la que a la larga se originan las confrontaciones diplomáticas y, finalmente, también los conflictos armados más graves.

3. Como de costumbre, para analizar el proceso de regulación, se requiere disponer de un modelo matemático. La teoría de los juegos proporciona uno adecuado.

Vamos a adoptar como prototipo, el caso de dos personas o grupos de ellas con intereses opuestos, y que llamaremos A y B, e impongamos la restricción de que los recursos o alternativas de que dispone cada participante para lograr sus objetivos no sean infinitos en número, sino limitados y bien definidos, y además, que la suma algebraica de la valuación que se haga de los resultados sea nula; lo que significa que, en la misma proporción en que A logra su finalidad, B no la alcanza, y viceversa. Estas condiciones, en la terminología de la teoría de los juegos, describen a un juego finito, de suma cero, entre dos personas; y recordemos de nuevo que al introducir en un tema típico de ingeniería, como lo es el proceso de regulación, al elemento humano, no se está tratando de antropomorfizar a ningún mecanismo inerte, sino únicamente de analizar las características del caso más general, y por tanto más difícil y desventajoso, de regulación: aquel en el cual el proceso de estabilización no se ve obstaculizado por "algo" más o menos fortuito, o constante, sino por "alguien" activo que dispone de recursos más o menos organizados y variables de oposición. La representación analítica de esta situación conflictiva, está contenida en un conjunto de elementos distribuidos en renglones y columnas, que forman una MATRIZ DE JUEGO, muy parecida a las ya empleadas aquí para otros conceptos. En ella, cada columna es un elemento de un espacio muestral

$$A = (A_1, A_2, \ldots, A_i, \ldots, A_n),$$

y cada renglón lo es de

$$B = (B_1, B_2, \ldots, B_j, \ldots, B_m),$$

constituyendo los diferentes cursos de acción o estrategias disponibles para A y para B respectivamente. Por tanto, aquí se entiende por ESTRATEGIA, simplemente una alternativa de acción, pero que por estar a su vez formada por un espacio muestral, es completa, y teóricamente no puede ser alterada por ninguna intervención deliberada o contingente, por la sencilla razón de que si efectivamente es tal, según lo dicho en 5.III, debe de abarcar todas las posibilidades significativas, y cualquier suceso, o intervención planeada para alterarla, estará tomada en cuenta como un ingrediente esencial constitutivo de la misma estrategia. Cada una de estas n estrategias potenciales de A, se pueden asociar con cada una de las de B, dando lugar a nxm resultados posibles que forman el cuerpo de la matriz de juego, que presenta el siguiente aspecto:

	A_1	A_2	$\cdots$	A_i	$\cdots$	A_n
B_1	r_{11}	r_{21}	$\cdots$	r_{i1}	$\cdots$	r_{n1}
B_2	r_{12}	r_{22}	$\cdots$	r_{i2}	$\cdots$	r_{n2}
$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdots$	$\cdot$
B_j	r_{1j}	r_{2j}	$\cdots$	r_{ij}	$\cdots$	r_{nj}
$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdots$	$\cdot$
B_m	r_{1m}	r_{2m}	$\cdots$	r_{im}	$\cdots$	r_{nm},

(1-3)VI

en ella, cada elemento r_{ij} representa el resultado de la elección por parte de A, de la estrategia A_i, y por parte de B, de la B_j. Hay que hacer notar que tanto A como B tienen libertad para elegir cualquier

alternativa de su espacio muestral, pero cada resultado es función — tanto de A como de B, no bastando para definirlo la intervención de un solo participante. En estas condiciones, el objeto de la teoría consiste en determinar la actitud óptima que, sobre los datos disponibles, — proporcione el máximo resultado teóricamente posible, suponiendo a las dos partes igualmente hábiles.

4. Si valuamos numéricamente a los resultados, y le adjudicamos a A el objetivo de obtener y mantener en cada etapa o elección de un partido, el resultado más alto posible contra la oposición de B, que tiene naturalmente el mismo objetivo, tendremos los elementos esenciales de - todo conflicto y la descripción analítica del caso más general de regulación; pero, para simplificar las cosas, podemos suponer que A trata de recibir una cantidad máxima, mientras que B se esfuerza por proporcio— nar una cantidad mínima; en estas condiciones, la inspección por parte de A de la matriz, le llevará a tomar en cuenta el valor mínimo de cada columna y a elegir de entre ellos la estrategia correspondiente al valor máximo. En forma análoga, B tomará nota del valor máximo que le — proporciona cada renglón, y de entre ellos eligirá el mínimo. De este - modo A se asegura una cierta ganancia mínima, y B, desde su propio punto de vista, también se garantizará que sus pérdidas no excederán a cierta cantidad. En el caso no muy frecuente de que coincidan los resultados - así elegidos, se tendrá un minimáximo o PUNTO DE SILLA ("saddle point": nombre adoptado por von Neumann en atención a su aspecto en una repre— sentación gráfica tridimensional), que define las respectivas estrate— gias óptimas a seguir; si A o B se aparta de ella, obtendrá un resultado inferior al que teóricamente podría obtener, pero el conocimiento — previo de la estrategia adoptada por el contrario no reporta ningún — beneficio adicional. En el caso más común, de que no exista punto de — silla, la actitud óptima requiere la adopción de estrategias mixtas y - el cálculo de valores expectados; el conocimiento previo de los planes de la parte contraria proporciona ventaja, por lo que desempeñan un papel las medidas de seguridad, introduciéndose en las elecciones un factor de azar, pero los resultados medios obtenidos se independizan de la actitud del adversario, cuyos errores dejan de proporcionar cantidades

adicionales. Todo esto da lugar a una teoría asociada bastante extensa y complicada; sin embargo, nuestro interés aquí no está orientado a la teoría de los juegos en cuanto tal, sino más bien a su capacidad para representar funcionalmente al proceso de la regulación, y así como se ha podido determinar una medida para valuar la capacidad informativa de un sistema, ahora se trata de definir un criterio adecuado para apreciar la capacidad reguladora de un elemento regulador. Al realizar esto, se habrá logrado poner de manifiesto la unidad esencial de las tres grandes teorías involucradas: la de los juegos, la de la regulación y la de la información.

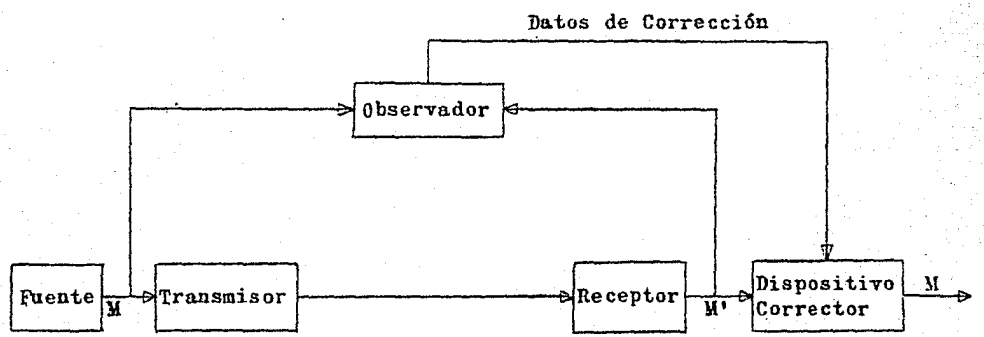
5. En el problema de la regulación se tiene habitualmente como dato un sistema, cuya salida r se trata de mantener dentro de un conjunto de valores satisfactorios, fijados por una unidad de mando o control C, independientemente de la influencia de una fuente B generadora de disturbios, que trata de sacar a r de su rango de valores; y además, se tiene también como dato un conjunto m de circunstancias y características referentes al sistema por regular y al sistema generador de disturbios, que incluye leyes naturales y características del ambiente, y cuyo modelo matemático está dado por la matriz (1-3)VI. En estas condiciones, se trata de integrar a todos estos elementos un nuevo elemento A, que reciba los disturbios de B, y mediante una transformación adecuada los recodifique en combinación con m, de manera que el resultado r, se mantenga constantemente dentro de los límites fijados por la unidad de control C, a pesar de la influencia de B. Desde el punto de vista de la información se trata simplemente, y refiriéndonos al diagrama (1-6)IV, de que la información generada por la fuente F alcance íntegra su destino D, independientemente de la presencia de la fuente de ruido R, que trata de contaminar la señal original; y el problema consiste entonces, en el diseño de un canal de corrección que restaure la fidelidad.

6. El Décimo Teorema de Shannon establece que:

"Si el canal de corrección tiene una capacidad igual a la ambigüedad $H(A \mid B)$, es posible codificar los datos de corrección para enviar-

los a través de este canal y corregir la información recibida, exceptuando una fracción tan pequeña como se quiera de los errores. Esto no será posible, si la capacidad del canal es menor de H(A | B)".- Y agrega Shannon:

"Aproximadamente entonces, H(A | B) es la cantidad adicional de información que debe ser suministrada por segundo al punto receptor para corregir el mensaje recibido". E ilustra su teorema con el siguiente diagrama esquemático:

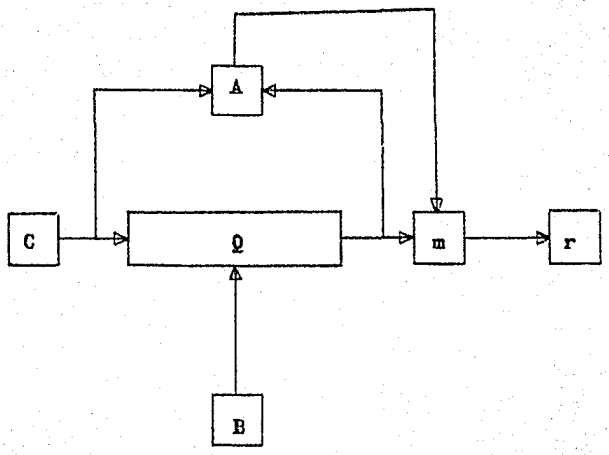


(1-6)VI

Este diagrama y este teorema los tomaremos como base para establecer formalmente la analogía de Ashby, y definir una medida para la capacidad reguladora.

En el diagrama (1-6)VI, se quiere indicar que en un sistema de comunicación afectado por la presencia de ruido, un observador que presenciara simultáneamente la entrada y la salida, alterada por el ruido, del canal, podría transmitir al punto receptor, a través de un canal adicional, las correcciones necesarias para restaurar el mensaje original.

Estos mismos conceptos, transferidos al proceso de regulación, nos permite presentar el siguiente diagrama equivalente,



(2-6)VI

en el que:

C = Unidad de control (o fuente de información).

A = Sistema regulador (u observador del canal con ruido).

Q = Sistema cuya salida se desea controlar (o canal con ruido).

B = Fuente de disturbios (o fuente de ruido).

m = Matriz de juego, elemento generalmente intangible o integrado a la misma estructura del sistema regulador, y que define el resultado de la acción conjunta del disturbio y del efecto regulador;(o dispositivo corrector).

r = Salida del sistema controlado (o mensaje corregido).

7. Por definición, la entropía AFIN, o incertidumbre del suceso simultáneo (A_i, B_j), está expresada por

$$H(A,B) = \sum_{i,j=1}^{n,m} P(A_i, B_j) \log_2 \frac{1}{P(A_i, B_j)}, \qquad (1-7)VI$$

que por las leyes de los logaritmos también se puede expresar como

$$H(A,B) = \sum_{i,j=1}^{n,m} P(A_i,B_j) \log_2 \frac{P(A_i)P(B_j)}{P(A_i,B_j)} + \sum_{i,j=1}^{n,m} P(A_i,B_j) \log_2 \frac{1}{P(A_i)P(B_j)}. \qquad (2\text{-}7)\text{VI}$$

Por otra parte sabemos, teniendo en cuenta (4-6)V, que

$$\sum_{i,j=1}^{n,m} P(A_i,B_j) \log_2 \frac{P(A_i)P(B_j)}{P(A_i,B_j)} = - \sum_{i,j=1}^{n,m} P(A_i,B_j) \log_2 \frac{P(A_i,B_j)}{P(A_i)P(B_j)} =$$

$$= - I(A;B), \qquad (3\text{-}7)\text{VI}$$

por tanto

$$H(A,B) = - I(A;B) + \sum_{i,j=1}^{n,m} P(A_i,B_j) \log_2 \frac{1}{P(A_i)} + \sum_{i,j=1}^{n,m} P(A_i,B_j)\log_2\frac{1}{P(B_j)}, \qquad (4\text{-}7)\text{VI}$$

o bien

$$H(A,B) = - I(A;B) + \sum_{i=1}^{n} P(A_i)\log_2 \frac{1}{P(A_i)} + \sum_{j=1}^{m} P(B_j)\log_2 \frac{1}{P(B_j)}; \qquad (5\text{-}7)\text{VI}$$

o sea que

$$H(A,B) = H(A) + H(B) - I(A;B), \qquad (6\text{-}7)\text{VI}$$

pero por (12-5)V

$$- I(A;B) = H(A \mid B) - H(A), \qquad (7\text{-}7)\text{VI}$$

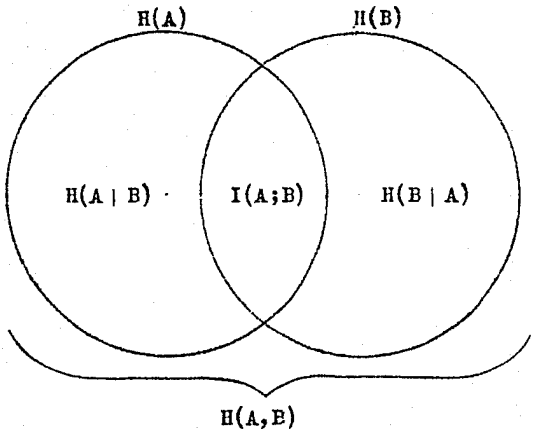
por lo que finalmente se ha demostrado que

$$H(A,B) = H(B) + H(A \mid B), \qquad (8\text{-}7)VI$$

y por simetría, también que

$$H(A,B) = H(A) + H(B \mid A). \qquad (9\text{-}7)VI$$

Abramson sugiere el siguiente diagrama como recurso mnemotécnico, y que resulta muy útil también para visualizar las relaciones entre las diversas entropías referentes a un canal.



(10-7)VI

Si ahora establecemos que en cada columna A_i de (1-3)VI no haya ningún elemento r_{ij} repetido, se tendrá que

$$H(r \mid A) \geq H(B \mid A), \qquad (11\text{-}7)VI$$

y a cada nuevo valor que adopte el disturbio o ruido B, el regulador u observador A, tendrá que saltar a una nueva columna, si ha de desempeñar su función de mantener a r en un nivel constante.

Pero por (8-7)VI y (9-7)VI, sabemos que

$$H(B) + H(A \mid B) = H(A) + H(B \mid A), \qquad (12\text{-}7)VI$$

y teniendo en cuenta la inecuación (11-7)VI, se puede también afirmar que

$$H(B) + H(A \mid B) \leq H(A) + H(r \mid A), \qquad (13\text{-}7)VI$$

lo que por (9-7)VI, equivale a

$$H(B) + H(A \mid B) \leq H(A,r); \qquad (14\text{-}7)VI$$

pero por otra parte, (6-7)VI nos dice que

$$H(A,r) \leq H(A) + H(r), \qquad (15\text{-}7)VI$$

y por consiguiente

$$H(B) + H(A \mid B) \leq H(A) + H(r), \qquad (16\text{-}7)VI$$

o sea, que finalmente

$$H(r) \geq H(B) + H(A \mid B) - H(A). \qquad (17\text{-}7)VI$$

En el caso en que $H(A \mid B) = 0$, es decir, cuando no exista incertidumbre respecto a A al conocer B, (17-7)VI se reduce a

$$H(r) \geq H(B) - H(A), \qquad (18\text{-}7)VI$$

que tiene una interpretación sencilla: la entropía de la salida de un sistema regulado, no puede ser menor que la diferencia entre la entropía de la fuente de disturbios y la entropía del sistema regulador; y en consecuencia, si $H(B)$ = cte., la única forma de mejorar la regulación consiste en incrementar la H del regulador.

Ashby llama a este principio, "Ley de la Variedad Necesaria", a

causa de que si existe equiprobabilidad en A y en B, las entropías correspondientes se reducen, según lo dicho en 7.IV, a la variedad logarítmica disponible para A y para B.

También se desprende de la inspección de (18-7)VI, que la regulación total, que implica H(r) nula, requiere que H(A) = H(B). Así pues, se ha llegado a establecer un criterio para valuar la capacidad reguladora de un sistema regulador, y además se ha logrado poner de manifiesto la naturaleza esencial del proceso de regulación: el bloqueo del flujo informativo procedente de fuentes distintas de las de mando o control, según se puede apreciar en la representación diagramática (2-6)VI, en la cual, la función de A y de m consiste en bloquear la información procedente de B, permitiendo que alcance la salida únicamente la información originada en C.

Ejemplo 1. Un piloto automático recibe información, respecto a la situación del espacio que rodea al avión en una extensión suficiente, a través de un canal múltiple formado por diez líneas diferentes, cada una de las cuales le puede proporcionar datos a razón de 100 bits/seg. Suponiéndole al avión suficiente maniobrabilidad, ¿es esta capacidad informativa suficiente para permitirle esquivar cien obstáculos potenciales diferentes, cada uno de los cuales puede estar o no presente en cada segundo?

Evidentemente esta capacidad es más que suficiente, puesto que en este caso

$$H(B) = 100 \text{ bits/seg.},$$

y

$$H(A) = 1000 \text{ bits/seg.}$$

Ejemplo 2. El dictador de un cierto país hipotético, que cuenta con una población civil de 10^7 habitantes, trata de controlar las actividades de todos y cada uno de los ciudadanos, a cada uno de los cuales se le puede estimar una actividad significativa promedio de 10^4 bits/día. Si dispone de una corporación de 10^4 agentes -no contados en la pobla-

ción citada- cada uno de los cuales está en posibilidad de reportear información a una agencia central, a razón de 10^2 bits/hora durante 10 horas diarias, y se le supone al dictador capacidad para asimilar toda la información recibida, es decir, la capacidad para tomar decisiones e imponerlas tomando en cuenta hasta el último bit de información; ¿le será suficiente la tal organización para estar en posibilidad de controlar todas las actividades diarias en todo el país?

Obviamente no, puesto que

$$H(A) = 10^7 \text{ bits/día.},$$

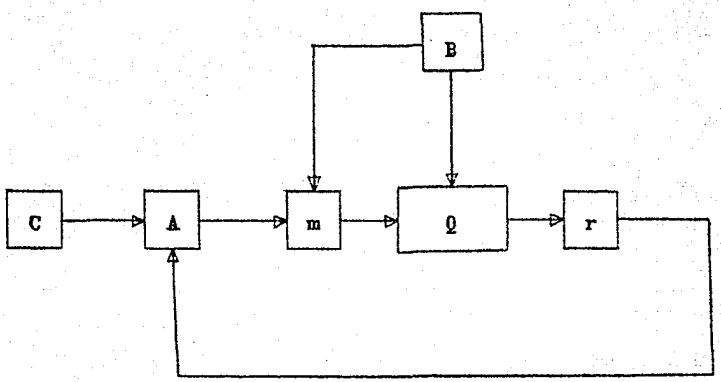
y

$$H(B) = 10^{11} \text{ bits/día.}$$

Le sería necesaria una mínima capacidad informativa 10 000 veces mayor. Si esto lo intentara resolver simplemente reclutando de algún otro lugar exterior más agentes, requeriría, desde el punto de vista estrictamente informativo, por lo menos diez agentes para vigilar y reportear las actividades de cada ciudadano.

8. En el sistema de control que describe el diagrama (2-6)VI, elementos sensores detectan el disturbio antes de que se actualice su efecto a la salida del sistema regulado, y de esta manera la regulación puede ser teóricamente perfecta. Pero hay un caso muy frecuente, en el que la información referente al disturbio no llega al elemento regulador A, sino hasta que el disturbio se ha hecho efectivo en la salida r del sistema, en calidad de ruido o error, siendo precisamente esta señal de error, o diferencia que existe entre la señal recibida y la orden original, la que inicia la acción compensadora. Por ser en estos casos la misma salida r, la portadora de la señal informativa para A, se tiene que permitir primero al disturbio que se manifieste, para saber que existe, y luego anularlo; y por tanto, la condición para que funcione la regulación es que esta no sea completa.

El diagrama correspondiente es el siguiente,



(19-7)VI

en el que se reconoce el característico circuito retroactivo de un servomecanismo. Para reducir al mínimo las variaciones permitidas, se suele agregar un elemento amplificador que hace actuar al regulador en forma más violenta y desde que el disturbio es incipiente, pero esto origina a su vez la inestabilidad oscilatoria del sistema, por sobrecompensación, y presentándose así el conocido conflicto de diseño entre sensibilidad y estabilidad, que requiere soluciones de compromiso a través del empleo de técnicas ya muy estudiadas y difundidas.

C O N C L U S I O N E S.

1. Estamos convencidos de que los conocimientos, si han de ser útiles, deben ser concisos y jerarquizados; es inútil o de resultados autoengañosos el intentar estudios avanzados sobre fundamentos ambiguos, contradictorios o decididamente falsos, asímismo, no puede haber motivación suficiente para una labor creadora, si en el estudio de una materia tan altamente abstracta como la presente, no se da primero un vistazo de conjunto que defina objetivos y promueva intereses; a nadie que pueda pensar con independencia le agrada que lo conduzcan sin explicaciones, bajo la promesa de una justificación posterior. Estas son las ideas que han impulsado la realización de este trabajo, porque es un hecho que el técnico o el experto sin un sólido núcleo de preparación teórico-científica, es un obrero que actúa pero no comprende, que puede tener una eficiencia de computadora para labores de rutina, pero cuya ciega tecnología es inoperante frente a un cambio de situación que implique nuevos factores. Es por esto también, que los ejemplos ilustrativos presentados y la exposición misma pueden parecer de una simplicidad pueril, porque de momento no nos interesan las aplicaciones concretas que requieran técnicas muy elaboradas y más o menos particulares; los problemas que enfrenta una ciencia en proceso de desarrollo y fundamentación, en la que todavía existen muchos puntos de divergencia e incluso opiniones contradictorias, son de naturaleza conceptual y genérica. Sin embargo, al terminar, tenemos que admitir que el material presentado constituye apenas la ordenación de un conjunto de notas marginales tomadas durante nuestra exploración del tema; las ideas centrales han sido aportadas por casi todos los autores citados en la bibliografía brevemente comentada que presentamos, y por varios más, y en el curso de su desarrollo nos hemos apoyado en las demostraciones de unos para aceptar las afirmaciones de otros, pero nos responsabilizamos plenamente por nuestras pequeñas y accesorias aportaciones personales, así como por sus posibles errores, y aunque todo el contenido es ciertamente muy elemental, esperamos que haya cubierto los temas más importantes

y proporcionado en forma suficientemente rigurosa y sistemática la verdadera naturaleza de la cibernética, la cual es posible que en el futuro constituya una actividad profesional independiente, pero de cualquier manera, siempre estará vinculada a la ingeniería, por su espíritu analítico y su precisión cuantitativa, y muy particularmente con la ingeniería electromecánica que es la que toma a su cargo, cuando esto es factible, la materialización en equipo real de sus concepciones. De este modo hemos llegado nada más que a los umbrales de esta ciencia, pero creemos que ya capacitados para comprender los teoremas de Shannon con sus en verdad sorprendentes demostraciones de que existe un límite inviolable para la capacidad informativa de una fuente y un canal dados, lo que evitará muchos esfuerzos infructuosos; la posibilidad de recibir información confiable a través de canales no confiables; etc., para continuar después la exploración de su extraordinaria vastedad temática.

2. La actividad más característica dentro de la ingeniería, en cualquiera de sus especialidades, es sin duda el diseño, respecto al cual se está generalmente de acuerdo en que constituye tanto un arte como una ciencia. Sin embargo, conviene reducir los aspectos "artísticos" al mínimo, ya que en este caso éstos son sinónimo de una especie de sutil intuición empírica producto de una larga experiencia personal en esta actividad, pero que deja mucho que desear en cuanto a eficiencia, y que desde luego no se puede transmitir formalmente a otra persona. Pues bien, creemos que en la conclusión de esta tesis de ingeniería, y siempre encuadrados en un criterio pragmático, bien cabe una pequeña sugerencia referente al impulso que podría dar la teoría de la información al diseño de ingeniería para hacerlo una actividad más científica y sistemática, susceptible de aprenderse en términos formales. En efecto, dado que el diseño es esencialmente un proceso de selección entre un campo de alternativas posibles, que procediendo por etapas va reduciendo las alternativas originales muy amplias, hasta definir un único sistema que mejor cumpla con el conjunto de condiciones que implica todo proyecto, y que incluye entre otros factores, leyes físicas, ventajas para su producción en masa, propiedades de los materiales empleados, consideraciones estéticas y aspectos económicos, queda situado dentro -

de los límites de la cibernética.

La discusión exhaustiva de este tema, tan complejo por la multitud de factores que involucra, queda más allá de nuestros objetivos y posibilidades actuales, pero con todo, creemos que un primer paso en la dirección indicada consistiría en la introducción de un criterio cuantitativo en el diseño; vale decir, de una "CANTIDAD DE DISEÑO" que proporcionara una indicación tanto de lo ya hecho y de lo que queda por hacer, respecto de un proyecto particular en marcha, como del terreno que queda por explotar en relación con una invención o descubrimiento básico. Claro está que este criterio no podría expresarse en términos absolutos, que originaría cálculos con cantidades astronómicas y difíciles de determinar, sino dentro del marco de una situación concreta con límites prácticos, por ejemplo, de una cierta empresa en una situación dada. Tratemos de explicarnos con un pequeño ejemplo. Supongamos que se trata del diseño de un amplificador electrónico que forma parte de un cierto equipo más amplio. Con este solo dato inicial, las posibilidades son muy extensas, pero cada una de las especificaciones de la lista de diseño de que hablábamos, constituye una etapa del gradual proceso selectivo que finalmente culminará en una unidad amplificadora bien definida. Por brevedad, y para destacar mejor la idea principal, supongamos que este diseño solamente involucra tres etapas, y que existen quince circuitos cuyas diferentes características presentan interés práctico. Ahora bien, en estas condiciones y sobre esta base, sería conveniente determinar numéricamente la magnitud de la labor selectiva que debe desarrollar el proyectista para efectuar este diseño. La teoría de la información nos dice inmediatamente que esta "cantidad de diseño" asciende a

$$\log_2 15 = 3.32 \log_{10} 15 = 3.32(1.176) = 3.9 \text{ bits.}$$

Si suponemos ahora, que en la primera etapa la selección de tal o cual juego de válvulas electrónicas proporcionó dos bits, y que en la segunda etapa la adopción de cierto conjunto de resistencias ha disminuido la variedad en un bit, fácilmente puede saber el diseñador, te-

niendo en cuenta el cálculo anterior y los principios de la cibernética, que la última etapa de su proyecto, si le ha de definir totalmente al circuito en cuestión, no deberá de proporcionarle menos de 0.9 bits.

Esta sugerencia, que por cierto sólo es original en su posible aplicación al diseño de ingeniería, puede ser de ventajas tan objetables como cualquiera otra innovación, pero de todas maneras, queda como terreno por explorar dentro de la siempre fascinante esfera de la actividad creadora.

3. Muchas etapas ha recorrido la humanidad desde que por primera vez alguien, en algún lugar de la Tierra, obligado por la hostilidad del ambiente y su situación relativamente indefensa, se detuvo a pensar, y objetivizando sus dificultades y los datos de su percepción, los transformó en pregunta, iniciando así su incursión en el mundo del pensamiento; desde entonces ya no dejó de hacerlo nunca, las ventajas que le confirió fueron abrumadoras desde los mismos albores de la inteligencia que le ha dado civilización y cultura, pero muy especialmente desde que tuvo autoconciencia de esta misma facultad de conocer. Auxiliares imprescindibles del hombre en su aventura intelectual han sido, también desde un principio, la función verbal y su peculiar habilidad manual. La primera le permitió literalmente pensar, pues al sustituir por sonidos articulados los objetos y situaciones de su ambiente, creó -o por lo menos empleó por vez primera- el concepto, así como el primer proceso de codificación y el primer canal para intercambiar ideas; la segunda, le permitió dejar la huella inconfundible de su presencia a través de las edades, y transformar en forma espectacular el aspecto y características de su ambiente de acuerdo con sus deseos y necesidades, y surgiendo así, de estas facultades específicas, la diferencia y también la amalgama entre pensar y hacer, entre ciencia y tecnología.

En cierto modo la cibernética, o cualquiera sea el nombre que se le quiera dar, no es sino una etapa más, la más moderna, del proceso descrito de conceptualización de la realidad, o si se quiere, de esta previa excursión por el mundo de las ideas y de las relaciones, antes de intervenir con eficacia amplificada en el mundo de las cosas; tras de ella, sólo se encuentra el método científico llevado conscientemente

hasta el extremo para explotar todas sus posibilidades, son los antiguos pilares del conocimiento, ya ampliamente difundidos y utilizados por los pensadores de la Grecia de hace veinticinco siglos: el análisis, la síntesis, la abstracción, la inducción y la deducción, aplicados a modernos problemas; es también continuación de la fecunda unión renacentista de experiencia y matemáticas, con sus poderosos métodos de análisis lógico, que consisten en remontarse regresivamente de consecuencia en consecuencia hasta establecer con evidencia la verdad o falsedad de una proposición dada. Se trata pues, nada menos, que de un recurso de que se vale el espíritu humano para satisfacer su inquietud intelectual, en la inacabable tarea en que se encontrará comprometido mientras pueda asombrarse de que el Universo sea como es, y mientras pueda hacerse preguntas y tenga el acicate de no poder contestarlas todas; en pocas palabras, es la más reciente manifestación del Homo sapiens en su aspecto más específico, en su perene y plenamente deliberada aceptación de todo aquello que, en el ámbito del Universo, constituya un reto a su entendimiento.

Finalmente, queremos expresar nuestro agradecimiento a todas aquellas personas e instituciones que, sabiéndolo o no, colaboraron en nuestra comprensión y verificación de los principios aquí contenidos; a todos los notables autores e investigadores, contemporáneos y del pasado, cuyas ideas se han citado, con nuestra admiración y respeto por su capacidad intelectual; y en particular, al Sr. Ing. Miguel Reyes Aguirre, por el apoyo prestado y las valiosas sugerencias hechas en las fases iniciales de este trabajo, y al Sr. Ing. Daniel Barrios Morales, bajo cuya amable dirección se realizó.

F I N.

J. M. Chanfón.
Otoño de 1969.

R E F E R E N C I A B I B L I O G R A F I C A.

1. Ashby W. R.: "An Introduction to Cybernetics".
New York: John Wiley, 1966.

Obra de introducción, que presupone poca preparación, con muchos ejercicios bien graduados. Además de su sentido pedagógico, es un libro recomendable por todos conceptos; aquí hemos adoptado en buena parte el esquema de pensamiento del autor.

2. Abramson N.: "Information Theory and Coding".
New York: McGraw-Hill, 1963.
(V. esp. Madrid: Paraninfo, 1966.)

Muy buena exposición del tema. El autor advierte sobre ciertas dificultades a los lectores no muy versados en matemáticas; aquí hemos tratado de desmenuzar esas dificultades. Obra de nivel medio.

3. Allendoerfer C. B. and Oakley C. O.: "Principles of Mathematics".
Tokio: McGraw-Hill, 2ª ed.

Particularmente recomendables son sus capítulos referentes a Lógica y Conjuntos, y a la Probabilidad. Los autores tienen muy presente que las matemáticas son algo más que el aprendizaje mecánico de una técnica para contar y medir. Obra de introducción a las matemáticas avanzadas.

4. Adler A.: "Über den Nervösen Charakter".
Munich: Bergman, 1912.
(V. esp. Bs. Aires: Paidós, 1959.)

Interesante teoría centrada alrededor de lo que podría llamarse la tercera Ley de Newton en sicología, y de los intentos infructuosos de regulación basados en ella.

5. Boole G.: "An Investigation of the Laws of Thought".
New York: Dover (primera impresión americana de la edición británica original publicada en 1854).

Primer tratado sistemático sobre lógica simbólica, extremadamente interesante; publicado hace más de un siglo, puede considerársele como la realización del sueño de Leibnitz de un lenguaje universal para la ciencia, exento de ambigüedades y de cuño algebraico. Aunque su com-

prensión sólo requiere nociones elementales de lógica y de álgebra, hay que leerlo con cuidado para captar plenamente su sentido.

6. Crosson F. J. and Sayre K. M.: "Philosophy and Cybernetics".
   Londres: University of Notre Dame Press.
   (Cat. Bibl. B. Franklin: 501 PHI).

Conjunto de disertaciones de alta calidad, escritas por varios especialistas.

7. Cossa P.: "La Cybernétique".
   París: Masson et Cie.
   (V. esp. Barcelona: Reverté, 1963).

Crítica a la cibernética; breve, dura, superficial. A nivel de divulgación.

8. Cheng D. K.: "Analysis of Linear Systems".
   Massachusetts: Addison-Wesley, 1961.

Texto de nivel medio sobre el tema, con una buena exposición del método de transformadas de Laplace para la solución de ecuaciones lineales diferenciales e integro-diferenciales.

9. Fenichel O.: "The Psychoanalytic Theory of Neurosis".
   New York: W. W. Norton.
   (V. esp. Bs. Aires: Paidós, 1966).

Interesante porque pueden descubrirse nociones decididamente cibernéticas en las líneas cardinales de sus premisas teóricas, en una área muy alejada de la ingeniería y de las ciencias exactas. Aunque especializado en su terreno, no presenta ninguna dificultad.

10. Fry T. C.: "Probability and Its Engineering Uses".
    (Cat. Bibl. B. Franklin: 519.1 F946P).

Volumen de nivel semiavanzado sobre el tema.

11. Hall H. S. and Knight S. R.: "Higher Algebra".
    Londres: Mac Millan.
    (V. esp. México: U.T.E.H.A., 1956).

Buen texto superior, con abundantes problemas y sus soluciones. Especialmente recomendable su capítulo sobre probabilidades.

12. Jacobowitz H.: "Electronic Computers Made Simple".
(V. esp. Gral. de ediciones. Cat. Bibl. B. Franklin: S 510.78 JAC).

Exposición elemental del tema.

13. Latil P. de: "La Pensée Artificielle. Introduction á la Cybernétique".
(V. esp. Bs. Aires: Losada, 1958).

Define muy claramente la naturaleza de los servomecanismos, y presenta, entre otros temas, una versión muy original de la probabilidad. Obra de popularización.

14. Liberman D.: "La Comunicación en Terapéutica Psicoanalítica".
Bs. Aires: Eudeba, 1966.

Laudable el esfuerzo, que además muestra hasta donde ha llegado la difusión del tema, pero creemos que una auténtica fusión de teorías, -- cualesquiera que estas sean, debe ir mucho más allá de una mera yuxta-- posición superficial de unas cuantas nociones. Tema especializado, con varias referencias interesantes a la teoría de la información.

15. Mendieta A. A.: "Tesis Profesionales".
México: Porrúa, 1967.

Indicaciones muy útiles acerca de todas las fases y aspectos de la realización de una tesis profesional.

16. Maltz M.: "Psycho-Cybernetics".
New Jersey: Prentice-Hall.
(V. esp. México: Herrero,1961).

Intento de aplicar algunas conclusiones de la cibernética al plano anímico. Obra a nivel de divulgación.

17. "Progress in Brain Research". Vol. 17.
Editado por N. Wiener y J. P. Schade.
(Cat. Bibl. B. Franklin: 501 WIE).

Diversos tópicos de alta especialización, y entre otras muchas cosas, una interpretación del mismo proceso de investigación científica - como un fenómeno retroalimentado, y una reseña histórica de la cibernética, dedicada a la memoria de su fundador.

18. Reyes A. M.: "Curso de Plantas Generadoras de Electricidad".
México: Ingeniería de México, A. C., 1964.

Apuntes bien hechos. El Cap. XVI da algunas nociones sobre estabilidad, retroalimentación y analogía.

19. Solodovnikov V. V.: "Introduction to the Statistical Dynamics of Automatic Control Systems".
New York: Dover, 1960 (traducción al inglés de la primera edición rusa).

Una de las primeras versiones rusas de la cibernética, de nivel semiavanzado, y de contenido no precisamente fácil.

20. Skilling H. H.: "Electromechanics".
New York: John Wiley, 1962.

Texto sobre la materia que, como todos los del autor, se caracteriza por su fino sentido pedagógico y su originalidad. Todo su contenido es recomendable.

21. Swadesh M.: "El Lenguaje y la Vida Humana".
México: Fondo de Cult. Ec., 1966.

Exposición amena y elemental del desarrollo del sistema de comunicación más antiguo y perfecto.

22. Shannon C. and Weaver W.: "The Mathematical Theory of Communication".
The University of Illinois Press, 1964.

Shannon, conocido por sus diversas aportaciones científicas, encierra en este pequeño libro y en escueto simbolismo matemático, varias de las ideas más trascendentes de este siglo.

23. Schniermann A. L.: "Bechterev's Reflexological School".
Clark University Press.
(V. esp. Bs. Aires: Paidós, 1963).

Breve exposición de una de tantas corrientes dentro de las llamadas escuelas objetivas de sicología, con cierta relación con lo presentado en los primeros capítulos de este trabajo. Tema especializado.

24. Sokolnikoff S. Ivan and Elizabeth:"Higher Mathematics for Engineers and Physicists".
New York: McGraw-Hill, 1941.

Un texto claro, que cubre los tópicos de matemáticas avanzadas que más frecuentemente se requieren en la ciencia aplicada.

25. Tolman R. C.: "The Principles of Statistical Mechanics".
Oxford Univers. y Press, 1938.

Amplio y clásico tratado sobre la materia. Sus primeros capítulos no ofrecen dificultades especiales.

26. Tsien H. S.: "Engineering Cybernetics".
McGraw-Hill (Cat. Bibl. B. Franklin: 531. T882e).

Buen texto sobre cibernética, en su aspecto más tradicionalmente ingenieril.

27. Von Neumann J. and Morgenstern O.: "Theory of Games and Economic Behavior".
New York: John Wiley, 1964.

El coautor principal, uno de los matemáticos de mayor influencia en la actualidad, y muy conocido también por su labor dentro de la física nuclear, de hecho inaugura, con este voluminoso tratado, toda una rama de las matemáticas; aunque no emplea recursos de matemáticas superiores, su lectura está lejos de ser fácil. Obra a la altura de la categoría intelectual de von Neumann.

28. Viaud G.: "L'Intelligence".
Presses Universitaires de France.
(V. esp. Bs. Aires: Paidós, 1959).

Origen y diferencia entre lo empírico y lo racional; entre el "Homo faber" y el "Homo sapiens". De fácil lectura, sin tecnicismos.

29. The Voice of America. Forum Lectures: "Philosophy of Science".
(Cat. Bibl. B. Franklin: 501 VOA).

Varios ensayos de mucha profundidad e interés.

30. Wiener N.: "Cybernetics or Control and Communication in the Animal and the Machine".
Cambridge, Mass.: The M.I.T. Press. 2ª ed., 1961.

Con la publicación inicial de esta obra, en 1948, nació oficialmente la cibernética. El Dr. Wiener, notable investigador pionero en muchos terrenos, y poseedor de varios grados académicos, acuñó el término,

sin embargo, parece que la referencia más antigua a esta palabra se encuentra en "Gorgias", uno de los diálogos de Platón en oposición a la sofística. Exceptuando los capítulos sobre series temporales, y retroalimentación, que son altamente técnicos, su contenido es accesible y hasta ameno.

31. Wiener N.: "The Human Use of Human Beings; Cybernetics and Society".
Boston: Houghton Mifflin, 1950.

Pequeño volumen de divulgación, bien escrito y muy interesante, en el que el autor muestra sus conocimientos enciclopédicos y sus puntos de vista sobre varios temas, en lenguaje llano.

32. Watson J. B.: "Behaviorism".
New York: W. W. Norton.
(V. esp. Bs. Aires: Paidós, 1961).

Exposición del núcleo de una teoría, que podría llegar muy lejos, si la objetividad de las observaciones se apoyara en el rigor de relaciones cuantitativas. Con alguna relación con lo aquí expuesto.

33. Williams J. D.: "The Compleat Strategyst".
New York: McGraw-Hill, 1966.

Libro de introducción, de contenido muy accesible, y con abundantes ejemplos ilustrativos que destilan buen humor.

Además se consultaron ocasionalmente, para comparar puntos de vista, así como resultados y conclusiones, las siguientes obras:

1. "Feedback Theory and its Applications".
P. H. Hammond.
2. "Fundamental Theory of Servomechanisms".
Leroy A. MacColl.
3. "Introduction to Automatic Control Systems".
Robert N. Clark.
4. "Industrial Electronic Control".
Cockrell.
5. "Simbolic Logic and Intelligent Machines".
Berkeley.

6. "Alcances de la Automación".
Woodbury.
7. "The Thinking Machine".
Pfeiffer.
8. "Los Principios de la Ciencia".
Eduardo Nicol.
9. "Fisiología del Sistema Nervioso".
John Fulton.
10. "Computadoras Digitales".
Paul Siegel.
11. "Digital Computer Principles".
Burroughs Corporation.
12. "An Introduction to Digital Computing".
Bruce W. Arden.
13. "Computer Basics". (Vol. 3).
Technical Education and Management, Inc.
14. "Sistema e Historia de las Doctrinas Filosóficas".
Francisco Larroyo.
15. "Electrical Engineering Circuits".
H. H. Skilling.
16. "Industrial Electronics and Control".
Royce G. Kloeffler.