

Universidad Nacional Autónoma de México

FACULTAD DE CIENCIAS



LA INTELIGENCIA ARTIFICIAL:

VALORACION Y POSIBILIDADES

TESIS PROFESIONAL

QUE PARA OBTENER EL TITULO DE:

F I S I C O

P R E S E N T A:

JOSE PEREZ - CARBALLO

MEXICO, D. F.

JULIO DEL 82



Universidad Nacional
Autónoma de México



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.

I N D I C E

INDICE.....	Índice-1
I. INTRODUCCION.....	I-1
II. SIMULACION COGNOSCITIVA.....	II-1
Traducción Automática.....	II-2
Juegos.....	II-7
Reconocimiento de Formas.....	II-13
Conclusión.....	II-16
III. PROCESAMIENTO SEMANTICO DE LA INFORMACION.....	III-1
Conclusión.....	III-5
IV. MICROMUNDOS.....	IV-1
Conclusión.....	IV-16
V. ROBOTICA.....	V-1
VI. REPRESENTACION DEL CONOCIMIENTO.....	VI-1
Cálculos de Predicados.....	VI-3
Programas Simples.....	VI-5
Lenguajes del Tipo Planner.....	VI-6
Producciones.....	VI-8
MYCIN.....	VI-9
Redes Semánticas.....	VI-14
Teoría de Marcos.....	VI-16
KRL.....	VI-21
VII. LA CRITICA DE DREYFUS.....	VII-1
Hipótesis Biológica.....	VII-5
Hipótesis Psicológica.....	VII-10
Hipótesis Epistemológica.....	VII-14
Hipótesis Metafísica.....	VII-20
CONCLUSION.....	Conclusión-1
APENDICE.....	A-1
BIBLIOGRAFIA.....	B-1

INTRODUCCION

Alguna vez, mi director de tesis me preguntó qué es lo que esperaba que un lector obtuviera de la lectura de mi tesis. Mi respuesta fue que yo esperaba que una gran variedad de lectores pudieran encontrar "algo" de interés en este trabajo. En general las tesis de licenciatura son completamente incomprensibles, no sólo para el llamado "hombre de la calle", sino para casi cualquiera que no sea el propio autor, su director de tesis y dos o tres técnicos más que están trabajando en el mismo campo. En general se trata de trabajos muy técnicos, especializados en campos muy esotéricos de la física o las matemáticas. Mi intención con este trabajo es, precisamente, la opuesta: quisiera que fuera comprensible, y que ofreciera material interesante, a cualquier persona. Afortunadamente, el tema es muy adecuado para ese propósito: "La Inteligencia Artificial".

Antes de entrar en materia conviene delimitar el ámbito de la discusión que sigue. Si mencionara todos los antecedentes relevantes de la Inteligencia Artificial tendría que empezar, por lo menos, desde los filósofos griegos y dedicar capítulos a la historia de las computadoras, sin olvidar uno para la época de la cibernética. Pero haría el trabajo de una longitud, variedad y complejidad difíciles de digerir, tanto para el lector, como para el autor. Por ello, decidí limitar la discusión, específicamente, a los esfuerzos dedicados a escribir programas "inteligentes" para computadoras. Esto me permite empezar el relato sin tener que remontarme en la historia más de 30 años y concentrar-

me en una clase muy específica de investigaciones.

¿Qué es la Inteligencia Artificial?. La respuesta más inmediata a esta pregunta es: la ciencia que estudia la posibilidad de construir me canismos inteligentes. Desgraciadamente, esta respuesta nos compromete a responder a otra pregunta: ¿Qué es la inteligencia?. Esta segunda - pregunta no puede responderse con la misma facilidad que la primera y, sin embargo, es necesario considerarla al inicio de un trabajo de la na turaleza del presente.

Carl Sagan considera que existe una especie de chovinismo instintivo en el ser humano que lo predispone en contra de las máquinas inteligentes. Abundando en este tema, menciona los marcapasos, pequeños aparatos capaces de percibir el latido del corazón y estimularlo al menor síntoma de fibrilación. El respecto dice:

No puedo imaginarme al portador de este aparato resentido con él por su inteligencia. Pienso que en un período corto habrá una aceptación similar para máquinas mucho más - inteligentes y complejas. (Sagan, ref. 1920)

Pero, ¿es razonable atribuir inteligencia a este dispositivo? El mismo criterio haría posible considerar inteligente a casi cualquier co sa. Con esto, quiero mostrar lo mal definidos que están estos concep - tos y el descuido con que los utilizan todo tipo de personas.

¿Sería posible definir algún criterio cuantitativo, cualitativo o ambos, que pudiera utilizarse en estos casos?

Turing propuso en 1950 lo que llamó el juego de la imitación:

Un interrogador se comunica por medio de un teletipo con otros dos

entes. Uno de ellos es una máquina y el otro es un ser humano. Es posible hacerles todo tipo de preguntas para tratar de decidir cuál es - cuál. No existen reglas ni restricciones para las respuestas que pueden dar. Si un análisis estadístico de varios intentos muestra que el interrogador no puede discernir correctamente entre las dos posibilidades se concluye que la máquina en cuestión es inteligente. Esta prueba, por supuesto, sólo es válida cuando se ha logrado diseñar un mecanismo con una inteligencia cualitativa y cuantitativamente humana. De manera que el criterio de Turing resulta inútil para analizar el trabajo realizado en Inteligencia Artificial hasta ahora, y lo más probable es que siga siéndolo por mucho tiempo todavía.

Entre los productos de la investigación en IA hay programas que juegan ajedrez, que demuestran teoremas, que simulan la habilidad de un experto humano en ciertos campos (tales como la manipulación algebraica o el diagnóstico de enfermedades infecciosas y la correspondiente receta de antibióticos) o que son capaces de responder a preguntas relativas a algún texto que se ha dado previamente al sistema. En todos estos casos, se trata de actividades que, de realizarlas una persona, nos parecerían manifestaciones de su inteligencia y, más aún, manifestaciones peculiarmente humanas.

Según P.H. Winston, los objetivos de esta ciencia son "...hacer la computadora más útil y entender los principios que hacen posible la inteligencia" (Winston, ref 2560). De hecho, es posible discernir esas dos grandes tendencias en la investigación de la IA: la Ciencia Cognoscitiva y la Ingeniería del Conocimiento. La primera trata de entender

los procesos inteligentes en hombres y máquinas y la segunda de resolver problemas prácticos usando programas de computadora muy refinados.

Actualmente, empieza a generalizarse el uso de robots en diversos tipos de industria, y algunas otras de las aplicaciones más complejas de la IA están llegando al público de muchas maneras distintas. Pero, por otro lado, el público tiene nociones muy adulteradas (cuando llega a tenerlas) de lo que es la IA y, en general, del papel de las computadoras en la vida moderna. E incluso existe una amarga polémica entre especialistas de diversas ramas de las ciencias y humanidades sobre los alcances, posibilidades o limitaciones de estas investigaciones. En resumen: el panorama es particularmente confuso.

En este país, además, se suma el desconocimiento a la confusión que rodea a estos temas. Se dice que la IA ha sido abandonada en todo el mundo, que los problemas que plantea no están al alcance de los medios de los que dispone la investigación en este país y que, en todo caso, son problemas cuyo planteamiento no se justifica en el contexto de las necesidades de México.

Por supuesto, las personas que sostienen estas opiniones, en general, tienen in mente la exploración espacial o, incluso, las complejissimas plantas robot donde el Departamento de Defensa norteamericano planea fabricar su maquinaria bélica en el futuro. Un futuro, por cierto, inevitable e inquietantemente cercano.

Mi intención es discutir los principales temas relacionados con la IA para tratar de esclarecer, en lo posible, la confusión existente, o al menos, dar al lector los elementos necesarios para que empiece a fun

damentar una opinión propia sobre estos asuntos.

Vale la pena hacer notar la diferencia que existe entre la Ciencia Cognoscitiva y la llamada Ingeniería del Conocimiento. La primera intenta resolver los problemas que plantea el estudio de la inteligencia y sus manifestaciones, tanto en hombres como en máquinas. La segunda se ocupa del diseño y desarrollo de sistemas computacionales no algorítmicos. Se podría decir que la Ingeniería del Conocimiento pone en práctica las ideas desarrolladas en el curso de la investigación en Ciencia Cognoscitiva. Sin embargo, esta última parece encontrar problemas cada vez más difíciles de resolver mientras que la rama ingenieril produce cada vez más aplicaciones al mundo real. Dreyfus ve la investigación en Inteligencia Artificial como:

... una solución ad hoc para un problema restringido, primero presentada con cautela, y, después, interpretada como el primer paso para llegar a métodos más generales (Dreyfus ref. 512).

El mismo Minsky escribe:

Métodos que funcionaron bien para problemas fáciles no pudieron generalizarse fácilmente a problemas más difíciles. Será necesario aplicar nuevas ideas para lograr un progreso continuo, ya que hay algunos problemas muy difíciles en nuestro futuro inmediato (Minsky, ref. 1024).

Al leer la literatura producida alrededor de la IA, es común encontrar predicciones sobre su futuro, más o menos inmediato, que proclaman la revolución más profunda de la historia humana como consecuencia de estas investigaciones. Similarmente, es fácil encontrar las críticas

más agrias y devastadoras, que consideran este esfuerzo inútil, ridículo y hasta peligroso.

¿Vale la pena siquiera ocuparse de estos asuntos o es tiempo perdido? ¿Estamos en el umbral de uno de los acontecimientos más trascendentes en la historia de la humanidad o se trata simplemente, de un juego inútil?

En todo caso, el hecho cierto es que en el nombre de la "Inteligencia Artificial" se han desarrollado algunas de las ideas más poderosas en computación; el tiempo compartido, por ejemplo, fué una idea de McCarthy producto de sus esfuerzos dentro del contexto de la IA. Estas ideas, cuando tienen éxito, se exportan a otros campos de la ciencia y la técnica y, finalmente, se olvida su origen. Pero, si alguien definiera la IA como una rama del conocimiento dedicada a la creación de un artefacto inteligente, podría decirse que sólo tiene una historia de fracasos, al menos hasta ahora. Pero estaría perdiendo de vista todos esos frutos secundarios, más o menos involuntarios, que es innegable que la IA ha dado a la ciencia y a la técnica. Ninguna actividad creativa humana es totalmente inútil. Es casi inevitable, en el desarrollo de la ciencia, que se encuentren ideas que valen la pena, aún si no se está buscando precisamente en esa dirección. La intención del presente trabajo es hacer un repaso muy somero de las ideas principales que se han manejado durante los primeros años de vida de lo que se ha dado en llamar Inteligencia Artificial, hablar un poco de sus posibilidades y, por supuesto, mencionar sus limitaciones.

Se ha abordado con mucho éxito, por medio de las técnicas de la IA, el tipo de problemas prácticos que pueden resolverse por medio de un mé

todo ad-hoc, ya que se desenvuelven en mundos muy restringidos donde la flexibilidad no es esencial. Ejemplos de sistemas de este tipo son los "sistemas expertos" como DENDRAL, MYCIN, o MACSYMA. Por otro lado, problemas que requieren una mayor flexibilidad, ya que se ocupan de mundos que no pueden restringirse sin trivializarse, no han podido resolverse - ni parece que sea posible en un futuro demasiado cercano. Los ejemplos obvios de esta clase son los problemas de visión o los de comprensión - del lenguaje natural.

Otro campo en el que la IA ha tenido algunos éxitos, si bien muy - limitados, es el de la robótica. Quizás porque un robot representa la culminación del esfuerzo en IA, se pierde de vista que las aplicaciones de esta ciencia abarcan un ámbito mucho más general que el de los robots. Aún si los robots fueran el único resultado posible del trabajo en este campo, valdría la pena, más aún, sería de la más alta importancia, que científicos mexicanos no fueran ajenos a estas investigaciones. Aunque no interese a México, por el momento, mandar un artefacto explorador a Marte o parezca que el esfuerzo de los países desarrollados por construir robots industriales puede, o hasta debe, ser ajeno a un país donde sobra la mano de obra barata, ningún país que pretenda desarrollarse en un mundo tecnológico puede ignorar una tecnología que nadie duda tendrá, en un futuro muy cercano, un impacto transcendental en la estructura del mundo moderno. Esto sin importar lo ajena a nuestras metas que pueda parecer esta tecnología en el momento presente.

En todo caso, ya existe en nuestro país una necesidad objetiva de resolver problemas a los que ya se aplican con éxito técnicas de la In-

teligencia Artificial. Un buen ejemplo, muy bien conocido en nuestro mundo académico, es la necesidad de investigadores mexicanos, en ramas muy diversas, de hacer "talacha" simbólica, es decir, de usar la computadora para realizar cálculos algebraicos (en oposición a los cálculos numéricos que se hacen tradicionalmente). Ahora, es común que un investigador en física o ingeniería, por ejemplo, tenga que invertir mucho esfuerzo y tiempo en realizar una gran cantidad de manipulaciones algebraicas esenciales para su trabajo que, sin embargo, pueden realizarse automáticamente. Es como si, existiendo la posibilidad de usar calculadoras electrónicas, un investigador tuviera que hacer una pausa en su trabajo para dedicar días a hacer sumas y restas con lápiz y papel. Más adelante, mencionaré MACSYMA y otros sistemas de manipulación algebraica que en este momento, no en un futuro imaginario, son una necesidad importante en México. Alrededor de la IA existe una polémica de carácter internacional a la que tampoco podemos ser ajenos. No es posible sentarse a diseñar programas o circuitos, concentrándose sólo en que funcionen, sin jamás dedicar un momento de reflexión a las ideas que hay detrás de este trabajo.

Uno de los primeros problemas que se presentan a alguien interesado en la literatura producida bajo el signo de la IA durante los últimos 25 años, que es, además, una de las fuentes de la mencionada polémica, es el entusiasmo que parecía dominar los primeros trabajos y las predicciones extraordinarias que se hacían sobre el futuro inmediato de estas investigaciones. Sin embargo, cuando uno busca las siguientes publicaciones de los mismos autores que, pocos años antes, habían hecho predicciones apasionadas, es fácil percibir un cambio de tono. Con de

masiada frecuencia se encuentran descripciones de sistemas que constituyan, según el autor, "un primer paso" hacia el diseño de máquinas inteligentes. Este tipo de máquinas parecía ser el objetivo único y final de la IA y nadie siquiera pensaba en colocar entre comillas la palabra inteligencia. El objetivo era igualar la inteligencia humana y, más aún, superarla. ¿Qué límite podría haber para la inteligencia artificial que la ciencia y la técnica todopoderosas podían otorgar a una máquina? Una vez construida la primera máquina inteligente, sería inevitable que su propia inteligencia fuera superada continuamente por nuevos modelos. Después de todo, contaríamos con su ayuda para construir la siguiente y con la de ambas para seguir adelante. Esta explosión intelectual terminaría en máquinas todavía superiores. Finalmente se esperaba que estas máquinas superinteligentes, que estarían más allá de nuestra comprensión, pudieran desviar su atención, de vez en cuando, de su elevada esfera de pensamiento para ocuparse de nosotros.

Estas no eran predicciones hechas por autores de ciencia ficción, sino que podían encontrarse en la mayor parte de los artículos escritos en los primeros años por los más serios científicos de la IA. Por ello, es más sorprendente el hecho de que no se cumpliera ninguna de tales predicciones. Sobre los programas que se proclamaban como un "primer paso" hacia el diseño de máquinas inteligentes Dreyfus comenta:

...Feigenbaum y Feldman informan que se están haciendo progresos tangibles, y definen "pro-

greso" muy cuidadosamente como "desplazamiento en la dirección de la meta final". De acuerdo con esta definición, el primer hombre en subir a un árbol podría decir que ha hecho un progreso tangible para acercarse a la Luna (Dreyfus, ref 512).

En este trabajo, quisiera mencionar las ideas más importantes, así como los más famosos "primeros pasos", las distintas técnicas que se usaron para resolver los problemas que se iban planteando y algunos de los más importantes puntos de vista.

SIMULACION COGNOSCITIVA

Uno de los primeros esfuerzos que se ha dado por clasificar más tarde dentro de la IA es el de dotar de capacidades intelectuales humanas a una máquina, tales como la capacidad de resolver problemas con gran flexibilidad. Alrededor de 1957, Simon y Newell, empezaron sus intentos por diseñar un programa capaz de resolver cualquier problema y no sólo cierto problema que el programador tiene en mente de antemano. Al analizar la manera de resolver un problema de lógica de un estudiante, notaron que, a veces, utilizaba procedimientos muy informales que no siempre conducían al resultado correcto pero que en muchos casos ayudaban a la solución de un problema aparentemente muy difícil. (Un ejemplo de este tipo de procedimientos es tratar sistemáticamente de sustituir una expresión corta por una larga).

Estas reglas, en general producto de la intuición o de una experiencia no verbalizada, fueron bautizadas como "heurísticas". Inmediatamente se empezó a utilizar este tipo de enfoque en el diseño de programas, en contraposición a los que utilizan algoritmos.

Los programas heurísticos no siempre obtienen el resultado óptimo pero suelen lograr uno bueno en casos en que la contrapartida algorítmica hubiera fracasado totalmente. Los programas algorítmicos garantizan una solución pero, en la mayor parte de los casos interesantes, es tal la complejidad del problema que la solución se obtiene en un tiempo tan largo que deja de ser práctico el uso del programa. La razón es que los algoritmos funcionan exhaustivamente, es decir, analizando todas las posibilidades una por una. En la mayor parte -

de los problemas el número de posibilidades, para todo efecto práctico, es infinito. Por otro lado, los programas heurísticos utilizan reglas para eliminar posibilidades o para escoger caminos que es posible que sean mejores aunque no siempre así sucede.

La técnica que usaron Simon y Newell para descubrir estas reglas heurísticas fué el análisis de protocolos de seres humanos. Se pedía a los estudiantes que explicaran de qué manera resuelven un problema y luego se analizaba el procedimiento para encontrar los atajos que inconscientemente habían seguido. Simon y Newell no sólo pretendían elaborar una teoría de la inteligencia basada en estas cuestiones sino que consideraban inmediata la posibilidad de convertir la teoría en una máquina inteligente. El razonamiento era que si es posible formalizar las reglas heurísticas, es decir, si es posible formalizar la intuición, más aún, convertir la intuición, la creatividad y el ingenio en un conjunto de reglas formales, nada podría impedir convertirlas en un programa de computadora que exhibiera estas facultades. El programa que resultó de estas investigaciones fué pomposamente bautizado "General Problem Solver". Otros de los primeros problemas que se abordaron con estas técnicas fueron la traducción automática y el reconocimiento de formas. Cada uno representa un campo de investigación clásico y merecen un comentario por separado.

TRADUCCION AUTOMATICA

No es difícil imaginar la fascinación que, en aquella época, de

be haber ejercido en políticos y científicos la posibilidad de disponer de una máquina capaz de traducir textos automáticamente (del ruso al inglés, por ejemplo). Cualquier documento podría traducirse en unos minutos. Hacia muy poco, durante la segunda guerra mundial, Turing había colaborado en el diseño de una máquina capaz de descifrar los códigos alemanes. Aún ahora partes de aquel proyecto, desarrollado en Inglaterra, siguen siendo material clasificado. Se ha comentado que no se ganó la guerra gracias a haber podido descifrar aquellos códigos pero que muy probablemente se hubiera perdido en otro caso.

No parecería haber una diferencia muy importante entre descifrar un código alemán o traducir un documento del ruso. Así que los años 50's se caracterizaron por la gran cantidad de investigaciones en traducción automática. Según Deyfrus éstas

... tuvieron el éxito más inmediato, abarcan las investigaciones más extensas y más caras y constituyen el fracaso más inequívoco. (Dreyfus, ref 512).

En efecto, una de las aplicaciones más inmediatas fué la creación de un diccionario automático (Anthony Oettinger, 1954). En esa época se resolvieron muchos problemas con relativa facilidad y se pensó que las únicas limitaciones eran técnicas. Se pensó que, con una memoria suficientemente grande, que permitiera almacenar más palabras y una velocidad mucho mayor para poder hacer rápidamente los análisis necesarios para decidir entre casos ambiguos, la traducción automática estaría resuelta. Sin embargo, uno de los productos de -

la investigación en traducción automática es que ahora sabemos que - también existen limitaciones conceptuales que, aún ahora, con memo - rias y velocidades en muchos órdenes de magnitud superiores a las de aquella época, impiden que sea fácil, en caso de que siquiera sea po - sible, resolver el problema.

Bertrand Raphael comenta que los lingüistas que trabajaban en - los problemas de traducción automática, esperaban que, si sus progra - mas eran capaces de captar suficientes diferencias sintácticas entre dos idiomas distintos, las oraciones traducidas mantendrían su signi - ficado sin alteración. Sin embargo, los resultados de estos experi - mentos "fracasaron miserablemente" según palabras del propio Raphael. El mismo cita el siguiente ejemplo famoso: La frase bíblica "the - spirit is willing but the flesh is weak" (el espíritu está dispuesto pero el cuerpo es débil), fué traducida del inglés al ruso y de nue - vo al inglés con el siguiente resultado: "the wine is agreeable but the meat has spoiled" (el vino es agradable pero la carne está podri - da).

Gradualmente, los investigadores que trabajaban en traducción - automática empezaron a admitir que había problemas que no podían re - solver todavía y que estos parecían ser especialmente difíciles. Yehoshua Bar-Hillel dice que el hecho de que, al principio se resol - vieran algunos problemas más o menos fácilmente hizo que se creara la ilusión de que el problema de la traducción automática podría re - solverse a corto plazo.

No parecía en ese momento que la distancia entre tales resultados (los primeros problemas resuel -

tos) y la verdadera traducción de alta calidad fuera enorme todavía, y que los resultados resueltos hasta entonces, a pesar de ser muchos, fueran sólo los más simples mientras que los "pocos" que aún quedaban eran los más difíciles, en realidad muy difíciles (Bar-Hillel, ref 128, citado por Dreyfus, ref 512).

Aún ahora, veinte años después de que fué publicado el párrafo anterior, los problemas esenciales de la traducción automática no han sido resueltos. ¿Qué es lo que hace estos problemas tan complicados? ¿Bastaría un mayor desarrollo de la tecnología de las computadoras para resolverlos? Después de terminar este trabajo he llegado a la conclusión de que hay un problema conceptual que va más allá de los problemas meramente técnicos.

Mucha gente, incluidos algunos de los investigadores más conocidos del campo de la IA, piensan que lo que hace falta es entender los principios básicos que permiten a un ser humano comprender el lenguaje natural, antes de intentar resolver el problema de la traducción automática.

Para traducir un texto es necesario primero "entenderlo" en el idioma original para poder después verterlo en el segundo idioma. El significado de este "entender" es el problema básico de la traducción automática, y, de hecho de la IA. Oettinger menciona el

muy misterioso proceso semántico que permite a

la mayor parte de la gente razonable interpretar la mayor parte de las oraciones razonables inequívocamente, la mayor parte del tiempo (Citado por Dreyfus, ref 512)

De hecho, en un contexto dado, las ambigüedades concebidas - en abstracto no se presentan. La oración es "ofda" en la forma - adecuada porque el contexto organiza la percepción. A esta habilidad de reducir el espectro de significados posibles, ignorando lo que, fuera de contexto, sería ambiguo, la llama Dreyfus "tolerancia a la ambigüedad". Este autor piensa que este proceso, junto - con lo que llama "conciencia marginal", de la que hablaremos más tarde, es esencial para comprender el lenguaje.

Cuando se menciona la dificultad de programar una máquina para que reproduzca ciertas conductas humanas muy complejas, como es el caso de la comprensión, parece muy natural sugerir un programa capaz de aprender como lo haría un niño. Esto, empero, no es sino darle la vuelta al problema. ¿Cómo enseñaríamos a una máquina? Este problema es tan difícil como el anterior; es, de hecho, el mismo.

¿Cómo se podría enseñar a una máquina? Wittgenstein decía - que el niño debe compartir, al menos, algunas de las metas e intereses del maestro. Un niño comparte una serie de cosas con las - personas que lo rodean y de las que aprende el lenguaje. Algunas de éstas son los elementos básicos en los que están fundamentados sus sistemas de percepción, no todos los cuales son aprendidos.

Dichos elementos forman un conjunto de principios fundamentales a partir de los cuales se construyen todas las estructuras maduras - de la percepción, así como las que manipulan los datos obtenidos - por aquélla. Es muy probable que estos dos tipos de estructuras sean de la misma naturaleza. La búsqueda y definición de estos - principios debe ser un objetivo central, no sólo de la IA, sino de los estudiosos de los procesos cognoscitivos en general.

JUEGOS

Las reglas heurísticas se han aplicado con gran éxito en los programas de juegos. Un juego puede representarse por un conjunto de "estados" que forman un árbol. Un "estado" se define como una configuración de los elementos del juego a la que se puede llegar por medio de jugadas legales. En el caso del ajedrez sería un tablero con las piezas colocadas en la forma en que podrían encontrarse en algún momento durante el proceso de un juego normal. Para - dar una idea de la complejidad de este tipo de problemas, vale la pena recordar cómo se construiría el árbol para el ajedrez. El - primer estado sería el primer nodo del árbol. A partir de éste , las reglas del ajedrez permiten llegar a veinte estados distintos, según la jugada que escojan las blancas. Esto significa que el - primer nodo está unido a veinte segundos nodos. Para cada primera jugada, las negras pueden responder con veinte jugadas posibles. Es por ello que cada uno de los nodos del segundo nivel estaría unido a veinte nodos a un tercer nivel. En promedio, cada jugada - se puede escoger entre alrededor de treinta posibilidades y la long

gitud promedio de un juego entero es de unas cuarenta jugadas. Según esto, habría un mínimo de 10^{120} juegos posibles. Esto hace imposible el diseño de un programa que permita un juego perfecto a través del algoritmo que analiza todas las posibilidades exhaustivamente.

Para diseñar programas heurísticos de juegos se utilizaron, - como en el caso de la simulación cognoscitiva, protocolos humanos. Esta técnica dió buenos resultados en algunos juegos, como las damas; en el caso del ajedrez, se encontraron problemas muy graves. Hasta ahora

... no ha sido posible encontrar heurísticas a nivel de maestro en el ajedrez. Todas las heurísticas usadas actualmente excluyen jugadas que los maestros hubieran encontrado o dejan abierto el riesgo del crecimiento exponencial (Dreyfus, ref 512).

Simón piensa que estas heurísticas realmente existen. Otras personas, como Dreyfus, sugieren que no hay evidencia para apoyar esta afirmación.

Al analizar los protocolos humanos, se encuentra que, a menudo, empiezan considerando cierta situación del tablero - una torre no defendida - por ejemplo. A partir de este punto, el jugador humano empieza a considerar las consecuencias de una serie de jugadas motivadas por esta primera observación. Pero el problema empieza mucho antes: empieza desde el momento en que se buscan -

las zonas del tablero que pueden ser interesantes. Esto no se hace, en el caso humano, analizando todas las posibilidades hasta encontrar aquéllas que pueden resultar promisorias. Dreyfus propone la existencia de otros procesos cognoscitivos responsables de este tipo de fenómenos. Un proceso de este tipo haría posible que un experto concentrara su atención desde el principio en una zona interesante sin necesidad de examinar muchas posibilidades.

El programa de Greenblatt, "MacHack", en cierta parte "dura" del juego llegó a calcular 26,000 alternativas, mientras que un jugador humano jamás considera más de 100 o 200. La jugada del programa fué buena, pero lo significativo es la diferencia entre 200 y 26,000.

Este orden de diferencia sugiere que, al jugar ajedrez, los seres humanos están haciendo algo diferente al simple conteo de alternativas, y la pregunta interesante es: qué es lo que hacen que les permite, considerando solamente 100 o 200 alternativas, encontrar jugadas más brillantes que las que encuentra la computadora después de buscar entre 26,000 (Dreyfus, ref 512).

Simon piensa que el proceso que permite a un ser humano concentrarse en el punto adecuado del tablero sin considerar cada una de las jugadas posibles, es de la misma naturaleza que el proceso de contar alternativas (y que, por lo tanto, puede simularse por medio de reglas heurísticas) sólo que es inconsciente y mucho más rápido que el resto del proceso, es decir, el tipo de reflexiones

que aparecen en un protocolo. Pero, se pregunta uno, si este proceso existe, dado que es mucho más eficiente y mejor que el otro - el consciente - por qué no continuarlo hasta el final en lugar de recurrir a otro proceso no tan bueno, esto es, al análisis explícito de alternativas. Todo el proceso podría ser inconsciente y la solución aparecería como una revelación y mucho más rápidamente. El hecho de que existan dos partes sugiere, al contrario, que la primera es de una naturaleza esencialmente distinta. Esta primera parte además está ligada inextricablemente al proceso de percepción.

Dreyfus sugiere la existencia de otro proceso cognoscitivo, para explicar este tipo de fenómenos, al que llama "conciencia marginal". Este proceso, nos permite percibir inconscientemente ciertos estímulos. A veces se percibe la existencia de un ruido continuo sólo en el momento en que el ruido cesa, lo que prueba que, de alguna manera, estaba siendo percibido a algún nivel. En el caso del ajedrez

...claves de todo el tablero, a pesar de permanecer en los márgenes de la conciencia, dirigen nuestra atención a ciertos sectores haciéndolos parecer prometedores, peligrosos, o simplemente dignos de interés (Dreyfus, ref 512).

Aparentemente la experiencia de un jugador de ajedrez constituye un contexto que le hace percibir el tablero con un contenido especial. Esta experiencia incluye la memoria reciente del desa -

rollo del juego y la intención con la que se han hecho ciertas jugadas, así como el recuerdo de juegos anteriores jugados u obser-
vados. En general, los programas consideran cada jugada como un -
problema aislado. Un programa que intentara conservar alguna in -
formación sobre los estados anteriores del juego, ocuparía muy -
pronto un espacio demasiado grande de la memoria.

Lo que se necesita es un programa que conserve se-
lectivamente aquellos datos que son significativos
a la luz de la estrategia presente y la estrategia
atribuida al oponente. Sin embargo los programas,
actualmente, simplemente no contienen una estrate-
gia a largo plazo (Dreyfus, ref 512).

Hasta ahora, los programas existentes consisten únicamente en
búsquedas en el espacio del juego que utilizan reglas heurísticas
complejas para tratar de jugar en la forma más eficiente posible.
En general, existe un equilibrio entre la profundidad de la búsque
da y el criterio que utiliza la máquina para decidir la mejor jug
da. Algunos de los programas de más éxito, tratan de utilizar al
máximo la mayor velocidad y la gran capacidad de memoria de las má
quinas modernas. Estas características les permiten profundizar -
en el árbol del juego sin necesidad de que el criterio para esco -
ger la jugada sea muy afinado. En algunos casos este criterio es
simplemente la diferencia entre el número de piezas de ambos con -
tricantes.

Ningún programa toma en cuenta otras técnicas o conceptos que

no sean contar alternativas explícitamente y el diseño de reglas heurísticas para discriminar entre éstas y, sin embargo

no existe prueba, introspectiva o de comportamiento, de que contar es el único tipo de "procesamiento de información" presente en el juego del ajedrez... al contrario, todos los protocolos sugieren que el ajedrez supone dos tipos de comportamiento: 1) determinación de los puntos esenciales, por medio de la organización global del campo perceptual en un área al principio en los márgenes de la conciencia a la cual otras áreas, también en los márgenes de la conciencia, hacen interesante; y 2) contar alternativas explícitamente (Dreyfus, ref 512)

El área de los programas de juegos ha permitido el desarrollo de técnicas de programación aplicables a problemas de interés más general. Las técnicas de búsqueda, por ejemplo, que originalmente fueron desarrolladas para este tipo de programas, tienen aplicaciones en todas las áreas en las que aparecen árboles de alternativas. Este tipo de programas puede tomar decisiones evaluando las posibles consecuencias de cada alternativa. Esto implica su aplicación potencial en gran variedad de ramas.

Los campos en los que se pueden esperar resultados más interesantes son aquellos en los que el universo del programa es suficientemente restringido como para hacer posible la enumeración explícita de posibilidades. Esto, como veremos más adelante, no pa-

rece ser posible en muchos casos y, sin embargo, los programas ya existentes, en los que se han aplicado estas técnicas, son suficientemente interesantes como para hacer que los esfuerzos en este campo hayan valido la pena.

RECONOCIMIENTO DE FORMAS

Los primeros programas que reconocían formas se basaban en la búsqueda de algunas características topológicas predeterminadas, - que usaban para comparar con las definiciones de cada una de las formas en términos de estos rasgos. El problema era encontrar rasgos relevantes, es decir, aquellos que se mantienen invariantes bajo ciertas transformaciones: tamaño, orientación, etc. Estos programas tuvieron éxito en el reconocimiento de algunas formas, en particular en aquellos casos en que el reconocimiento dependía de un número pequeño de rasgos específicos. Sin embargo, ninguno de estos programas representa un verdadero avance en el reconocimiento de formas dado que

Cada uno es un pequeño triunfo ingenieril, una solución ad-hoc para un problema específico, sin aplicabilidad general (Dreyfus, ref 512).

Dado que las computadoras deben reconocer las formas en términos de una lista de rasgos específicos, uno de los primeros problemas técnicos que se presentan es el crecimiento exponencial. Esto sugiere que los seres humanos utilizan otro tipo de método para analizar la percepción. Dreyfus piensa que el reconocimiento de formas necesita de una combinación de "intuición", "conciencia mar

ginal" y "tolerancia a la ambigüedad", más allá de las posibilidades de las máquinas digitales.

Parece claro que la percepción humana no funciona por medio de una enumeración explícita de alternativas. Un ejemplo que ilustra este punto, es cierto tipo de afasia estudiado por Gelb y Goldstein. Algunas de las víctimas de esta enfermedad han perdido la capacidad de reconocimiento perceptual. Para estos pacientes todo reconocimiento se convierte en un problema de clasificación. El paciente tiene que usar listas y procedimientos de búsqueda igual que lo haría una computadora digital. Algunos de estos afásicos sólo pueden reconocer una figura, como un triángulo, listando sus características explícitamente, esto es, contando sus lados, recordando que un triángulo tiene tres lados y concluyendo, pues, que la figura en cuestión es un triángulo. Esta percepción conceptual consume mucho tiempo y es ineficiente, tanto que las víctimas de esta enfermedad son incapaces de desenvolverse en el mundo cotidiano.

Dreyfus cita este ejemplo para mostrar que deben existir -- otros procedimientos, además de los que usan las máquinas digitales, para resolver los problemas de la percepción. En general, estos procesos se realizan en el inconsciente, es decir, ciertas cosas "salen" de pronto al plano consciente sin haber sido el resultado de un proceso del que el sujeto tenga consciencia.

Dreyfus parte de la proposición de Wittgenstein sobre otra forma de procesamiento de la información, indispensable para la

percepción de formas, a la que llama "agrupamiento perspicaz", que es, a su vez, una combinación de "conciencia marginal", "intuición" y "determinación por contexto". Esta forma de proceso es necesaria para reconocer formas bajo las siguientes condiciones:

1) La forma puede estar torcida, incompleta, deformada o rodeada de ruido;

2) Los rasgos requeridos para el reconocimiento pueden ser tan finos y numerosos que, aún pudiéndose formalizar, una búsqueda a través del árbol formado por tales rasgos, pronto se volvería demasiado grande a medida que se añadieran nuevas formas;

3) Los rasgos pueden depender de contextos internos o externos, de manera que una especificación libre de contexto no es posible;

4) Puede que no haya rasgos comunes sino una complicada red de similitudes interrelacionadas, capaz de asimilar nuevas variaciones.

Cualquier sistema que pretenda igualar la eficiencia humana debe ser capaz de distinguir las características esenciales de las no esenciales en una configuración dada, usar sugerencias que están en los bordes de la conciencia (cualquiera que sea lo que esto significa en el caso de una máquina), tomar en cuenta el contexto y percibir lo individual como típico, es decir, situar lo individual con respecto a un paradigma. Este último punto permitiría reconocer configuraciones similares.

El hecho de que aún el reconocimiento de moderada complejidad

pueda requerir de todas estas formas de procesamiento de la información, es, probablemente, la causa de que el trabajo en reconocimiento de formas no haya progresado más allá del reconocimiento de patrones muy simples, nada comparable a los problemas de percepción que requiere el mundo real. Desgraciadamente, se acepta, en general, que es necesario un progreso importante en reconocimiento de formas para permitir progresos en otras áreas de la IA.

CONCLUSION

Los esfuerzos en la investigación en IA, en los primeros años estuvieron caracterizados por una esperanza desmedida en futuros progresos, motivada por la solución relativamente fácil de los primeros problemas. Sin embargo, en todos los campos que hemos tratado, hemos visto cómo, tarde o temprano, se encontraron dificultades que ya no se pudieron resolver con la misma facilidad. En el caso de los juegos, el crecimiento exponencial del árbol de alternativas requiere de una restricción de los caminos que pueden seguirse. En juegos complicados, como el ajedrez, los programas no pueden seleccionar los caminos más prometedores. En la resolución de problemas no se trata sólo de cómo dirigir una búsqueda selectiva entre alternativas explícitas, sino de cómo estructurar el problema, de manera que se pueda empezar el proceso de búsqueda. En traducción automática, no es claro ni siquiera cuáles son los elementos que deben manipularse, debido a las ambigüedades intrínsecas del lenguaje natural. En reconocimiento de formas, las tres dificultades están inextricablemente entrelaza-

das.

Se diría, pues, que estos primeros intentos de la Simulación Cognoscitiva tropezaron con barreras que no han podido superar. Y, sin embargo, la importancia de estos estudios radica, precisamente, en haber encontrado y empezado a definir estas barreras. Puede que no sean problemas tan fáciles de resolver como se pensaba al principio, pero eso sólo los hace más fascinantes.

PROCESAMIENTO SEMANTICO DE LA INFORMACION

Los primeros sistemas diseñados para manipular lenguaje natural, manejaban sólo la sintaxis, es decir, resolvían los problemas de ambigüedad que, necesariamente, implica el lenguaje, utilizando reglas referentes a la estructura gramatical de las oraciones. Al demostrarse que esto sólo conducía al fracaso, se intentó utilizar el contenido semántico. Minsky afirma que

Algunos proyectos no han progresado tanto como se esperaba, en particular los de traducción del lenguaje y demostración de teoremas matemáticos. Ambos casos, en mi opinión, representan intentos prematuros de manipular formalismos complejos sin representar, además, su significado (Minsky, ref 960, p 258)

Esta idea de representar el significado, que parece mucho más adecuada, resultó ser un problema tan difícil que no parece que pueda resolverse pronto. Los primeros esfuerzos en tal sentido fueron expuestos en el libro de Minsky "Semantic Information Processing" (Minsky, ref 1152). Los trabajos contenidos en él se caracterizan por ser soluciones ad-hoc a problemas cuidadosamente escogidos, a pesar de lo cual producen la ilusión de una actividad intelectual muy compleja.

El primer programa que describe Minsky es STUDENT, escrito por Bobrow. Este programa resuelve problemas de álgebra de secundaria, que le son dados en lenguaje natural. A pesar del resultado aparentemente impresionante hay dos cosas que hacen que el pro

blema no sea tan difícil como aparenta; éstas son: el hecho de - que se trabaja en un contexto muy reducido y que se conoce un tipo de estructura de datos muy adecuada para representar este tipo de información, esto es, las ecuaciones algebraicas.

El programa trabaja de manera muy simple. Rompe las oraciones del enunciado del problema en unidades, usando como base palabras clave como "times", "of", "equals", etc. Sustituye estas unidades por letras que representan variables y trata de establecer ecuaciones simultáneas. Si las ecuaciones no pueden resolverse, apela a otras reglas para romper las oraciones en otras unidades y vuelve a probar. Evidentemente el interés del programa no radica en la solución en sí de los problemas, lo cual se reduce a realizar unas cuantas operaciones elementales, sino a la comprensión del enunciado, que es también el mayor problema para una persona.

Existe una restricción, que hace que el programa funcione, - que no existe en textos ordinarios, ésta es, que los trozos de las oraciones, representados por variables, deben establecer ecuaciones solubles. Esto mismo es una limitación a la generalidad del programa. Bobrow mismo llama la atención sobre el siguiente ejemplo

... la frase "the number of times I went to the movies" que debería interpretarse como una sola variable, sería interpretada incorrectamente como el producto de dos variables: "number of" y "I went to the movies" dado que "times" siem-

pre se considera un operador.

En realidad STUDENT "entiende" un subconjunto muy pequeño del inglés y esto en un contexto muy restringido. Lo que hace el sistema es una mera manipulación sintáctica; de hecho, el contenido semántico está determinado de antemano por el programador. Sin embargo, Bobrow piensa que es posible generalizar el programa:

El sistema STUDENT es un primer paso hacia la comunicación, por medio del lenguaje natural, con las computadoras. Si se trabaja más sobre la teoría semántica propuesta, resultará posible la creación de sistemas mucho más complejos. (Citado por Dreyfus ref 512)

El resto de los trabajos en el libro de Minsky está constituido también por soluciones ad-hoc de un problema restringido que, sin embargo, se esperaba poder extender a situaciones más generales. Estas generalizaciones no han sido posibles hasta ahora.

Otro programa descrito en el libro es el que resuelve problemas de analogías, escrito por Evans. Este tipo de problemas aparece con mucha frecuencia en pruebas de inteligencia. Se dan ocho dibujos, tres marcados con las letras A, B y C respectivamente, y los restantes numerados del 1 al 5. El problema consiste en encontrar un dibujo X, de entre los cinco propuestos, tal, que A sea a B como C es a X. Es decir, se trata de encontrar la transformación que convierta a C en X y que sea lo más parecida posible a la que transforma A en B.

Este programa está compuesto por tres módulos. El primero construye una lista con la descripción de los dibujos A y B en términos de las relaciones entre las figuras que lo componen. Las posibles relaciones son: "arriba", "abajo", "adentro", etc. Este módulo también construye una descripción de la manera en que las figuras que componen el dibujo A son transformadas para obtener el dibujo B y cada uno de los dibujos numerados. Las posibilidades son que las figuras se hagan más pequeñas, más grandes, - sean rotadas, reflejadas o alguna combinación de estas transformaciones. También es posible que se añadan figuras o se omitan.

El segundo módulo compara la regla que transforma A en B con las que transforman C en cada uno de los dibujos numerados. Si no se encuentra una transformación exactamente igual, el tercer módulo trata de simplificar la transformación A-B para hacerla lo más parecida posible a la regla con la que se está comparando. Finalmente, se escoge la transformación que requirió una menor simplificación.

Este programa exige de una descripción formal de las figuras que analiza. El método es difícil de ampliar a propósitos más generales. La mente humana, por otro lado, es probable que funcione de una forma más global. Cuando se percibe una figura no se está construyendo, consciente o inconscientemente, una lista explícita de características.

Otro programa descrito por Minsky es el de Quillian. Este programa es un intento de estructurar una memoria semánticamente.

Procesamiento Semántico de la Información

Quillian comenta en su tesis (Quillian, ref 1856 p 54) que programas como el de Bobrow trabajaban a un nivel casi puramente sintáctico y expone la necesidad de extender el ámbito del lenguaje que un programa puede manejar. Para este fin, propone Quillian, es necesario incorporar un número creciente de hechos semánticos. El resultado es un programa heurístico que almacena y recupera el significado de palabras y "cualquier cosa que pueda ser expresada a través del lenguaje, captada por la percepción o de cualquier otra manera conocida y recordada" (Quillian, ref 1888 p 221). Esta información constituye una enorme red interconectada. La idea detrás de todo esto es que la comprensión del lenguaje natural implica la construcción de un todo estructurado a partir de un número enorme de partes explícitamente listadas. Quizás en este caso, igual que en el del ajedrez y en la resolución de analogías, los fenómenos gestálticos (es decir, globales) juegan un papel importante en el ser humano. Pero, si el sistema funcionara, no habría objeción a usar métodos distintos a los que usó la naturaleza. Sin embargo, la estructura crece demasiado rápidamente al añadir nuevas definiciones, lo cual la hace impropia como modelo de la memoria humana y poco práctica para otro tipo de aplicaciones.

CONCLUSION

Ninguno de los programas descritos por Minsky plantea ningún principio general de la habilidad de la mente humana para comportarse inteligentemente. En general, se trata de soluciones muy -

hábiles para problemas particulares, como en el caso Bobrow y Evans, o modelos demasiado simplificados como en el caso de Quillian.

Se evita el problema central de cómo estructurar y almacenar una cantidad enorme de datos, aunque cabe preguntarse si es realmente necesario almacenar "una cantidad enorme de datos". Sin duda que la cantidad necesaria para simular un comportamiento inteligente, aún uno muy simple, como el de un niño de 18 meses, es mucho mayor que la manejada hasta ahora por los programas existentes, ninguno de los cuales ha podido mostrar la suficiente flexibilidad o generalidad. Quizás se debería empezar a pensar en métodos totalmente distintos y, posiblemente, incluso en nuevas tecnologías. Hasta ahora, todos los esfuerzos han estado dirigidos a programar una máquina digital más o menos poderosa. Quizás se debería intentar otro tipo de máquinas, probablemente híbridas (parte digitales y parte analógicas) o heterárquicas (máquinas compuestas de varias partes que trabajan en paralelo para resolver un problema) o ambas cosas. La estructura de la memoria de una máquina digital, de hecho la naturaleza misma de los datos que almacena, es lo que hace que el número necesario de datos para simular o reproducir un comportamiento inteligente, aún uno muy simple, crezca demasiado. Almacenar otro tipo de datos y de una manera distinta podría ayudar a resolver el problema.

A pesar de todo, Minsky piensa que los programas descritos en su libro representan "pasos hacia formas de manipular el cono-

cimiento" y piensa que el hecho de que

todavía parecen tener ámbitos de aplicación muy restringidos no indica una falta de progreso hacia la generalidad. (Minsky, ref 1152 p 14)

Dreyfus piensa que los problemas que se presentan a los programas en IA se "evitan", en el caso de los seres humanos, en lugar de resolverse, es decir, son problemas que ni siquiera se presentan en ellos, dado que la mente humana no utiliza las técnicas discretas de proceso de la información que utilizan las máquinas y que son el origen de todos los problemas.

Por ello, no es de ninguna manera obvio que el progreso, por poco que sea, de Minsky hacia formas de "manipular el conocimiento" ... sea un progreso hacia una inteligencia artificial. (Dreyfus, ref 512 p 146)

Dreyfus explica el hecho de que Minsky y Bobrow piensen que estos programas constituyen un "primer paso", a pesar de la dificultad encontrada para generalizarlos, proponiendo la existencia de una serie de hipótesis:

Su optimismo general en el sentido de que alguna solución computable debe funcionar, sin embargo, puede pensarse como la consecuencia de una suposición metafísica fundamental concerniente a la naturaleza del lenguaje y del comportamiento humano inteligente, a saber, que todo comportamiento ordenado realizado por seres humanos puede, en principio, ser formalizado y (por lo tanto) procesado por computadoras digitales.

les. (Dreyfus, ref 512)

Esta y otras hipótesis, a las que Dreyfus atribuye el optimismo de la IA, son suficientemente interesantes como para merecer una discusión más extensa en un capítulo posterior.

MICRO-MUNDOS

La dificultad de representar el contenido semántico en ámbitos muy generales, como es el caso del lenguaje natural en el contexto del mundo humano, trajo como consecuencia la invención de los Micro-Mundos. El Micro-Mundo es un subconjunto del conocimiento humano que puede analizarse aisladamente. Según palabras de Winston, es un campo suficientemente rico como para permitir el estudio de los puntos principales pero, al mismo tiempo, suficientemente simple como para hacer posible el progreso. SHRDLU y SEE son los ejemplos que discutiremos de este tipo de enfoque.

SHRDLU es un sistema creado en 1971 por Terry Winograd como parte de su tesis de doctorado en MIT. Este sistema simula en una pantalla de televisión el brazo de un robot que es capaz de mover cualquiera de los bloques de distintas formas y colores de un conjunto representado en la pantalla. Este sistema hace posible que una persona mantenga una conversación con la computadora. SHRDLU responde preguntas, "entiende" declaraciones u órdenes, siempre dentro del contexto de este pequeño mundo de bloques móviles. SHRDLU es capaz de usar la sintaxis, la semántica y, en general, hechos que conoce sobre los bloques, para "entender" el lenguaje natural restringido a este ámbito. Reconoce formas coloquiales de discurso correctamente, además de frases formales completas ("gracias", "de nada", etc.). Interpreta correctamente -

los pronombres y decide a favor de la versión más significativa - de casi todas las construcciones ambiguas. Para hacer todo esto se vale de una estrecha interacción entre el análisis sintáctico, el análisis semántico, la habilidad para resolver problemas y la observación del universo de discurso, que en este caso, es el brazo simulado del robot y el conjunto de bloques. Otra importante innovación del sistema es el uso de procedimientos, en lugar de - estructuras estáticas de datos, para representar el conocimiento. El diccionario contiene un pequeño programa para cada palabra que denota una acción.

El diálogo que es capaz de sostener SHRDLU es sorprendente - mente natural. El siguiente ejemplo está tomado de "Five Lectures on Artificial Intelligence" (Winograd, ref 2416). (Decidí no traducir la conversación para no desvirtuar el ejemplo. Las frases que escribe el operador están en minúsculas y las respuestas de - SHRDLU en mayúsculas. El resto del texto, que está traducido, - son los comentarios que agregó Winograd)

Un diálogo con SHRDLU:

1. pick up a big red block

OK.

El sistema responde "OK" cuando lleva a cabo una orden. Para tomar el bloque rojo tuvo que quitar el - que estaba encima y para eso fué necesario encontrar le un lugar a este último.

2. grasp the pyramid.

I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN

Hay tres posibilidades en la escena y el programa sabe que las frases que empiezan con "the" se refieren a un objeto específico que el interlocutor tiene en mente.

3. find a block which is taller than the one you are holding and put it into the box.

El sistema tiene que encontrar cuál es el significado de "one" e "it", así como utilizar la deducción para decidir qué bloque está sosteniendo.

BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.

Había dos posibilidades para "it": la seleccionada y la que se refiere al bloque que está sosteniendo. El sistema tiene heurísticas para decidir entre ambas.

OK.

[...]

7. is at least one of them narrower than the one - which I told you to pick up?

la frase "at least one of them" [al menos uno - de ellos] debe analizarse y referirse a la pregunta anterior y "the one which..." [aquél que] también debe referirse a bloques. Los grupos nominales pueden tener referencias numéricas -

complicadas como "at least one", "more than three", "exactly seven", etc. Se utiliza la memoria de eventos pasados para deducir cuál bloque es "the one I told you to pick up".

YES THE RED CUBE.

[...]

9. can the table pick up blocks?

NO

Las reglas heurísticas incluyen el conocimiento de que el acto de "tomar" [picking up] debe realizarse por un objeto animado, y que la mesa no es un objeto animado.

[...]

13. the pyramid is mine.

I UNDERSTAND

Toda la información sobre formas, colores, tamaños y posiciones de los objetos está almacenada de antemano en la memoria del sistema. La posición se deja como una propiedad desconocida de los objetos a la cual puede hacerse referencia. El sistema almacena esta información en su base de datos.

14. I own all blocks which are not red, but I don't own anything which supports a pyramid.

I UNDERSTAND

El resultado del entendimiento de estas oraciones es la creación de nuevos procedimientos para deducir si poseo algo o no.

15. do I own the box?

NO

[...]

34. put the blue pyramid on the block in the box.

Esta oración es sintácticamente ambigua, pero se entiende sin ambigüedad dado que "poner la pirámide azul, que está sobre el bloque, dentro de la caja" no tiene sentido (dada la posición en ese momento de los bloques), mientras que "coloca la pirámide azul sobre el bloque que está dentro de la caja" sí lo tiene.

35. is there anything which is bigger than every pyramid but is as wide as the thing that supports it?

El sistema entiende cuantificadores tales como "every", "any", "some", etc., y estructuras tales como "bigger than", "as wide as", etc. En este caso "it" se refiere a un grupo nominal dentro de la misma oración pero dos niveles arriba en el árbol analizador.

YES THE BLUE BLOCK.

SHRDLU es capaz de "entender" el lenguaje natural, restringi

do al microcosmos de los bloques, lo suficientemente bien como para llevar a cabo la mayor parte de las órdenes que se le dan y entender las declaraciones que se hacen al respecto, así como las preguntas, no sólo sobre el estado actual de la escena, sino sobre la historia de las modificaciones que se han hecho a lo largo de la sesión.

Winograd intentó simular por medio de este sistema el manejo de las estructuras cognitivas básicas necesarias para entender un texto. A continuación se describen muy brevemente algunas de estas ideas básicas.

En primer lugar, uno de los problemas reales más complejos, a saber, la percepción del mundo y el almacenamiento de la información obtenida, en este caso simplemente no existe. El programa utiliza un modelo de su mundo (el mundo de los bloques) detallado que describe el estado actual de la escena junto con su conocimiento de procedimientos para cambiar ese estado y para hacer deducciones sobre él. Sin embargo este modelo no está en términos espaciales o analógicos sino que es una descripción simbólica, preconstruida por el programador, en el sistema

...que hace la abstracción de aquellos aspectos -
del mundo que son relevantes para las operaciones
que se usan para trabajar con él y discutirlo. -

(Dreyfus, ref 512)

La abstracción de los detalles que son relevantes es crucial para cualquier sistema inteligente. En este caso, está hecha de

antemano por el programador.

Además del conjunto de hechos y procedimientos que forman el conocimiento sobre el mundo físico, el sistema incorpora una serie de programas que realizan el análisis semántico para resolver problemas del tipo que ocurren al tratar de decidir a qué se refiere una cuestión dada, y, finalmente, una gramática que determina la estructura sintáctica. Una de las ideas más interesantes de SHRDLU es el intento de coordinar estas tres componentes distintas.

El lenguaje no puede reducirse a las áreas separadas de la sintaxis, semántica y pragmática, con la esperanza de que el entendimiento de cada una por separado llevará a un entendimiento del todo. La clave de la función del lenguaje como medio de comunicación es la manera en que estas áreas interaccionan (Winograd, ref 2400)

Este sistema no funciona haciendo un análisis sintáctico primero, uno semántico después y, finalmente, utilizando deducción para obtener una respuesta. Cuando un trozo de estructura sintáctica empieza a tomar forma, se utiliza un programa semántico para ver si puede tener sentido y la respuesta resultante puede dirigir el análisis sintáctico, a su vez. Para decidir si tiene sentido, la rutina semántica puede llamar a procesos deductivos para hacer preguntas sobre el mundo real (parte pragmática).

Winograd dice que ninguno de los tres componentes por separa

do puede servir para simular la comprensión humana. En efecto, - podemos entender oraciones que no están construidas correctamente y, similarmente, podemos entender oraciones formadas por algunas palabras cuyo significado desconocemos, mediante el uso del análisis sintáctico. El proceso humano parece ser una mezcla coordinada de los tres puntos de vista donde una gran variedad de información sintáctica y semántica puede ser relevante y donde se aprovecha lo que pueda ser más útil para entender una parte dada de una oración. Este sistema es interesante porque lleva a cabo los diferentes tipos de análisis al mismo tiempo. Sin embargo la estructura del control del sistema no es un buen modelo de la mente humana:

Queda mucho por hacer para entender la manera de diseñar programas de computadora en los cuales varios procesos paralelos trabajen en forma coordinada, sin estar bajo el control jerárquico primario de uno de ellos. Un modelo del lenguaje que sea capaz de permitir realizar el tipo de "heterarquía" que se encuentra en los seres vivos (como la coordinación entre distintos sistemas de un organismo) se acercará mucho más a una teoría psicológica válida. (Winograd, ref 2400)

En 1966 Adolfo Guzmán diseñó el sistema SEE para presentarlo como su tesis de doctorado en MIT. Este sistema analiza proyecciones, en dos dimensiones, de escenas muy complicadas, que repre

sentan grupos de cubos u otros cuerpos rectilíneos similares. El objetivo del programa es realizar un análisis de la escena que permita llegar a conclusiones respecto al número de objetos presentes, sus relaciones y propiedades. El sistema de Guzmán efectúa el análisis a partir de los datos obtenidos de los vértices. El primer paso es hacer una clasificación de los vértices según su forma: L, flecha, T, K, X, tenedor, pico hacia arriba, pico hacia abajo. Con estas ocho primitivas y una serie de reglas heurísticas para su uso, el programa puede hacer análisis muy acertados de muchas escenas, en muchos casos, con mayor éxito que un ser humano. Estas reglas están basadas en el hecho de que los distintos tipos de vértices sólo pueden combinarse, junto con las superficies que forman las figuras, de ciertas maneras definidas.

Más adelante, el programa de Guzmán fué generalizado por Waltz, quien aumentó tres primitivas más que le permitían decidir si una línea era parte de una sombra, una rajadura, un borde oscurecido por una sombra o una fisura interna. Este método resolvió un problema de apariencia muy difícil de manera sorprendentemente fácil. Sin embargo el problema resuelto es demasiado particular para tener un interés práctico y hasta ahora los intentos por generalizarlo han sido infructuosos. En todo caso, parece imposible generalizar un sistema basado en el análisis de los vértices de figuras rectilíneas para que funcione con cualquier tipo de figura. Este es un buen ejemplo en el que la aplicación de los Micro-Mundos sólo conduce a un callejón sin salida.

Parece más factible la generalización del tipo de principios aplicados en SHRDLU para la comprensión del lenguaje natural. Incluso parece posible la aplicación de este tipo de ideas a casos prácticos en los que el contexto esté suficientemente restringido y en el que sea posible crear una representación del universo de discurso totalmente simbólica (tal podría ser el caso de las reservas de una compañía aérea). Pero la comprensión del lenguaje en el mundo real implica más que el análisis sintáctico, semántico o la aplicación de procedimientos: implica un mundo humano.

Veamos los problemas que se presentan cuando se intenta extender este tipo de sistema al mundo real:

Las historias infantiles son un campo favorito porque se piensa que su simplicidad debe facilitar mucho su comprensión. Minsky y Papert analizan las posibilidades de una máquina para entender la conocida fábula del cuervo y la zorra. El cuervo tiene un pedazo de carne en el pico y se dispone a comerlo parado en la rama de un árbol. Aparece una zorra que le dice que ha oído hablar de su bella voz y que quisiera oírlo cantar. El cuervo, adulado abre el pico para cantar, dejando caer el pedazo de carne que la zorra devora inmediatamente. La pregunta clásica para probar el grado de comprensión de un niño es: ¿La zorra pensaba que el cuervo tenía una voz muy bella?

El problema aquí no está en escoger el significado de cada palabra de entre un conjunto finito de posibles significados; tam

poco ayuda el diseño de un sistema lógico muy poderoso, y no parece posible elaborar un modelo del mundo humano que pueda representarse simbólicamente y que contenga el conocimiento necesario.

En palabras de Minsky y Pappert:

El problema central es que nadie ha construido - los elementos de un conjunto cognoscitivo, sobre tales asuntos que sea adecuado para entender el - relato (Minsky, ref 1088 p 42-44).

Minsky y Pappert se preguntan qué es lo que se debe saber para entender esta fábula. Proponen que sería interesante contar con el concepto de adulación. Para esto sugieren una microteoría de la ADULACION que describen como

...una colección, susceptible a ser ampliada, de hechos o procedimientos que describen las condiciones en las que podríamos encontrar adulación, qué formas toma, cuáles son sus consecuencias, etc.

(Minsky, ref 1088)

Pero esta vez no se está restringiendo el ámbito a un mundo separable del total del universo humano. En este caso cada concepto está inscrito en la esencia misma del mundo humano y no es posible, como antes, hablar de micromundos que se pueden estudiar separadamente. Esto se desprende del mismo artículo de Minsky y Pappert:

... sería muy difícil describir la adulación a nuestro entendedor, si éste no sabe ya que se pueden hacer afirmaciones con propósitos distintos al de transmitir información literal sobre hechos.

Esto obliga a los autores a proponer la inclinación del concepto de OBJETIVO o INTENCION. Obviamente este nuevo concepto implica la necesidad de otros y estos a su vez de otros más. Según Minsky y Pappert la lista es finita,

simplemente es muy grande, y se necesita un gran conjunto de conceptos para organizarla. Después de un tiempo, encontraremos que se hace cada vez más difícil agregar nuevos conceptos, y los nuevos parecerán menos indispensables. (Minsky, ref. 1088)

Dreyfus piensa que es totalmente injustificada la creencia de que la referencia, aparentemente interminable, a otras características humanas convergerá, de manera que puedan ser estudiados aisladamente micro-mundos simples.

La mayor parte de los problemas a los que se enfrentan las máquinas que tratan de duplicar funciones intelectuales humanas, como en el caso que acabo de describir, son una consecuencia de la naturaleza misma de las computadoras digitales, que obliga a estructurar el conocimiento de manera explícita y formal. Estos modelos se basan en la descripción simbólica para simular el comportamiento inteligente.

En general, se ha tratado de extender los métodos formales discretos que han tenido éxito en campos como la lógica formal a otros campos, como el de la percepción, por ejemplo, en donde su aplicabilidad es menos obvia. El hecho de que los resultados de

estos intentos no hayan sido tan buenos como se esperaba debería hacer interesante la consideración de enfoques radicalmente distintos, o, al menos, debería provocar la reflexión sobre la validez conceptual de esos primeros intentos.

Uno de los problemas más comunes que se presentan al tratar de estudiar los fenómenos cognoscitivos es su carácter global. En este sentido resulta muy sugerente considerar las ideas propuestas por la teoría de la gestalt. Las primeras aplicaciones de esta teoría fueron en el campo de la percepción. Se propone, por ejemplo, que la visión es un fenómeno global, es decir que el análisis de una escena no se hace parte por parte para construir con los fragmentos de información, así obtenidos, un todo, sino que la escena se percibe toda entera al mismo tiempo. La digitalización nos ha obligado a romper el mundo real en pedazos que tenemos luego que integrar, sin embargo, es posible que el cerebro implique un funcionamiento global que no necesita de la fragmentación del mundo real y su consiguiente y problemática reconstrucción.

Erich Goldmeier propone una alternativa al análisis de la percepción en términos de características formales (Goldmeier, ref 704 p 128). Goldmeier habla de distinciones, entre fondo y figura, asunto y forma, aspectos esenciales y accidentales, normas y distorsiones, etc. , que él piensa que no pueden ser explicados en términos de características formales conocidas. Goldmeier propone la existencia de "zonas de resonancia" en el cere-

bro para explicar estos fenómenos neurológicos. La estructura de las máquinas digitales no permite manejar el conocimiento de manera que sea posible simular estas zonas de resonancia de manera práctica (desde un punto de vista de tiempo real, por ejemplo), - al menos no con las técnicas conocidas.

John Haugeland, (736) en "The Plausibility of Cognitive Psychology", hace ver las similitudes que existen entre los hologramas ópticos y la memoria humana. Un holograma es una estructura que guarda memoria sobre algunas características de un frente de ondas, utilizando los mismos medios químicos que utiliza una fotografía tradicional. Cuando se ilumina con luz blanca una fotografía común y corriente, el frente de ondas reflejado reproduce las amplitudes que tenía el frente de ondas original que se usó para imprimir la fotografía. De la misma manera, cuando se ilumina un holograma con luz coherente (de una sola fase) el frente de ondas resultante coincide con el original pero, esta vez, no sólo en amplitudes sino en fases también. Esta es la razón que hace que el frente de ondas reflejado por el holograma sea una reproducción exacta del frente de ondas original. En otras palabras, no hay diferencia entre ver la luz reflejada por el holograma y ver la luz que originalmente lo imprimió.

Algunas de las características interesantes de los hologramas son las siguientes:

a) Toda la imagen puede reconstruirse a partir de cualquier pedazo suficientemente grande del holograma; cuanto menor sea el

pedazo menor será la resolución, pero toda la imagen estará presente. Se puede decir que toda la imagen está en cada uno de los puntos del holograma.

b) Similarmente, en el caso de un holograma de varios objetos, no es posible decir qué parte del holograma corresponde a qué objeto.

c) Si un holograma es iluminado por la luz reflejada por cierto objeto, aparecen manchas brillantes indicando la presencia y localización de todas las imágenes de ese objeto en el holograma, y manchas menos intensas que indican objetos "similares".

De manera que alguna forma neurofisiológica de codificación holográfica podría modelar varias características imprescindibles de la memoria y el reconocimiento visual, incluyendo su velocidad, algunas de sus invariantes, y el hecho de que grandes lesiones en áreas relevantes del cerebro sólo dañan levemente estas funciones (Dreyfus, ref 512)

CONCLUSIONES

La complejidad de los problemas planteados obligó a los investigadores a crear el concepto del Micro-Mundo, que les permitía considerar sólo parte del problema, aisladamente del resto del mundo. Esto permitió el diseño de algunos programas que eran capaces de funcionar, con relativo éxito, en campos muy restringidos.

Restringir el campo permite resolver versiones muy simplifi-

casas del problema con técnicas ad-hoc, sin embargo, la misma especificidad de las soluciones hace muy difícil, imposible en algunos casos, la generalización.

Los problemas resueltos son poco o nada significativos para el estudio de la mente humana, dado que carecen totalmente de generalidad; sin embargo, algunos de los programas representan técnicas que pueden usarse con éxito para resolver problemas prácticos por medio de las computadoras.

Las dificultades encontradas parecen ser consecuencia de la naturaleza de las máquinas digitales. Podrían resultar más interesantes otros métodos para resolver los problemas planteados por estos estudios. Los hologramas son un ejemplo de la posibilidad de modelar adecuadamente fenómenos cognoscitivos complejos por medio de técnicas no digitales.

ROBOTICA

Antes de empezar a hablar de robots conviene hacer algunos comentarios sobre el significado de la palabra "robot". Definirla es casi tan difícil como tratar de definir la inteligencia; de hecho, los problemas planteados por la definición de "robot" se heredan de la problemática de la definición de la inteligencia.

Algunas de las posibilidades de definición son las siguientes:

a) Un robot es un dispositivo mecánico que se comporta, al menos en ciertos aspectos, de la misma manera, o similar, que un ser humano.

b) Un robot es un dispositivo mecánico capaz de interactuar con el medio siguiendo un conjunto de órdenes; donde la interacción, a su vez, es modificada por los estímulos del medio.

La primera definición se reduce a una comparación con el ser humano, de la misma manera que nuestra definición de inteligencia. La segunda es un poco más específica aunque aún no lo suficientemente buena como para excluir ciertas máquinas automáticas que normalmente no consideramos robots. Una tornamesa automática, por ejemplo, puede ejecutar una serie de movimientos fijos o incluso programables, puede modificar su comportamiento de acuerdo a cambios en las condiciones del medio y, sin embargo, no la consideraríamos un robot. Ciertos modelos ni siquiera carecen de cierta flexibilidad (para actuar de la manera adecuada según el

tamaño del disco, por ejemplo) pero después de todo no tienen la flexibilidad necesaria para poder efectuar un trabajo distinto del de tocar discos. Estaría más dispuesto a aplicar el término "robot" si el dispositivo fuera capaz de realizar varias funciones muy distintas sin ningún cambio de su configuración esencial. Una máquina, o una línea de ensamblaje, capaz de armar cualquier, o casi cualquier aparato sin que para hacer el cambio sea necesario hacer cambios en su configuración física (es decir sólo a través de cambios de software), se acercaría más a la idea que yo tengo de lo que es un robot.

Bertrand Raphael (ref 1904) menciona los siguientes ejemplos cuando habla de robots: los robots industriales "Unimate" utilizados en la planta armadora de General Motors en Lordstown, Ohio; los carros que transportan pertrechos en el Fairfax County Hospital y SHAKEY, el robot desarrollado en el Stanford Research Institute - 1969.

Los robots Unimate (de Unimation Inc.) son capaces de hacer las soldaduras necesarias de cualquier modelo de automóvil. Cada vez que se va a modificar el modelo que va a ser armado en la línea, un operador humano dirige las acciones del brazo de los robots; más tarde, el robot es capaz de repetir todas las sucesiones de movimientos por sí mismo. Estos robots tienen dispositivos sensores que registran la presencia del vehículo colocado en la posición adecuada y realizan la serie de movimientos memorizados.

El producto más reciente de Unimation es PUMA (Programable - Universal Machine for Assembly), un pequeño robot que consiste únicamente en un brazo mecánico controlado por microprocesadores. PUMA es parte de la vanguardia de una multitud de robots pequeños y baratos que empiezan a invadir algunos campos de la industria norteamericana. Algunas de las ventajas de los robots del tipo PUMA sobre los métodos tradicionales de ensamblaje son las siguientes:

a) Flexibilidad. Al cambiar el modelo del objeto que se está armando, o el objeto mismo, era necesario, usualmente, descartar las viejas herramientas y fabricar otras nuevas. Puede enseñarse a PUMA a ensamblar una gran variedad de objetos mediante cambios de software solamente.

b) Los robots pueden trabajar en condiciones peligrosas o incómodas para un ser humano, sin descanso durante largos periodos y a cualquier hora del día, todos los días de la semana.

c) Es más barato mantener a un robot que a un obrero.

d) Y, por supuesto, "el sueño dorado del capitalista: ¡no es sindicalizable!"

Entre los más importantes clientes para el naciente mercado de los robots se encuentran los fabricantes de automóviles. Estos, utilizan los robots, hasta ahora, principalmente para operaciones de soldadura y, casi tan frecuentemente, para la aplicación de pintura. En el futuro cercano, será posible utilizarlos en muchos más aspectos de la fabricación de los automóviles.

El reciente desarrollo del PUMA ha hecho que General Electric estudie la posibilidad de utilizarlo para el ensamblaje de partes pequeñas, como las usadas en aparatos electrónicos. En un futuro no muy lejano, los robots del tipo PUMA serán un lugar común en la industria.

La Fuerza Aérea de los Estados Unidos está patrocinando investigaciones que podrían tener como consecuencia la creación de una planta totalmente automatizada que produciría aviones militares.

Entre los problemas que quedan por resolver está el de la visión. Los sistemas de visión artificial no tienen problemas para reconocer una pieza expuesta aisladamente en una superficie plana. Pero cuando las partes están revueltas y superpuestas el problema se vuelve muy complicado.

Otro ejemplo que menciona Raphael en su libro (ref 1904) es el de los carritos del Fairfax County Hospital, vehículos que se mueven, por corredores y elevadores especiales a través del hospital. La misma computadora que controla los elevadores, dirige un carro, por ejemplo, desde el almacén hasta cierto cuarto en el hospital donde se necesita algo. La computadora selecciona la ruta a seguir de manera que no haya colisiones ni embotellamientos.

Finalmente, SHAKEY, es el dispositivo más parecido a un robot, tal como se concibe este término popularmente. SHAKEY es un vehículo provisto de sensores, cuyos movimientos son controlados por una computadora demasiado grande para poder colocarse en el

vehículo mismo. La computadora recibe una serie de datos provenientes de la cámara de televisión y de los otros sensores, para utilizarlos para corregir los planes que SHAKEY sigue para resolver un problema. Estos problemas consisten, en general, en moverse a través de cuartos y pasillos del laboratorio, buscar ciertos objetos y, quizás, recolocarlos siguiendo ciertas instrucciones.

La finalidad del proyecto SHAKEY fue diseñar un sistema capaz de interactuar con el medio que lo rodea. Este medio era percibido por una variedad de sensores, desde una cámara de televisión, hasta dispositivos muy similares a los bigotes de un gato. SHAKEY era capaz de ejecutar órdenes muy simples, del tipo: "PUSH THE BOX OFF THE PLATAFORM" o "BLOCK DOOR 5 WITH BOX 2 FROM ROOM 3".

De los proyectos que hemos descrito, probablemente el menos ambicioso científicamente es el de los robots industriales y es también el que, sin duda, ha alcanzado un éxito más evidente. Los proyectos del tipo SHAKEY son, probablemente, más ambiciosos porque tratan de incorporar una gran variedad de ramas de la IA en un sólo sistema. SHAKEY debía tener la capacidad de comprender el lenguaje natural, en particular la visión. Todo esto debía funcionar en el mundo real, no sólo en una pantalla de televisión como SHRDLU (que, por cierto, se diseñó más tarde), que trabaja con un mundo totalmente simbólico. Aquí, el sistema mismo tenía que estructurar la representación del mundo real.

En el momento en que se intentó el desarrollo de SHAKEY, las ramas de la IA involucradas no se encontraban en un estado de de-

sarrollo suficientemente avanzado. Entre los problemas que plantearon estos primeros intentos se encuentra el de la representación del mundo físico, así como de la posición de cada una de las partes del propio cuerpo del robot en relación al mundo en cada momento. Otro de los grandes problemas es el de la percepción, que incluye los sistemas de visión, así como otro tipo de sensores a través de los cuales se recibe información, que debe usarse no sólo para hacer planes sino para modificarlos mientras se están llevando a cabo.

La magnitud de los problemas encontrados obligó a abandonar casi totalmente la investigación en robótica alrededor de 1973. Según Raphael, ello fue consecuencia de la decisión del gobierno norteamericano de suprimir los fondos destinados a estos proyectos. Algo similar ocurrió en otras partes del mundo en que se desarrollaban proyectos de investigación en el mismo campo. Pero como hemos visto antes, los robots, lejos de desaparecer, han pasado a constituir ahora, al principio de la década de los 80's, una industria de billones de dólares. Se predice, actualmente, que estamos en el umbral de una invasión de pequeños robots del tipo PUMA que harán todo tipo de trabajo, desde trasquilar ovejas en Australia (proyecto que, de hecho, está en marcha), ensamblar vehículos y otro tipo de maquinaria además de intervenir en los procesos de pintura y soldadura, hasta el ensamblaje de cassettes.

Los robots industriales serán los que tendrán un impacto más importante y palpable en la sociedad y, sin embargo, no serán los

Únicos robots en el mundo o en sus cercanías. Durante los años - 80's se enviarán robots exploradores, cada vez más perfeccionados al espacio. En particular, se enviará a Marte un robot que se construye en los laboratorios JPL de California.

Como hemos mencionado antes, uno de los problemas más importantes que plantea, no sólo el estudio de robots, sino casi cualquier rama de la IA, es el de estructurar la enorme cantidad de conocimiento que parece necesaria para interaccionar sustancialmente con el mundo real, además de definir la naturaleza misma de este conocimiento. En efecto, en este caso no se trata simplemente de almacenar una serie de datos simbólicos del tipo de axiomas y teoremas formales, sino del conocimiento necesario para interactuar con el mundo real, que bien puede resultar ser de un tipo cualitativamente distinto.

REPRESENTACION DEL CONOCIMIENTO

Hemos visto que el problema de la representación del conocimiento es central, tanto en el ámbito de la Ingeniería del Conocimiento como en el de la Ciencia Cognoscitiva. La primera trata de estructurar el conocimiento necesario para resolver cierto problema específico de la manera más eficiente posible (solución ad-hoc). La segunda, pretende representar el conocimiento necesario para obtener una conducta inteligente. Este segundo objetivo no sólo presenta la dificultad de la estructuración, igual que el primero, sino que, en este caso, no está clara siquiera la naturaleza misma del conocimiento en cuestión. El problema de definir la cantidad y la calidad de conocimiento necesario para reproducir una conducta inteligente es, esencialmente, un problema conceptual cuya solución no parece fácil.

Antes se ha mencionado que el problema de almacenar una gran cantidad de datos, de manera que puedan ser utilizados de una forma eficiente por un sistema inteligente, es un problema difícil.

Dreyfus lo menciona como

... el problema de cómo reestructurar y recuperar datos en situaciones en donde cualquier cosa puede ser relevante. (ref 512 p 33)

Es decir, no sólo es un problema estructurar los datos, sino decidir cuáles son los que son relevantes para un problema dado.

Recordemos que esta decisión debe tomarla el propio sistema y que -
"no se vale" que el programador la tome de antemano.

En opinión de Winograd, el problema de la representación de co
nocimiento es el principal que enfrentan los investigadores de la -
IA.

La IA podría ser caracterizada como la parte de la
ciencia de la computación que estudia este problema
de representación (ref' 2416 p 7)

En STUDENT, por ejemplo, la idea de la representación era muy
simple: significado = ecuaciones. Pero esta idea sólo puede funcio
nar, como he mencionado, en el caso de mundos formales muy restrin
gidos. En el caso de un sistema inteligente donde los datos se in
terrelacionan de una manera muy compleja, es necesario un método de
representación más general.

Según Winograd (ref 2496), cuando se diseña un sistema para re
presentar conocimiento en una computadora, existen tres puntos prin
cipales que son los importantes:

Primero, interesa que la manera en que el sistema hace uso de
la información sea eficiente. Es importante la forma en que la efi
ciencia cambia al aumentar la cantidad de conocimiento.

El segundo punto se refiere a la adición de conocimiento nuevo
a la estructura ya existente. Winograd propone que el conocimiento
sea modular, de forma que no sea necesario preocuparse en detalle -
de la manera en que está interconectada toda la estructura con los
datos nuevos.

Finalmente, se debe considerar el problema esencial de la complejidad de la estructura en relación con su generalidad. Es necesario escoger entre un sistema muy general pero demasiado complejo o uno demasiado particular pero lo suficientemente simple como para ponerse en práctica.

Al construir un sistema, se presenta el problema de tener que equilibrar todos estos puntos de la manera adecuada al caso concreto.

En las siguientes secciones se describen algunos de los sistemas de representación más importantes.

CALCULO DE PREDICADOS

Un ejemplo muy a la mano de representación, es el usado en matemáticas y lógica simbólica: el cálculo de predicados. En este tipo de formalismo se define una serie de operaciones. Las unarias, por ejemplo, permiten establecer a qué categoría pertenece cierto elemento. De esta manera, se puede construir una expresión atómica del tipo PERRO(fido). Una operación binaria permite establecer la relación que existe entre dos elementos: AMO(Juan, fido). También es posible usar cuantificadores para representar situaciones más complejas como "los perros son animales" o "cualquier perro tiene dueño":

$$\forall x \text{ PERRO}(x) \Rightarrow \text{ANIMAL}(x) \quad (\text{Teorema I})$$

$$\forall x \text{ PERRO}(x) \Rightarrow \exists y \text{ AMO}(x, y) \quad (\text{Teorema II})$$

Una serie de manipulaciones simbólicas permitiría almacenar -

cierta cantidad de conocimiento de esta manera y usarlo para responder preguntas o resolver problemas. La pregunta "¿fido tiene amo?" podría contestarse considerándola un teorema que debe demostrarse a partir de los teoremas almacenados en el sistema. Cuando el problema se plantea como la demostración de un teorema como el anterior, cuya representación sería:

$\exists y \text{ AMO}(\text{fido}, y)$

el sistema debe decidir que el teorema II es un hecho relevante, debe asociar la "x" del teorema con fido, usar algunas reglas de la lógica, y concluir que existe "y" tal que se cumple la propiedad $\text{AMO}(\text{fido}, y)$.

En la mayor parte de los sistemas de demostración de teoremas que utilizan cálculo de predicados se recurre al "procedimiento de prueba uniforme". Este procedimiento es tal que, si hay suficiente conocimiento en el sistema para probar algo, se logrará hacerlo tarde o temprano. Pero si el sistema no es muy pequeño esto ocurrirá más bien tarde. El problema es que el sistema probará muchos teoremas que no son relevantes antes de encontrar el bueno (o buenos). Los teoremas que se prueban se escogen de acuerdo a las leyes de la lógica. Hay heurísticas que hacen el procedimiento un poco más rápido, pero no hay nada en el cálculo de predicados que

permita decidir lo que es relevante, de una manera eficiente.

Algunas de las ventajas de esta representación son su generalidad y modularidad. El principal problema es la falta de eficiencia.

PROGRAMAS SIMPLES

Otro tipo de representación tan común como el cálculo de predicados y que presenta características que resulta interesante comparar, lo constituyen los "programas simples". Este es el título bajo el que Winograd incluye los programas algorítmicos sin ningún elemento heurístico.

La primera diferencia entre el cálculo de predicados y la programación es la de que, en el segundo caso, el conocimiento está separado en dos partes distintas: programa y datos, en oposición a la representación más homogénea que ofrece el primero.

En programación, el control es totalmente explícito. El programador ha determinado de antemano el conocimiento que se debe aplicar en cada momento (esto se logra por medio de subrutinas). En el caso del cálculo de predicados, la decisión para escoger el conocimiento que se va a utilizar es de una naturaleza mucho más general.

La eficiencia es mucho mayor en el caso de la programación. No se pierde tiempo probando distintas posibilidades o tomando decisiones heurísticas sutiles. Sin embargo, los programas simples y el cálculo de predicados son polos opuestos. Los programas son eficientes pero poco generales, mientras que las representaciones

del tipo de cálculo de predicados son mucho más generales pero mucho menos eficientes.

LENGUAJES DEL TIPO DE PLANNER

La idea básica de este tipo de lenguajes es la de conservar la eficiencia de la programación contando al mismo tiempo con representaciones tan generales y poderosas como el cálculo de predicados.

Los sistemas que utilizan este tipo de representación, generalmente, cuentan con una base de datos con afirmaciones simples del tipo "A está sobre B". Estos hechos se representan como (SOBRE A B). Además, hay una serie de teoremas consecuentes y teoremas antecedentes que constituyen un tipo de conocimiento más complejo. Obsérvese el siguiente ejemplo, adaptado del artículo de Winograd (ref 2496).

TEOREMA I (CONSECUENTE (X Y Z) (SOBRE ?X ?Y)

 (META (SOBRE ?X ?Z))

 (META (SOBRE ?Z ?Y)))

TEOREMA II (CONSECUENTE (X Y) (SOBRE ?X ?Y)

 (META (SUPERFICIE-LIBRE ?Y))

 (META (TOMA ?X))

 (META (VE-A ?Y))

 (META (SUELTA ?X)))

TEOREMA III (ANTECEDENTE (X Y) (SOBRE ?X ?Y)

 (BORRA (SUPERFICIE-LIBRE ?Y)))

Estos son tres teoremas en un lenguaje del tipo de PLANNER.

El primero, codifica el hecho de que es posible establecer que un objeto X está encima de otro objeto Y, si se puede establecer que X está sobre Z que, a su vez, está sobre Y. Este teorema de PLANNER no sólo contiene la información, como podría contenerla el teorema escrito en cálculo de predicados, sino que enumera una serie de acciones que deben tomarse si se tiene como objetivo establecer que (SOBRE X Y).

El segundo teorema es muy parecido a un programa que llamara subrutinas. La información que codifica es la necesaria para poner X sobre Y: primero es necesario quitar lo que hubiera encima de Y, después tomar X, trasladarlo a una posición encima de Y y, finalmente, soltarlo.

En lugar de llamar a una subrutina por nombre, un teorema de PLANNER especifica una configuración del resultado al que se quiere llegar, en este caso (SOBRE A B). Los teoremas se guardan usando un índice especial que puede ayudar a decidir cuáles se ajustan a la configuración objetivo. Cuando se tiene un objetivo particular, el sistema automáticamente prueba los diferentes teoremas que están indexados como útiles para este caso. Los teoremas escritos arriba serían llamados por cualquier objetivo de tipo (SOBRE ? ?).

El último teorema sirve para modificar la estructura de conocimiento existente. Cuando el sistema añade información que indique que X está sobre Y, este teorema borrará la información presente que codifique el hecho de que no hay nada encima de Y.

Probablemente una de las aplicaciones más famosas de este tipo de lenguajes es el sistema SHRDLU escrito por Winograd.

PRODUCCIONES

El sistema de producciones fué diseñado por Newell y Simon. Consiste en la aplicación de un conjunto de reglas situación-acción. El conocimiento se representa por medio de un conjunto, linealmente ordenado, de este tipo de reglas, a las que se llama producciones. Estas operan en una colección de patrones que constituyen la llamada memoria a corto plazo. Una producción se parece mucho a un teorema antecedente en un lenguaje del tipo de PLANNER.

Si los patrones en la memoria a corto plazo corresponden con el que indexa a la producción, entonces se realizan las acciones especificadas por la producción.

A continuación, se cita el ejemplo, también basado en el mundo de los bloques, que utiliza Winograd para ilustrar las producciones.

MEMORIA A CORTO PLAZO:

(SOBRE A B) (SOBRE B C) (#META (SOBRE A C))
(LOCALIZACION C (100 200 100)) ...

PRODUCCIONES:

(SOBRE X Y) (SOBRE Y Z)
METE-MCP (SOBRE X Z)
(#META (SOBRE X Y)) (SOBRE X Y)
METE-MCP (META-LOGRADA (SOBRE X Y))

donde MCP significa memoria a corto plazo.

Los patrones de las producciones se comparan con toda la memoria a corto plazo, no sólo con una sola de las afirmaciones. De esta manera, una producción puede ser disparada por una colección de hechos en una forma que sería muy difícil para PLANNER.

La decisión de qué producción utilizar está basada en una comparación sintáctica entre los patrones de la memoria a corto plazo y de las producciones.

La acción de una producción es una secuencia explícita de operaciones en la memoria a corto plazo. En esto también hay diferencia con un lenguaje como PLANNER, ya que una producción no dispara directamente a otra, de la misma manera que un teorema puede invocar a otro, todo lo que puede hacer es dejar modificada la memoria a corto plazo, de manera tal que disparará otras producciones cuando se haga de nuevo la correspondencia de patrones.

Otro ejemplo muy conocido de este tipo de representaciones es el sistema MYCIN.

MYCIN

Este programa diagnostica enfermedades bacterianas y recomienda antibióticos adecuados. En su artículo de 1975, Davis, Buchanan y Shortliffe (ref 448) describen MYCIN como un proyecto interdisciplinario realizado en la Universidad de Stanford. La finalidad era crear un sistema de consulta para diagnóstico y tratamiento de enfermedades infecciosas, en particular infecciones bacterianas en la sangre. Este problema fué escogido, en parte, para respon-

der a una necesidad real: un estudio (Kagan 1973) (ref. 448) de -
mostró que una de cada cuatro personas en Estados Unidos había si -
do tratada con penicilina durante ese año y que en casi el 90% de
los casos el tratamiento había sido inadecuado. Esto hizo pensar
en la necesidad de un sistema consultivo para los médicos que te -
nia que ser confiable y fácil de usar

Este sistema tendría que hacer uso de una cantidad enorme de
datos técnicos que cambian constantemente. Esta fué una de las -
razones que hizo a los autores escoger la representación de cono -
cimiento por producciones.

Una tercera característica que se quiso dar al sistema es la
capacidad de mantener un diálogo en lenguaje natural con el usua -
rio. Esto permite dar una serie de explicaciones coherentes sobre
los resultados. Por ello, MYCIN utiliza un sistema de razonamien -
to simbólico en lugar de funcionar de una manera puramente esta -
dística. Este diálogo tenía que estar diseñado de tal manera que
hiciera fácil el uso del sistema por personas no acostumbradas a
manejar computadoras.

Muchas veces la identificación del organismo culpable de una
infección requiere un tiempo demasiado largo. Por ello, a veces
se hace necesario empezar un tratamiento sin saber exactamente -
qué bacteria es la responsable de la enfermedad. Cuando la grave -
dad del estado del paciente obliga al médico a actuar en estas -
condiciones, en general se recetan una o varias drogas de espec -
tro amplio para cubrir todas las posibilidades. Otra opción con -

siste en utilizar todos los datos disponibles para tratar de recetar una droga más específica y, por tanto, más eficaz. En general, la falta de tiempo y de experiencia suficientes para aprovechar los datos disponibles obliga a los médicos a optar por tratamientos de amplio espectro.

MYCIN establece un diálogo con el operador para tratar de obtener la información que necesita. Estos datos sobre el paciente los utiliza para decidir cuál o cuáles pueden ser las bacterias responsables de la infección y, de acuerdo con ello sugiere los antibióticos que pueden ser más efectivos para combatirla.

La idea no es la de reemplazar al médico sino darle una herramienta para que su trabajo sea más eficiente. En general, este tipo de sistemas necesitan, en algún momento de su operación, el juicio peculiarmente humano. En el caso de MYCIN, por ejemplo, se pide en cierto momento al operador que califique el estado del paciente en una escala del 0 al 4, según su gravedad. En este caso, el programa está pidiendo un tipo de información que no se podría obtener a través de análisis médicos objetivos pero que se logra recurriendo a la experiencia del operador y a su juicio. Probablemente, ninguno de los dos por separado, hombre o máquina, hubieran podido llegar a un resultado tan bueno como el que obtienen juntos.

MYCIN funciona utilizando una serie de reglas del tipo situación-acción, en las que la situación está definida en términos de

unos cuantos parámetros que determinan qué regla heurística específica es la relevante. Un ejemplo de este tipo de reglas es el siguiente:

PREMISA (AND (SAME CNTXT INFECT PRIMARY-BACTEREMIA)
(MEMBER CNTXT SITE STERILSITES)
(SAME CNTXT PORTAL GI)
ACCION (CONCLUDE CNTXT IDENT BACTEROIDES TALLY-7)
(Davis, ref 448)

Es decir:

si 1) La infección es una bacteremia primaria, y
2) el lugar de donde se tomó la muestra es uno de los clasificados como "sterilsites" y,
3) se sospecha que la vía de entrada al organismo es el tracto gastrointestinal,
entonces se puede concluir que existen circunstancias que sugieren (el programa es capaz de cuantificar en qué medida) de que la identidad del organismo es bacteroide.
(Davis, ref 448)

Al igual que la mayor parte de este tipo de programas, MYCIN fué escrito después de estudiar muchos protocolos donde expertos humanos se enfrentaban al tipo de problemas que se quiere plantear más adelante al sistema. Estos protocolos se utilizan para diseñar las reglas heurísticas necesarias.

El ejemplo del apéndice muestra un diálogo con MYCIN en el que el sistema trata de obtener la información pertinente acerca de los síntomas del paciente, su condición general, su historia -

clínica y los resultados de los análisis que se hayan realizado (ver apéndice 1)

El sistema posee alrededor de 200 producciones. MYCIN adopta una de varios cientos de hipótesis y hace las preguntas necesarias para determinar si se cumplen las premisas correspondientes. De esta manera, las preguntas siempre parecen seguir una línea definida. La puesta en práctica de MYCIN fue extremadamente complicada, entre otras cosas vale la pena recordar que en el campo en que se aplica, las conclusiones nunca son seguras, así que hubo - que diseñar una teoría de razonamiento plausible que utiliza un "factor de certeza" para tomar decisiones. Aparentemente, éste es un punto que presenta problemas que no se han resuelto, todavía, a satisfacción de los autores.

Evidentemente, el primer requerimiento para crear un sistema de la naturaleza de MYCIN, que trabaja conforme a estas reglas situación-acción, consiste en que sea posible establecer estas reglas. Como los autores mismos admiten, no en todos los casos esto es posible.

Parece que se requiere un campo que haya llegado a un cierto grado de formalización, que incluye quizás un conjunto de primitivas reconocidas y un mínimo entendimiento de procesos básicos (Davis, ref 448 p 17)

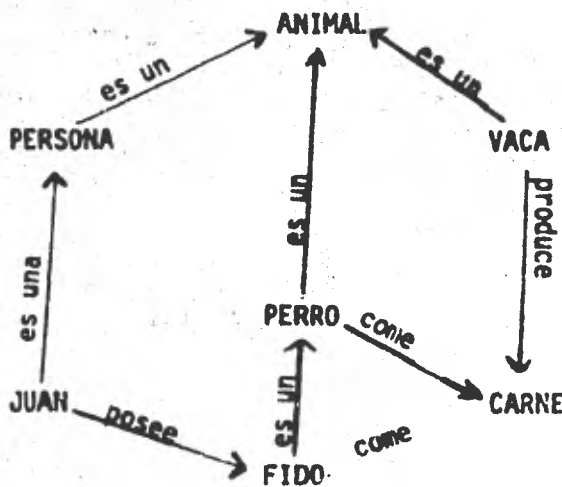
Otra limitación está impuesta por el hecho de que las primitivas no pueden estar interrelacionadas de una manera demasiado compleja. Aparentemente, el número de factores que interaccionan

en una premisa para disparar una acción tiene un límite práctico, si bien no parece que exista un límite teórico.

Finalmente, concluyen los autores, este tipo de sistema no parece el indicado para campos que exigen una fuerte interacción muy compleja entre objetivos ni para aquellos en los que es difícil diseñar reglas confiables.

REDES SEMANTICAS

Las redes semánticas son una representación muy usada en sistemas de comprensión del lenguaje natural. Una red se forma uniendo entre sí una serie de elementos simples. Cada unión es de un tipo distinto y contiene información sobre el tipo de relación que existe entre los elementos unidos. El siguiente es un ejemplo de red semántica:



Algunos tipos de información, como la jerarquía clase-subclase, se prestan especialmente bien para este tipo de representación.

Hay dos tipos de operaciones que son posibles en las redes. El primero es una deducción simple: la pregunta "Fido come carne?" podría contestarla un sistema que usara ciertos procedimientos para atravesar la red que le permitieran obtener la información: "Fido es un perro" y "perro come carne".

La operación es una especie de búsqueda. Se usa para deci - dir qué nexos son relevantes para lo que se está preguntando. Si se menciona, por ejemplo, "Fido" y "carne", sin especificar expli - citamente una relación, la red podría usarse para encontrar una. Se puede imaginar el proceso como una señal que se propaga de no - do en nodo, uno por unidad de tiempo. Una señal de este tipo se originaría en cada uno de los nodos en cuestión. Al intersectar - se las señales, es decir, al llegar ambas al mismo nodo, se habrá encontrado el conjunto más pequeño de nexos que conecta a los dos nodos iniciales. En el caso que nos ocupa, habríamos encontrado la relación "Fido come carne". Uno de los problemas de esta bús - queda es que crece exponencialmente al aumentar el número de no - dos. Si la búsqueda se extiende a más de uno o dos nexos se en - contrarán muchas conexiones y no todas serán relevantes.

Otro problema estriba en que esta estructura sólo permite al macenar un tipo de información muy simple. Las afirmaciones que requieren el uso de cuantificadores no podrían ser representadas, fácilmente, en una red.

El punto central de esta representación es la facilidad con que se encuentran objetos relacionados entre sí. Es decir, es - una solución al problema planteado más arriba: la recuperación -

de la información relevante.

TEORIA DE MARCOS

En 1974 (ref. 928) Marvin Minsky propuso la Teoría de Marcos de Referencia (Frame Theory). Esta teoría se concibió como una generalización para unificar los distintos puntos de vista de otras representaciones.

Según Minsky, enfrentarse a una nueva situación implica escoger una estructura de memoria que él llama marco (frame). Un marco contiene una serie de datos recordados sobre situaciones similares. Hay tres tipos distintos de información contenida en un marco. Parte de esa información es sobre la utilización misma del marco, otra es sobre lo que se puede esperar que pase en un futuro inmediato y, finalmente, qué puede hacerse si estas expectativas no se cumplen.

Esta estructura no es fija, puede adaptarse a la realidad cambiando los detalles necesarios. Un marco es similar a una red con nodos y relaciones. Los niveles superiores son fijos y representan las cosas que siempre son ciertas respecto a una situación dada. Los niveles inferiores tienen muchas terminales que deben llenarse con los datos pertinentes a la situación específica. La información contenida en cada terminal, en general, está en forma de sub-marcos.

Minsky propone la formación de sistemas de marcos que agrupen marcos relacionados entre sí. Los efectos de acciones importantes se representan como transformaciones entre los diferentes marcos del sistema.

En el caso de análisis de escenas visuales, los di
ferentes marcos de un sistema describen la escena
desde diferentes puntos de vista, las transforma -
ciones de un marco a otro representan los efectos
de moverse de un lado a otro. (Minsky, ref. 928)

Este párrafo requiere una explicación más detallada sobre la
manera de representar "diferentes puntos de vista" por medio de -
los marcos. Sin duda, se puede ver un objeto de tres dimensiones
desde diferentes ángulos y desde cada uno se verá muy distinto.
No es posible tener un marco para cada punto de vista. Quizás se
ría posible tener un marco para varios puntos de vista, de manera
que cada uno de los que faltaran fuera muy parecido a alguno de -
los existentes. Esto presenta no sólo problemas teóricos sino ob
vios problemas de realización.

Minsky propone que los diferentes marcos de un sistema tengan
terminales comunes, lo cual permite coordinar la información obte
nida desde distintos puntos de vista.

Una de las aportaciones más importantes de los marcos es que
consideran las expectativas o presunciones que se puede tener -
respecto a distintas situaciones. Las terminales de un marco con
tienen, en general, asignaciones por "default". Estas asignacio-
nes, sin embargo, son lo suficientemente flexibles como para ser
desplazadas por las características reales de la situación especí
fica.

Dreyfus (ref. 512) llama la atención sobre el antecedente de

la teoría de marcos: el noema del análisis fenomenológico de Husserl.

Para él (Husserl), el noema, o representación mental de cualquier tipo de objeto, aporta un contexto u - "horizonte interior" de expectativas... (ref 512)

El noema es una representación simbólica de todas las características que se pueden esperar al explorar un cierto objeto.

Eventualmente, el análisis de Husserl se enfrentó a graves problemas. Este filósofo pasó muchos años tratando de definir los componentes de los noema de objetos comunes y se encontró con la necesidad de incluir cada vez más información sobre el contexto general. Heidegger, más tarde, sugirió que el contexto de costumbres culturales era esencial para determinar las características relevantes y, por tanto, requisito previo para estructurar el "horizonte interior".

El problema más interesante al que se enfrenta la teoría de marcos propuesta por Minsky, es el de la representación del conocimiento cotidiano. Las dificultades de este problema son precisamente las mismas con las que se enfrentó Husserl y que, al final de su vida, le hicieron pensar que había tratado de hacer algo esencialmente imposible: aplicar los noema al mundo real.

Todo lo que implica el reconocimiento y comprensión de un objeto tan cotidiano como una silla parece no ser una cuestión simple. ¿Qué es lo que permite reconocer a una silla como tal? Sin duda, no es su forma, dado que existen una variedad prácticamente

infinita de diseños. ¿Podríamos encontrar una serie de primitivas, es decir, de características esenciales que permitieran construir el marco de una silla?

Minsky propone usar las características generales propias de todas las sillas para llenar los niveles superiores del marco correspondiente. Las terminales se llenarían con la información propia de un caso particular y permitirían diferenciar una silla de otra.

Según Dreyfus lo que convierte a un objeto en una silla es su función, y lo que le da sentido a esta función es su lugar en un contexto general. Este contexto está formado por el resto de los objetos que le dan sentido a una silla, así como de la cultura misma, que le da sentido, en general, a todos estos. Este contexto incluye, por supuesto, las características anatómicas y fisiológicas que hacen necesaria y posible a cada silla.

Dreyfus argumenta que no es posible encontrar características libres de contexto que permitan definir un objeto dado, una silla, por ejemplo. Según él, las partes de una silla, tales como patas, respaldo, etc. son reconocidas como tales una vez que el todo ha sido reconocido como una silla a través del contexto general.

El formalismo de los marcos ha sido tratado de aplicar en algunos sistemas; entre los más importantes están SCRIPTS de Schank y KRL de Winograd.

La idea de los SCRIPTS es tratar de definir un muy pequeño

número de primitivas, a partir de las cuales se puedan describir una serie de actividades estereotipadas. Esto debería funcionar para entender relatos sencillos.

Schank definió 11 primitivas, de las cuales, son ejemplos las dos siguientes: ATRANS - transferencia de relación abstracta (posesión, por ejemplo); PTRANS - movimiento físico de un objeto. Con estas primitivas creó una serie de escenarios estereotipados que le permiten al sistema llenar lagunas y entender pronombres - en historias simples.

Schank define un SCRIPT como

... una cadena causal predeterminada de conceptualizaciones que describen la secuencia normal en que ocurren las cosas en una situación familiar.

De esta manera, hay un SCRIPT para ir a un restaurant, uno para una fiesta de cumpleaños, uno para un juego de football, otro para un salón de clase, etc. (Schank, ref. 2240 p. 131)

La idea de Schank para entender el relato consiste en escoger el SCRIPT adecuado, lo que nos da una idea de cómo se debe desarrollar la actividad en cuestión, normalmente. El SCRIPT es un marco de referencia gracias al que sabemos qué esperar como desarrollo normal de la situación. Esto debería hacer posible que el sistema comprendiera una serie de detalles gracias a su conocimiento del contexto en que está inscrita la historia. Desgraciadamente, el mismo problema que se presenta teóricamente para la teoría

de marcos se hace patente en la práctica para los SCRIPTS de Schank.

En palabras de Dreyfus:

... no importa qué tan estereotipado pueda ser, ir a un restaurant no es un juego auto-contenido, sino un conjunto muy variable de comportamientos que están relacionados con el resto de la actividad humana. (ref. 512)

Esta es la razón por la que no existe una forma "normal" de ir a un restaurant. No es posible analizar el problema aislada - mente del resto de la experiencia humana, pues es inútil tratar - de segmentar el mundo de esta manera.

Por otro lado, no es posible determinar de antemano cuáles son los puntos relevantes de una historia, ya que estos dependen siempre de la historia misma. En cada caso, quien lee un relato, acepta ciertas convenciones del autor. Ello le permite encontrar relevantes los mismos detalles que, en otra historia no lo hubie - ran sido.

Otro problema que no parece fácil de resolver es el que plantea la manera de escoger el SCRIPT adecuado en cada caso. Decir cuál es el marco adecuado ya presupone una cierta comprensión de la situación.

KRL

Otra aplicación de los marcos de Minsky es KRL (knowledge - Representation Language). Este es un lenguaje desarrollado por - Winograd y colaboradores cuya finalidad es tratar de formalizar - toda la estructura de necesidades y conocimiento del mundo que ha

ce posible (hasta cierto punto) la comunicación entre dos - personas. Se espera que KRL pueda captar en estructuras simbólicas el conjunto de creencias (beliefs) que constituyen el "modelo del mundo" de una persona. KRL hace particular énfasis en el hecho de que los detalles relevantes están en función de su contexto.

KRL describe los conceptos según su relación con ciertas descripciones prototípicas. Algo parecido a "es como Juan pero más alto y pelirrojo". Para KRL la persona definida por la frase anterior estaría descrita por las mismas características que describen al prototipo Juan salvo por la altura y el color del pelo. Estas pueden ser características definidas por defecto para Juan o posibilidades que se dejan abiertas para adecuarse a cada caso particular.

Estos prototipos están descritos por medio de estructuras simbólicas y su número debe ser, no solamente finito, sino suficientemente pequeño como para que sea posible manejarlos. El problema de definir los prototipos adecuados es interesante y de gran importancia.

Otro problema, ligado con el anterior, es la dificultad de escoger el marco adecuado. Una computadora tendría que revisar un gran número de posibilidades antes de encontrar la buena. Es muy plausible que la mente utilice otro procedimiento más eficiente. Sin embargo, KRL representa un intento muy interesante de incorporar el contexto al almacenar el conocimiento.

LA CRITICA DE DREYFUS

El primer punto que se debe tocar al considerar la posibilidad de las máquinas inteligentes es un tanto espinoso y de naturaleza ideológica, puesto que exige aceptar la hipótesis de que la "mente" es consecuencia del comportamiento de una estructura material y desechar la de que, detrás de la inteligencia, hay algo intangible o misterioso, como el alma o el espíritu. Se trata, de acuerdo con la primera hipótesis, "simplicemente" de la consecuencia del comportamiento de una serie de estructuras formadas por proteínas; algo tan material y tangible como un conjunto de engranajes o de microcircuitos impresos, aunque quizás no tan fácil de analizar. Así sería posible que el mismo tipo de comportamiento pudiera reproducirse con otro tipo de estructuras y con materiales distintos de las proteínas, de la misma manera que se puedan realizar una serie de operaciones aritméticas lo mismo con una máquina electromecánica que con una moderna calculadora construida a base de microcircuitos y también es posible que en un futuro más o menos lejano podamos utilizar el mismo tipo de materiales y reconstruir las mismas estructuras (es decir, máquinas hechas de proteínas). Si la naturaleza lo hizo, quizás no haya razón para que tarde o temprano no podamos repetirlo.

Pero casi todos los esfuerzos hechos para tratar de reproducir la inteligencia humana se basa en un solo tipo de máquina: la computadora digital, con las limitaciones que ello implica. De estas limitaciones han surgido algunas de las críticas que se han

hecho a la Inteligencia Artificial. Una de las más interesantes es la de Dreyfus, por lo que se escogió para ser discutida, a continuación, en la parte final de esta Tesis.

Hubert L. Dreyfus escribió una serie de artículos sobre las limitaciones de la Inteligencia Artificial y su libro "What Computers Can't Do" (ref. 512) reúne las ideas fundamentales de su crítica.

Para empezar, Dreyfus insiste particularmente, en las afirmaciones de investigadores como Simon y Minsky, hechas en los primeros años de la I.A. De las muchas citas que aporta Dreyfus, vale la pena recordar una especialmente significativa. Se trata de un párrafo escrito por Simon en 1957 cuando, en palabras del propio Dreyfus, "el General Problem Solver parecía estar iniciando la Era de la Inteligencia Artificial":

No es mi intención sorprenderlos o asustarlos... Pero la manera más simple en que puedo resumirlo es -
diciendo que hay ahora en el mundo máquinas que -
piensan, aprenden y crean. Más aún, su habilidad -
para estas cosas se incrementará rápidamente hasta
que -en un futuro visible- el ámbito de los problemas que puedan resolver sea el mismo que el de la -
mente humana (Ref 512 p. 316)

He subrayado la frase "en un futuro visible", en el párrafo anterior, porque ilustra el espíritu del tipo de afirmaciones que Dreyfus critica. Aún ahora, casi 25 años después, afirmar que -

existen en el mundo máquinas que piensan, aprenden y crean, me parece, en el mejor de los casos, excesivo. Quizás, cabría discutir qué es lo que Simon trataba de decir (o, mejor aún, por qué trataba de decirlo). El párrafo de Simon, además, no parece sugerir ningún tipo de limitación, por lo que resulta engañoso.

Además, Simon hizo una serie de predicciones que esperaba - que se cumplieran alrededor de 1967. Incluan que una computadora sería campeón mundial de ajedrez, que sería descubierto y probado por una computadora un importante y original teorema matemático y que la mayoría de las teorías psicológicas tendrían la forma de programas de computadora.

En esos primeros años, los más importantes investigadores - pensaban o, al menos, así lo decían en sus publicaciones, que sólo se requería una "moderada extrapolación de las capacidades de los programas existentes" (ref 512) para cumplir este tipo de predicciones.

Transcurrido el doble del plazo propuesto por Simon, ninguna de las dos primeras predicciones se ha cumplido ni parece que puedan cumplirse en un futuro visible. Respecto a la última, se puede decir que, en efecto, ha habido una importante influencia de las ideas y de la terminología de la computación sobre la Psicología, aunque, quizás, las mismas condiciones que hacen posible que exista la Inteligencia Artificial favorecen un tipo afín de Psicología. Este tipo de Psicología no ha logrado resultados muy interesantes y, en todo caso, no parece deseable un enfoque que limi-

ta las perspectivas de estudio de la mente humana a las de la com
putadora digital.

Al analizar las dificultades encontradas por la Inteligencia Artificial, a lo largo de sus 25 años de historia, Dreyfus conclu
ye que el futuro visible del que habla Simon, parece alejarse cada vez más. Sin embargo, los investigadores que trabajan en Inte
ligencia Artificial aún se muestran optimistas y protestan ante las críticas, diciendo que 25 años es muy poco tiempo para juzgar los verdaderos alcances de una ciencia naciente. Según Dreyfus este optimismo está basado en una serie de hipótesis más o menos conscientes que estos investigadores adoptaron. Dreyfus formuló en un principio cuatro y, en una revisión posterior de su libro, agregó una quinta hipótesis, estas son:

1) Hipótesis biológica: A algún nivel - usualmente las -
neuronas - el cerebro procesa información en operaciones discre-
tas, por medio de algún equivalente biológico de interruptores -
on/off.

2) Hipótesis psicológica: La mente puede ser vista como un
aparato que opera en pedazos de información, de acuerdo con reglas
formales.

3) Hipótesis epistemológica: el conocimiento es formaliza-
ble. Todo lo que puede ser entendido puede ser expresado en tér-
minos de estructuras simbólicas.

4) Hipótesis ontológica: todo lo esencial para el comporta-
miento inteligente debe poder ser reducible, en principio, a un -
conjunto de elementos libres de contexto. Es posible descomponer

lo que es en un conjunto de hechos lógicamente independientes unos de otros.

5) Hipótesis metafísica: el fondo(contexto) puede ser tratado como otro objeto y es representable en el mismo tipo de descripción estructurada que se usa para los objetos comunes.

HIPOTESIS BIOLOGICA

Los primeros estudios de la electrofisiología del cerebro parecían apoyar la hipótesis de que éste funciona como una computadora digital. Si la evidencia empírica corroborara esta hipótesis, habría razón para esperar resultados interesantes al tratar de reproducir la inteligencia humana por medio de una máquina digital. En palabras de Dreyfus:

En la raíz del optimismo de los investigadores de la Inteligencia Artificial está la convicción de que el comportamiento humano inteligente es el resultado del procesamiento de información por medio de una computadora digital y, dado que la naturaleza logró un comportamiento inteligente con esta forma de procesamiento, la programación adecuada debería permitirnos obtener tal comportamiento en computadoras digitales, ya sea imitando a la naturaleza o programando, incluso, mejor que ella (ref. 512)

Antes de argumentar que el cerebro, probablemente, no es una máquina digital, vale la pena recordar cuál es la diferencia en -

to tiempo de proceso a las grandes computadoras modernas que, sim
plemente, dicha simulación resulta inaceptable. Por supuesto, -
existe la posibilidad de encontrar nuevas técnicas matemáticas -
que permitan elaborar nuevos formalismos que hagan práctica la si
mulación digital.

Suponer que el cerebro puede simularse mediante una computa-
dora digital implica la hipótesis de que existe un formalismo que
permite una representación práctica.

El hecho de que las computadoras sean máquinas universales -
no es la única razón por la que han acaparado casi toda la aten-
ción de la Inteligencia Artificial; otra razón, señala Dreyfus, -
es de orden filosófico:

... tales máquinas (las computadoras) sólo hubie-
ran sido máquinas de sumar superdesarrolladas si
no fuera porque el punto de vista platónico, refi-
nado por 2000 años de metafísica, encontró en ellas
su realización. Finalmente se disponía de una má-
quina que opera por medio de reglas sintácticas en
fragmentos de información. Más aún, las reglas es
tán incorporadas a los circuitos de la máquina.

Una vez que la máquina ha sido programada, ya no -
hay necesidad de interpretación, ya no se recurre
a la intuición y juicio humanos. Esto es lo que
Hobbes y Leibniz necesitaban. M. Heidegger, justi
ficadamente, vió en la cibernética la culminación

de la tradición filosófica (ref 512)

De la hipótesis que hemos mencionado la biológica es la que resulta más fácil de comparar con la evidencia empírica.

El cerebro humano está compuesto por un número de neuronas - del orden de 10^{11} . Una neurona típica consiste en un cuerpo celular de alrededor de 100 micrómetros de diámetro, del cual sale una fibra, más o menos larga, llamada axón, y un gran número de pequeñas ramificaciones llamadas dendritas. Hablando de una manera muy general, se puede decir que, a través de las dendritas y del cuerpo celular, la neurona recibe señales provenientes de otras neuronas; de alguna manera, se combinan estas señales en el cuerpo celular y se produce un impulso que se transmite a través del axón. En general, el axón se ramifica en su extremo terminal, lo que hace posible que el impulso se transmita a varias neuronas a la vez. Se estima que el número aproximado de sinapsis, conexiones entre distintas neuronas, en el cerebro humano, puede ser del orden de 10^{14} . Existen dos tipos de sinapsis: inhibitorias, que reducen la propensión de la neurona postsináptica para emitir un impulso, y excitatorias que aumentan dicha propensión.

Los impulsos que se transmiten a través de las redes neuronales son eléctricos sólo en parte. Lo que se transmite a través del axón es una onda de despolarización a lo largo de la membrana. Por medio de un intercambio muy selectivo de iones con el exterior, la célula es capaz de mantener el interior del axón a un cierto potencial negativo respecto al exterior. Si algún -

estímulo hace aumentar el potencial más allá de un cierto umbral, ocurre un cambio espectacular: un rápido intercambio de ciertos iones, entre el interior y el exterior de la membrana del axón, hace que la pared interna se vuelva localmente positiva; esto sirve de estímulo para que la zona adyacente se despolarice igualmente, mientras que se restablecen las condiciones iniciales en la primera zona despolarizada. Cuando el impulso llega a la terminal del axón, se libera un transmisor químico en la membrana pre-sináptica de la terminal, se difunde a través de la corta distancia que separa ambas células y afecta a la membrana post-sináptica de cualquiera de dos maneras:

En una sinápsis excitatoria, el transmisor reduce el potencial de la membrana post-sináptica, de manera que la célula post-sináptica tiende a generar impulsos con mayor frecuencia. En una sinápsis inhibitoria, el efecto del transmisor es estabilizar el potencial de membrana post-sináptica, con lo que se hace más difícil la labor de despolarización de las sinápsis excitatorias y, por ello, se evita que se produzcan nuevos impulsos o se reduce su frecuencia. (Hubel, ref 800)

Si el estímulo es demasiado pequeño, no se produce ningún impulso en el axón, pero, una vez que se ha sobrepasado el umbral requerido para que se genere un impulso, éste se produce sin importar la magnitud del estímulo. Este hecho es conocido como la

ley del todo o nada y es el único argumento que se podría proponer para apoyar la tesis de que el cerebro es una máquina digital.

Es posible comparar el cerebro con una computadora digital, ya que ambos son máquinas que procesan información, sin embargo, la comparación no puede ir más lejos, porque la computadora es una máquina digital y el cerebro no parece serlo, si tomamos en cuenta los párrafos anteriores. Para la computadora, cada estado discreto de cierto sistema representa un símbolo en un lenguaje descriptivo, es decir, representa un trozo específico de información. En una máquina analógica son variables físicas continuas las que representan la información que se está procesando. Si en el cerebro cada pulso pudiera relacionarse con algún símbolo en cierta forma de procesamiento de la información, podríamos considerarlo como una máquina digital; sin embargo, la frecuencia a la que los pulsos se transmiten es una parte esencial del proceso y ésta es una variable continua, lo que hace al cerebro más parecido a una máquina analógica que a una digital.

HIPOTESIS PSICOLÓGICA

La experiencia empírica no permite sostener la hipótesis biológica, es decir, que el cerebro funciona como una máquina digital, lo cual no invalida necesariamente los esfuerzos para reproducir el comportamiento inteligente por medios digitales. Es posible que el cerebro no sea una máquina digital pero eso no significa que la mente no funcione como si lo fuera. Esta hipótesis, llamada "psicológica", es la que se usa para justificar el uso de -

los modelos de computadora en psicología.

La mayor parte de quienes trabajan en el campo de la IA, no consideran esta hipótesis como tal, es decir, una afirmación que puede ser aceptada o descartada por medio de argumentos lógicos o por comparación con la evidencia empírica; simplemente, la aceptan a priori como si fuera un axioma.

El funcionamiento del cerebro humano es consecuencia de una serie de procesos físicoquímicos "continuos" (a un nivel macroscópico, pero nadie está sugiriendo una simulación de la mente a nivel cuántico). En principio, como hemos apuntado más arriba, es posible formalizar y calcular discretamente estos procesos. Esto permitiría a una computadora simular el comportamiento del cerebro. ¿Pero, en cuánto tiempo?

Uno de los ejemplos de procesos para los que existe un formalismo matemático que no es posible tratar en forma práctica por medio de una computadora digital, es el caso de dos sustancias en solución que evolucionan hasta alcanzar el equilibrio. Dreyfus se pregunta si esta solución sigue los mismos pasos discretos que seguiría una computadora al resolver las ecuaciones que describen el proceso. En este caso, apunta "la solución estaría resolviendo en momentos lo que costaría siglos a la máquina - en caso de que la máquina pudiera resolverlo". Así pues, el hecho de que los procesos concretos que ocurren en el cerebro puedan, en principio, formalizarse, no significa que el funcionamiento global pueda simularse.

Decir que la mente humana funciona como una computadora digital significa que debería ser posible reducir todos los procesos cognoscitivos a combinaciones de aquellos procesos que una computadora es capaz de realizar: comparación, clasificación, búsqueda de listas, etc. Sin embargo, hay algunos procesos realizados por la mente humana que no ha sido posible tratar de esta manera. Como ejemplo, se pueden mencionar el tipo de fenómenos a que se refiere la teoría de la Gestalt. Se trata de fenómenos globales, entre los cuales la visión es probablemente el ejemplo más claro. Al ver una escena no se "digitaliza", es decir, no se fragmenta en trozos muy pequeños para analizarlos uno por uno y construir, a partir de la información de cada trozo, la comprensión de la escena entera. Lo que, en realidad, parece que ocurre es que se comprende la escena globalmente, es decir, toda de una vez. Eventualmente esto permite entender cada una de las partes.

Quizás uno de los mejores ejemplos de la diferencia entre las computadoras digitales y la mente humana se encuentra en los intentos de diseñar programas para jugar ajedrez. Ya se ha mencionado antes la diferencia cualitativa que existe entre un jugador humano y una máquina. Vale la pena agregar un comentario sobre la nota aparecida en el Scientific American de septiembre de 1979 referente al campeonato de ajedrez para computadoras (North American Computer Chess Championship). En esta nota se habla de BELLE, programa capaz de examinar 5000 posiciones por segundo y que incluye en su memoria 150,000 posiciones de apertura. La ma-

por parte de los programas mencionados en la nota examinan exhaustivamente cada jugada posible y las correspondientes respuestas - posibles hasta cierta profundidad. Sin embargo, la clasificación de BELLE es de 1900 lo cual lo sitúa en clase A (la clase A va - de 1800 a 1999, las clases superiores son: experto, de 2000 a 2199 y maestro, de 2200 en adelante. El campeón mundial Anatoly Karpov alcanza 2705). La nota acepta que los maestros de ajedrez - juegan de una manera cualitativamente distinta:

Usualmente examinan en profundidad sólo dos o cuatro jugadas, que reconocen inmediatamente como - las importantes para el desarrollo a largo plazo - de la posición. Los maestros pueden jugar un juego notablemente fuerte usando sólo cinco minutos por jugada. La fuerza de un maestro no radica en la - habilidad de examinar cientos de secuencias posi - bles de jugadas sino en la habilidad de reconocer patrones y recordar "instantáneamente" posiciones similares que ha jugado él mismo o ha visto a otros jugar.

A pesar de que el cerebro no es capaz de realizar las operaciones de búsqueda y clasificación con la rapidez con que lo hace una máquina, utiliza otros procedimientos con los que obtiene resultados superiores. Algunos de los puntos que parecen importantes en este ejemplo son el reconocimiento de lo esencial y la capacidad para identificar patrones similares. Estos procesos cog-

noscitivos son consecuencia de la estructura peculiar del cerebro y hasta ahora no han podido duplicarse por medio de máquinas digitales.

HIPOTESIS EPISTEMOLOGICA

Aún admitiendo que el comportamiento humano puede no ser explicable si se supone que el cerebro realmente sigue reglas heurísticas cuando realiza operaciones inconscientes, cabe la posibilidad de plantear la hipótesis de que la inteligencia puede ser formalizable en términos de tales reglas, lo que haría posible su simulación por medio de una máquina digital.

En efecto, se puede pensar en un formalismo matemático que sirva para describir un fenómeno, sin pretender que este fenómeno se realiza reproduciendo de alguna manera el mismo formalismo. Se puede describir el comportamiento de los cuerpos que forman el sistema solar, hasta cierto punto, por medio de una serie de ecuaciones diferenciales basadas en la física newtoniana; pero no se pretende con esto que los planetas están, literalmente, resolviendo ecuaciones diferenciales cuando se mueven a través de sus órbitas. Sin embargo estas ecuaciones permiten construir un modelo del sistema solar en una computadora digital. La mente humana podría estar en el mismo caso: aún sin conocer las reglas, en caso de que éstas existan, en las que está basado el funcionamiento de la mente, podría ser posible formalizar este comportamiento por medio de algún tipo de reglas. Esta idea es lo que constituye la hipótesis epistemológica.

Dreyfus señala la diferencia que existe entre la hipótesis epistemológica y la psicológica:

(...) aquellos que aceptan la hipótesis psicológica
(...) suponen que las reglas usadas en la formalización del comportamiento son precisamente las mismas reglas que lo producen, mientras que los que adoptan la hipótesis epistemológica (...) sólo afirman que cualquier comportamiento no arbitrario puede formalizarse de acuerdo a ciertas reglas y que dichas reglas, cualesquiera que sean, pueden usarse en una computadora para reproducir el comportamiento. (ref 512)

Esta hipótesis puede resumirse en dos puntos: a) independientemente de la forma en que realmente funcionan el cerebro y la mente humana, la inteligencia es formalizable porque todo comportamiento no arbitrario puede ser formalizado, y, b) este formalismo puede usarse para reproducir el comportamiento en cuestión.

Como en los casos anteriores, es importante hacer notar que se está adoptando esta hipótesis al abordar los problemas de la IA y que no se trata sino de eso: una hipótesis, que puede resultar equivocada

En cierto sentido, el comportamiento humano puede considerarse como el resultado de una serie de reglas, si con ello, simplemente se quiere sugerir - que es ordenado. Pero aceptar que las reglas en -

cuestión son del tipo que puede transformarse en un programa de computadora o en algún formalismo equivalente, es una afirmación muy distinta y mucho más radical, que exige mayor justificación.

(ref. 512)

Dreyfus dice que la habilidad que permite a un ser humano hablar no puede ser completamente formalizada, que la hipótesis epistemológica no sólo no es plausible sino que lleva a contradicciones. Recuerda que Wittgenstein afirmaba que el lenguaje no se utiliza de acuerdo a reglas estrictas y que tampoco se aprende por medio de tales reglas. Uno de los argumentos de Wittgenstein empezaba suponiendo que, en efecto, todo el comportamiento no arbitrario debe estar basado en reglas y reducía esta hipótesis al absurdo al buscar reglas que se usan para aplicar las reglas y así sucesivamente.

Para tener una teoría completa de lo que la gente capaz de hablar puede hacer, no sólo es necesario un conjunto de reglas gramaticales y semánticas, sino también otras reglas que permitirían a una máquina o a una persona reconocer el contexto en el que las reglas deben aplicarse. Así que debe haber reglas para reconocer la situación, las intenciones de las gentes que hablan y así sucesivamente. Pero si la teoría, por consiguiente, re -

quiere todavía más reglas para poder explicar cómo deben aplicarse las reglas, según sugiere el punto de vista de los intelectuales puros, nos encontramos ante una regresión infinita. Pero como, de hecho, podemos usar el lenguaje, esta regresión no puede ser un problema para los seres humanos. Si la IA es posible, tampoco debe ser un problema para las máquinas. (ref. 512)

Para Wittgenstein no hay un final para esta regresión, simplemente hay tantos niveles de reglas como la situación específica requiere. En algún nivel, según la situación, la interpretación de la regla es evidente y la regresión se detiene. Según Dreyfus; en el caso de la IA, la regresión también se detiene con la interpretación que es evidente en sí misma, pero, dado que la computadora "no está en una situación", esta interpretación no tiene nada que ver con la situación. La solución del teórico de la computación depende de, finalmente, poder reducir la información a fragmentos libres de contexto, completamente determinados y que no requieren mayor interpretación para ser entendidos.

Una refutación total de la hipótesis epistemológica requeriría un argumento que probara que el mundo no puede ser analizado en términos de datos libres de contexto. Entonces, dado que la hipótesis de que existen elementos no ambiguos es la única manera de salvar a la hipótesis epistemológica de una re-

gresión infinita de reglas, el formalista, atrapado entre la imposibilidad de siempre tener reglas para la aplicación de reglas y la imposibilidad de encontrar datos primarios no ambiguos, tendría que abandonar totalmente la hipótesis epistemológica.

Esto lleva a Dreyfus a plantear otra hipótesis, la que supone que el mundo puede ser exhaustivamente analizado en términos - de datos libres de contexto o hechos atómicos, a la que llama "la más profunda hipótesis subyacente en el trabajo sobre la Inteligencia Artificial". La denomina hipótesis ontológica.

Dreyfus analiza algunas afirmaciones en el sentido de que no es posible, y nunca lo será, la traducción automática. Según Dreyfus esta posición, junto con otras similares, está basada específicamente en las limitaciones tecnológicas presentes: la afirmación no se hace apoyada en una imposibilidad de principio.

Sólo si uno rechaza la hipótesis ontológica de que el mundo puede ser analizado como un conjunto de - hechos-fragmentos de información- se puede legítimamente ir más allá de la imposibilidad práctica.

La mayor parte de los investigadores en Inteligencia Artificial, acepta actualmente la necesidad de colocar la computadora - en una "situación", como dice Dreyfus, o en un contexto que le -- permita resolver las ambigüedades y, en general, entender el lenguaje natural. Dreyfus encuentra que los mismos dos problemas - que se presentan en la resolución de ambigüedades que hicieron ne

cesario apelar a "la situación", se presentan de nuevo cuando se trata de reconocer el contexto. Estos dos problemas son: primero, encontrar, entre el gran número posible de hechos, aquéllos que - sean relevantes para aplicar un criterio de selección que permita reconocer el contexto; segundo, que, aún suponiendo que el programa pueda encontrar los detalles relevantes, los mismos serían ambiguos, ya que serían capaces de definir muchos contextos distintos hasta ser interpretados.

Evidentemente, tendría que usarse un contexto más amplio para determinar cuál de las características infinitas es relevante y cómo debe ser entendida - cada una. Pero si, a su vez, el programa debe permitir a la máquina identificar el contexto más amplio en términos de sus características relevantes - y ésta es la única manera en que procedería una computadora, que opera en términos de elementos discretos- el programador debe aceptar que algunas características son intrínsecamente relevantes y tienen un significado fijo independientemente del contexto -una posibilidad ya excluida en la apelación original al contexto- o bien tendrá que afrontar una regresión infinita de contextos. Parece que sólo hay una salida: En lugar de escalar el árbol buscando contextos cada vez más amplios, la computadora debe trabajar desde un contexto primario- lo que Weizenbaum -

llama "la cultura que compartimos"

Dreyfus piensa que, en efecto, existe algo a lo que se puede llamar el contexto primario, pero argumenta que éste es tan improgramable como la regresión que trata de evitar. Dreyfus propone que la única alternativa a la hipótesis ontológica es olvidar la independencia de los hechos y entenderlos como un producto de la situación. Esto equivale a argumentar que sólo en términos de su relevancia, determinada por la situación, son hechos. Ello evita, según Dreyfus, el problema de cómo reconocer la situación desde afuera, que plantea que, para que una Inteligencia tenga siquiera hechos que interpretar, ya debe estar en una situación.

LA HIPOTESIS METAFISICA

En el prólogo escrito en 1979 para su libro "What Computers Can't do", Dreyfus propone una hipótesis más, la hipótesis metafisica. Al describir el trabajo reciente en Inteligencia Artifi - cial en el cual se ha hecho énfasis en el problema del contexto, Dreyfus dice que los investigadores han tratado implícitamente de considerar el contexto más amplio como un objeto con su propio - conjunto de características descriptivas pre-seleccionadas.

Esta hipótesis de que el contexto más amplio puede tratarse, simplemente, como otro objeto que puede representarse con el mismo tipo de descripción es - tructurada en el que se representan los objetos cotidianos, es esencial para toda nuestra tradición filosófica. De acuerdo con Heidegger, quien es el pri

mero en identificar y criticar esta hipótesis, la llamaré "Hipótesis Metafísica".

Argumentar en torno a esta hipótesis es más difícil que en el caso de las anteriores. Dreyfus dice que es probable que no exista ningún argumento en contra de ella, dado que los hechos que se podrían mencionar para mostrar que el contexto más amplio es irrepresentable, se convierten entonces en hechos que se ha demostrado que son representables.

Para entender una historia, por ejemplo, es necesario tener un contexto de costumbres, entre otras cosas, que hagan inteligibles los detalles del comportamiento de los personajes. Para argumentar sobre esta hipótesis, Dreyfus recuerda la historia de Charniak (en esta historia dos niños están considerando comprar un regalo para un tercero; uno de ellos sugiere comprar un cometa pero el otro responde que su amigo ya tiene uno. Para entender esta historia es necesario saber que en el caso de algunos objetos, a la persona que lo recibe no le interesa que le regalen otro si ya tiene uno). En este caso, la única manera de entender el comportamiento de los personajes es por medio del conocimiento sobre los detalles de una serie de costumbres humanas. Pero el problema es cómo almacenar "los hechos" posibles o qué es posible comentar acerca de la acción de regresar regalos. Simplemente, puede haber una infinidad, o un número no definido, de razones para regresar un regalo, y cada una de estas razones presenta, a su vez, una cantidad indefinida de excepciones de las excepciones.

Dreyfus, simplemente, analiza una razón posible para regresar un regalo: si uno ya tiene un cierto objeto no le interesa tener otro igual. Por supuesto, para esto hay excepciones, como el caso de billetes, galletas o canicas. Vale la pena mencionar que, aún en el caso de las cometas, no es claro que este precepto funcione. Sería necesario, pues, incluir, junto con la regla, las posibles excepciones. ¿Pero, cuáles son las posibles excepciones? ¿Es siquiera posible enumerarlas? Además, habría que tener en cuenta - las excepciones de las excepciones, como, por ejemplo, las propuestas por Dreyfus, a saber: una galleta tan grande (3 pies de díametro), que una es suficiente; mil canicas que son en opinión de - Dreyfus, más de las que un niño normal puede manipular, etc. Es necesario especificar todas las excepciones a las excepciones. Pero estas excepciones también pueden tener excepciones y el proceso sigue así indefinidamente. El grupo de MIT opina que la mayor parte de las situaciones tienen suficiente en común con situaciones previamente encontradas como para que las características fundamentales puedan ser preanalizadas y almacenadas específicamente para la situación.

Dreyfus critica esta solución haciendo notar el omnipresente problema de decidir a qué situación pre-analizada es similar la presente. Además de otro problema: aún si toda la experiencia humana pudiera analizarse en función de situaciones estereotípicas, cualquier marco de referencia aplicable al mundo real debe describirse en términos de "lo normal", y el intento de caracterizar -

Las condiciones que determinan la aplicabilidad de la norma a un caso específico lleva necesariamente a una regresión. Sólo nuestro sentido general de lo que es típico puede usarse para decidir, y esta comprensión, por definición, no puede ser específica para una situación.

La idea que está detrás de todos los intentos de la IA es - que, independientemente de la naturaleza del conocimiento necesario para entender situaciones específicas, tal conocimiento debe estar representado, de alguna manera, en los seres humanos capaces de dicho entendimiento, y su representación debe ser en forma de una estructura explícita de datos.

Sin embargo, no es obvio que este conocimiento tenga que estar en forma explícita. Dreyfus, en este punto, utiliza el ejemplo del aprendizaje de una habilidad psicomotriz, a saber, aprender a nadar:

puedo aprender a nadar practicando hasta que desarrolle los patrones de respuesta necesarios, sin necesidad de representar mi cuerpo ni movimientos musculares en forma de alguna estructura de datos.

(ref. 512)

De esta misma manera, sugiere Dreyfus, se puede adquirir y almacenar el tipo de conocimiento necesario sobre las costumbres culturales, por ejemplo, que permita reconocer y actuar en situaciones específicas. Así señala que no es necesario que sea explícito alguna vez este tipo de conocimiento.

Ello implica una afirmación muy interesante, al comparar la habilidad psicomotriz con otro tipo de habilidades cognoscitivas, es decir, entre la habilidad de conducir una bicicleta con la de reconocer una escena visual. A pesar de las diferencias, muy importantes, que van desde la naturaleza de los conceptos que se manejan hasta las áreas relacionadas del cerebro, parece plausible pensar que existe una relación. ¿De qué otra manera se podría explicar la génesis de las operaciones cognoscitivas elementales si no es considerándolas fundamentadas en las operaciones psicomotrices elementales equivalentes? Por supuesto, existe la posibilidad de atribuir un origen genético a estas operaciones cognoscitivas, pero parece una explicación más admisible el considerarlas un producto de la interacción, a través de las operaciones psicomotrices, con el medio externo. Por su parte, esta habilidad psicomotriz estaría determinada, en una pequeña parte, genéticamente pero, a su vez, dependería para su desarrollo de la interacción con el medio, percibida y comprendida a través de la habilidad cognoscitiva. Son dos cosas, pues, que dependen mutuamente la una de la otra para desarrollarse.

Dreyfus concede un papel importante a la imaginación, al considerar algunos casos en los que se revela la presencia de representaciones que, sin embargo, no son representaciones formales. "[...] son usualmente representaciones no formales, más parecidas a imágenes, por medio de las cuales exploro lo que soy o lo que sé". Un ejemplo de dicho tipo de representaciones sería colocar

se uno, imaginariamente, en una situación específica y evocar todo el contexto de sensaciones, sentimientos, etc. que uno experimentaría al encontrarse realmente en tal situación. Según Dreyfus, ello permite, en el caso de la historia de Charniak, colocarse en el lugar de Jack y preguntarse cómo reaccionaría uno si recibiera una segunda cometa, sin tener que hacer explícito todo lo que sería necesario decir a una computadora para que pudiera llegar a la misma conclusión. Se trata de representaciones concretas (imágenes o memorias) basadas en la experiencia propia para las que no es necesario hacer explícitas las estrictas reglas formales requeridas por las representaciones simbólicas abstractas.

Dreyfus concluye:

"al volver la vista hacia los últimos 10 años de investigación en IA podríamos decir que el punto básico que ha salido a la luz es que, 'dado que la inteligencia tiene que estar situada, no puede estar separada del resto de la vida humana'".

CONCLUSION

Minsky decía que era absurdo juzgar una rama de la ciencia - por los resultados que ha obtenido en sólo 30 años. Desgraciadamente en el caso de la Inteligencia Artificial no tenemos más que eso. El resto es especulación. En este trabajo he tratado de - proporcionar al lector el estímulo y los datos necesarios para - que saque sus propias conclusiones sobre lo que puede ser el futuro de la Inteligencia Artificial.

No hay duda de que, en el futuro, las computadoras tendrán - un impacto cada vez mayor en la estructura de toda la sociedad. Pero esto no será una consecuencia inmediata de las investigaciones en Inteligencia Artificial. Como hemos visto a lo largo de - todo este trabajo, el viejo sueño de construir un artefacto cuya inteligencia sea comparable a la humana, en caso de que esto sea posible, no está cerca ni mucho menos. Y no estoy hablando de - componer sinfonías como Beethoven o escribir como Cervantes, sino "simplemente" de exhibir un comportamiento inteligente comparable al de un niño de dos años. Una de las cosas que hemos aprendido con estos estudios es a respetar la mente humana: presenta problemas tan complejos que harán falta muchas generaciones de investigadores para empezar a comprender sólo algunos de ellos. Espero que no haya nadie en el mundo capaz del tipo de declaraciones que hacía Simon hace 30 años. Si algo hemos adquirido en estos años es humildad.

Durante la discusión de algunas de las ideas centrales de la

Inteligencia Artificial (sólo he mencionado aquellas que me parecieron interesantes para la discusión del tema) me pareció imprescindible mencionar alguna de las críticas. Escogí la de Dreyfus porque me parece que es una de las más interesantes, pero tampoco es la única. Otra, quizás más conocida, es la de Weizenbaum, que tiene un enfoque totalmente distinto.

La esencia de la crítica de Dreyfus es señalar que existen una serie de hipótesis que es necesario adoptar si se pretende - que es posible la Inteligencia Artificial. Algunas de estas hipótesis son más difíciles de argumentar que otras, pero el punto - central no es rebatirlas exhaustivamente sino hacer notar que existen y que son eso: hipótesis, no verdades incontrovertibles. Desgraciadamente, algunas de estas hipótesis se escapan del ámbito - puramente científico para incursionar en el filosófico. Debo admitir que la mayor parte de los argumentos filosóficos de Dreyfus los entiendo sólo hasta cierto punto. Pero me parece que es importante mencionarlos, aunque sólo sea de la manera superficial - en que lo hice aquí, para insistir en que se trata de hipótesis - que es importante conocer.

Mi conclusión, después de terminar este trabajo, es la de que parece haber limitaciones intrínsecas para reproducir el comportamiento inteligente por medio de computadoras digitales. Posiblemente, será posible obtener soluciones específicas para ciertos - problemas, en un futuro muy cercano, pero no se podrá reproducir la inteligencia, ni al nivel más burdo, en mucho tiempo, en caso

de que esto sea posible. Es decir, el hecho de que sea posible - hacer reservaciones de boletos de avión hablando por teléfono con una computadora no significa que se haya resuelto el problema general de la inteligencia. Se tratará de otra ingeniosa solución ad-hoc nada más.

Si alguna vez nos aproximamos a la solución, seguramente no será con computadoras digitales tal como las conocemos ahora, sino con otros artefactos que construiremos utilizando nuestra experiencia con éstas.

Por otro lado, aunque el problema general de la inteligencia esté muy lejos de ser resuelto, la investigación en la Inteligencia Artificial ha producido, y producirá, muchas herramientas muy útiles en todas las ramas del conocimiento. Ejemplos de esto son los programas de consulta o de manipulación simbólicas.

A pesar de que no estamos cerca de cumplir las primeras profecías de Simon respecto al futuro de la Inteligencia Artificial, los frutos que hemos obtenido hacen que este estudio valga la pena.

A P E N D I C E

Conversación con MYCIN (ref. 448)

-----Paciente-1)-----

1) Nombre del Paciente (nombre-apellido)

##FRED SMITH

2) Sexo

##MASCULINO

3) Edad

##55

4) ¿Se han obtenido cultivos positivos de un lugar en el que Fred Smith tiene una infección?

##SI

-----INFECCION-1-----

5) ¿Qué infección es?

##BACTEREMIA-PRIMARIA

6) Por favor escriba la fecha en la que aparecieron por primera vez síntomas de la bacteremia-primaria (INFECCION-1) (mes/día/año)

##Mayo 5, 1975

Al primer organismo significativo de este cultivo (CULTIVO-1) se le llamará:

-----ORGANISMO-1-----

9) Identifique al ORGANISMO-1:

##DESCONOCIDO

10) ¿EL ORGANISMO-1 es un bacilo o un coco (etc.)?

Conversación con MYCIN (ref. 448)

-----PATIENT-1-----

1) Patient's name: (first-last)

##FRED SMITH

2) Sex:

##MALE

3) Age:

##55

4) Have you been able to obtain positive cultures from a site at which Fred Smith has an infection?

##YES

-----INFECTION-1-----

5) What is the infection?

##PRIMARY-BACTEREMIA

6) Please give the date and approximate time when signs or symptoms of the primary-bacteremia (INFECTION-1) first appeared. (mo/da/yr)

##May 5, 1975

The first significant organism from this blood culture (CULTURE-1) will be referred to as:

-----ORGANISM-1-----

9) Enter the identity of ORGANISM-1:

##UNKNOWN

10) Is ORGANISM-1 a rod or coccus (etc.)?

##BACILO

11) Tintura de Gram del ORGANISMO-1

##GRAM NEGATIVO

12) ¿Se han encontrado cultivos POSITIVOS de lugares que NO están asociados con infecciones sobre las que usted está solicitando consejo?

##NO

13) ¿Crece aeróbicamente el ORGANISMO-1?

##SI

14) ¿Crece anaeróbicamente el ORGANISMO-1?

##SI

15) ¿Cuál se sospecha que es la vía de entrada del ORGANISMO-1 a este lugar estéril?

##GI (Gastro intestinal)

...(siguen más preguntas, para acumular suficiente información para poder inferir las posibles identidades de los organismos)

Mi recomendación sobre una terapia estará basada en la(s) siguiente(s) infección(es) y organismo(s) responsable(s) probable(s):

El ORGANISMO-1 puede ser PSEUDOMONAS-AEURGINOSA

El ORGANISMO-1 puede ser KLEBSIELLA-PNEUMONIAE

El ORGANISMO-1 puede ser E. COLI

(...)

46) En una escala del 0 al 4 en la que los números mayores indican mayor severidad, cómo calificaría usted el grado de enfermedad de Fred Smith?

##3

A-4

##ROD

11) The gram stain of ORGANISM-1:

##GRAMNEG

12) Have there been POSITIVE cultures from sites which are NOT as associated with infections about which you are seeking advice?

##NO

13) Did ORGANISM-1 grow aerobically?

##YES

14) Did ORGANISM-1 grow anaerobically?

##YES

15) What is the suspected portal of entry of ORGANISM-1 into this sterile site?

##GI

(...siguen más preguntas, para acumular suficiente información para poder inferir las posibles identidades de los organismos)

My therapy recommendation will be based on the following probable infection(s) and potential causative organism (s):

INFECTION-1 is PRIMARY-BACTEREMIA.

The identity of ORGANISM-1 may be PSEUDOMONAS-AERUGINOSA

The identity of ORGANISM-1 may be KLEBSIELLA-PNEUMONIAE

The identity of ORGANISM-1 may be E. COLI

(...)

46) On a scale of 0 to 4, where higher numbers indicate increasing severity how would you rate Fred Smith's degree of sickness?

##3

47) Does Fred Smith have a clinically significant allergic reaction to any antimicrobial agent?

##NO

48) Fred Smith's weight in kilograms (or number POUNDS):

##70

(nôtese que dos drogas cubren 6 organismos potenciales)

Rec 1. My preferred therapy recommendation is as follows:

In order to cover for Items 1 2 3 4 5 6 :

Give GENTAMICIN

Doses: 119 mg (1.7 mg/kg) q8h IV or IM for 10 days

Comments: Modify dose in renal failure

In order to cover for Item 4:

Give: CLINDAMYCIN

Doses: 595 mg (8.5 mg/kg) q6h IV or Im for 14 days

Comments: If diarrhea or other GI symptoms develop, patient should be evaluated for possible pseudomembranous colitis.

Do you wish to see the next choice therapy?

##YES

Which, if any, of the above drugs do you wish eliminated from consideration?

##CLINDAMYCIN

(...)

#63

47 ¿Padece Fred Smith alguna reacción alérgica, clínicamente significativa, a algún agente anti-microbiano?

#64

48) Peso de Fred Smith en kilogramos (o número libras)

#70

(nótese que dos drogas cubren 6 organismos potenciales)

Rec 1 La recomendación de terapia que prefiero es la siguiente:

Para cubrir posibilidades 1 2 3 4 5 6:

Recete GENTAMICIN

Dosis: 119 mg (1.7 mg/kg) q8h IV o IM por 10 días

Comentarios: Modifíquela en caso de falla renal.

Para cubrir la posibilidad 4:

Recete CLINDAMYCIN

Dosis 595 mg (8.5 mg/kg) q6h IV o Im por 14 días

Comentarios: Si aparecen diarrea u otros síntomas GI, debe evaluarse la posibilidad de colitis pseudomembranosa.

¿Desea ver la siguiente opción de terapia?

#65

¿Cuál, si hay alguna, de las drogas anteriores le gustaría eliminar?

#CLINDAMYCIN

(...)

BIBLIOGRAFIA

- 64.- Arnheim, Rudolf: "Intelligence Simulated"; Midway; University of Chicago; Illinois; Junio 1967.
- 128.- Bar-Hillel Yehpshua: "The Present Status of Automatic Translation of Languages"; Advances in Computer Science; Academic Press; New York; 1960.
- 160.- Bobrow, Daniel: "Natural Language Input for a Computer - Problem Solving Program"; MAC-TR-1; MIT, abstract of thesis; Mass; 1962.
- 256.- Catell, Raymond B.: "Intelligence"; Encyclopaedia Britannica; Vol XII; 1970.
- 384.- Chaplin, James P.: "Systems and Theories of Psychology"; Holt, Rinehart and Winston; New York; 1968.
- 448.- Davis, Randall, Edward Shortliffe y Bruce Buchanan: "Production Rules as a Representation for a Knowledge-Based Consultation Program"; Stanford University Artificial Intelligence Laboratory Memo AIM 266; California; Octubre 1975.
- 512.- Dreyfus, Hubert: "What Computers Can't do"; Harper and Row; New York; 1979.
- 640.- Feigenbaum, E. A. y Julian Feldman eds.: "Computers and - Thought"; Mc Graw Hill; New York; 1973.
- 704.- Goldmeier, Erich: "Similarity in Visually Perceived Forms"; International Universities Press; New York; 1972.
- 736.- Haugeland, John: "The Plausibility of Cognitive Psychology"; The Behavioral and Brain Sciences; Vol 1, No. 2; 1968.
- 800.- Hubel, David H.: "The Brain"; Scientific American; Vol 241; No. 3; Septiembre 1979.
- 832.- Husserl Edmund: "Cartesian Meditations"; Martinus Nijhoff; La Haya; 1960.

- 864.- Kandel, Eric R.: "Small Systems of Neurons"; Scientific American; Vol 241; No. 3; Septiembre 1979.
- 896.- Marr, D.: "Artificial Intelligence--A Personal View"; MIT Artificial Intelligence Laboratory Memo No 355; Mass; Marzo 1976.
- 922.- Minsky, Marvin y Seymour Papert: "Artificial Intelligence"; Condon Lectures; Oregon System of Higher Education; Oregon; 1973.
- 928.- Minsky, Marvin: "A Framework for Representing Knowledge"; MIT Artificial Intelligence Laboratory Memo 306; Mass; Junio 1974
- 960.- Minsky, Marvin: "Artificial Intelligence"; Scientific American; Vol 215; No 3; Septiembre 1966.
- 1024.- Minsky, Marvin: "Descriptive Languages and Problem Solving"; Semantic Information Processing; MIT Press; Mass; 1969.
- 1088.- Minsky, Marvin y Seymour Papert: "Draft of a Proposal to ARPA for Research on Artificial Intelligence at MIT"; Mass; 1970-1971.
- 1152.- Minsky, Marvin, ed.: "Semantic Information Processing"; MIT Press; Mass; 1969.
- 1280.- Neisser Ulric y O. G. Selfridge: "Pattern Recognition by Machine"; Computers and Thought; Mc Graw Hill; New York; 1963.
- 1408.- Nilson, Nil: "Artificial Intelligence"; IFJP; 1974.
- 1536.- Pask, Gordon: "Cybernetics"; Enciclopedia Britannica; - Vol VI; 1970.
- 1664.- Piaget, Jean: "Seis Estudios de Psicología"; Seix Barral; Barcelona; 1973.
- 1792.- Piaget, Jean: "The Child and Reality"; Penguin; New York; 1976.

- 1856.- Quillian, Ross: "Semantic Memory"; Bolt Beranek and Newman, Inc; paper AFCRL-66-189; Octubre 1966.
- 1888.- Quillian, Ross: "Semantic Memory"; Semantic Information - Processing; MIT Press; Mass; 1969.
- 1904.- Raphael, Bertrand: "The Thinking Computer"; W. H. Freeman Press; San Francisco 1976.
- 1920.- Sagan, Carl: "Broca's Brain"; Random House; New York; 1979.
- 1984.- Shank, Roger y Kenneth Colby eds.: "Computer Models of Thought and Language"; W. H. Freeman Press; San Francisco; 1973.
- 2016.- Shank, Roger: "Panel on Natural Language Processing"; - IJCAI-77; Proceedings; 1977.
- 2176.- Selfridge, O. G. y Ulric Neisser: "Pattern Recognition by Machine"; Computers and Thought; Mc Graw Hill; New York; 1963.
- 2240.- Shank, Roger: "Using Knowledge to Understand"; Theoretical Issues in Natural Language Processing; Mass; Junio 1975.
- 2304.- Wertheimer, Max: "Productive Thinking"; Harper and Bros; New York; 1945.
- 2378.- Winograd, Terry: "A Procedural Model Of Language Understanding"; Computer Models Of Thought and Language; W.H. Freeman Press; San Francisco; 1973.
- 2400.- Winograd, Terry: "Artificial Intelligence and Language Comprehension"; Artificial Intelligence and Language Comprehension; National Institute of Education; Washington, D.C.; 1976.
- 2416.- Winograd, Terry: "Computer System for Natural Language"; Five Lectures on Artificial Intelligence; Stanford - University Artificial Intelligence Laboratory Memo AIM 246; California; Septiembre 1974.

- 2432.- Winograd, Terry: "On Some Contested Suppositions of Gene
rative Linguistics About the Scientific Study of Lang-
uage"; COGNITION, Vol 5, 1977.
- 2496.- Winograd, Terry: "Representation: Formalisms for Knowledge";
Five Lectures on Artificial Intelligence; Stanford Uni-
versity Artificial Intelligence Laboratory Memo AIM 246;
California, Septiembre 1974.
- 2560.- Winston, P. B.: "Artificial Intelligence"; Addison-Wesley;
Mass; 1977.