

POSICION FILOSOFICA Y ESTADO TECNOLÓGICO DE LA INTELIGENCIA ARTIFICIAL

Luis Guillermo Restrepo

El objeto del presente artículo es, en primer lugar, mostrar un marco filosófico en el cual, a juicio del autor, se mueve la investigación en Inteligencia Artificial (I. A.), al parecer diferente a la visión de la inteligencia que se tiene popularmente por "sentido común". El enfoque se basa en una confianza en los logros de la I. A. a largo plazo. En segundo lugar, presentar un resumen del estado actual en cuanto a primeros logros y áreas de exploración.

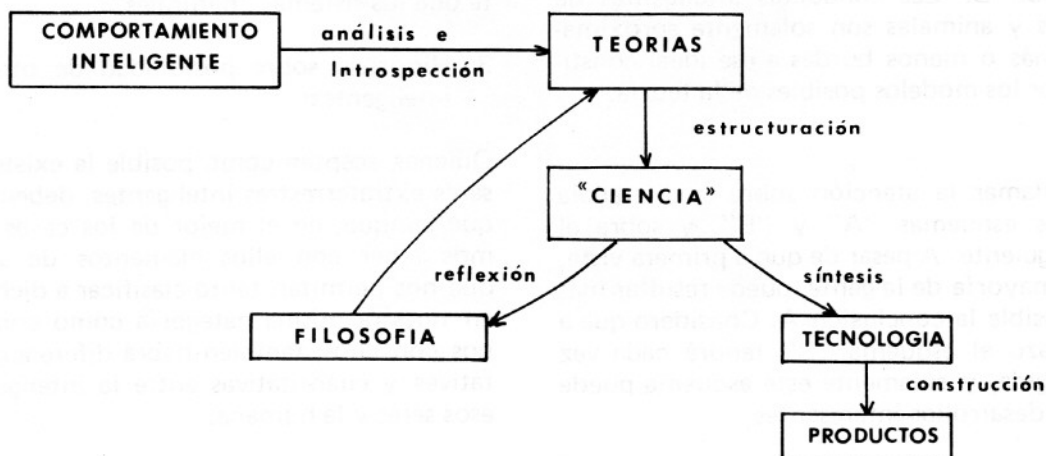
Muchos investigadores, como Roger Schank ven dos aspectos al responder qué es la I. A.: Una ciencia y una tecnología que se tratan de desarrollar (1). Aunque la I. A. aún no se puede considerar como una ciencia, al tratar de darle forma a ésta, a partir de teorías fragmentarias, se van obteniendo implicaciones filosóficas y resultados tecnológicos, como he tratado de esquematizar en la figura 1.

Supuesto filosófico implícito en la I. A.

A mi juicio, el principal supuesto filosófico implícito en la I. A. consiste en asignar a algunas de las tareas que requieren de la inteligencia humana un status menos "misterioso" que el asignado tradicionalmente.

A pesar del interés y publicidad que hay actualmente acerca de la I. A., es preciso reconocer que estamos aún en la infancia de esta disciplina y sus logros hasta ahora son más bien limitados.

Así lo expresan algunos investigadores: "La I. A. es un campo en donde el desacuerdo es más común que la teoría sólida, y las ideas interesantes más comunes que los programas de computador refinados. Aun así, va tomando forma un núcleo teórico y práctico aceptado (aunque no universalmente aceptado)" (2).



Siendo optimistas respecto al progreso en el desarrollo de la I. A. esperamos que el referido cuerpo de teorías se vaya estructurando hasta formar, a largo plazo, una teoría formal de las capacidades demostradas por los seres inteligentes. Esto es interesante, ya que en el fondo se aparta de una idea preconcebida muy popular, idea que podemos caracterizar con el esquema de pensamiento siguiente:

Hipótesis A: La mayor manifestación de inteligencia que puede existir corresponde al ser humano. Las demás, como las inteligencias de los animales son necesariamente inferiores.

Conclusión A: Los posibles logros de la I. A. son solamente aproximaciones siempre burdas a ese ideal constituido por la inteligencia humana.

Este esquema "A" parece razonable a primera vista, pero trataré de demostrar que es más correcto el siguiente, bastante opuesto en su conclusión.

Hipótesis B: Con base en los comportamientos inteligentes exhibidos por humanos y animales, es posible inducir una teoría de la inteligencia, cuyos modelos pueden ser posteriormente investigados y desarrollados en abstracto, sin mucha referencia a la existencia de entes que posean los comportamientos obtenidos por la teoría.

Así los modelos teóricos pueden trascender las capacidades de los entes conocidos y guiarnos en el diseño de nuevos sistemas.

Conclusión B: Las conductas inteligentes de humanos y animales son solamente aproximaciones más o menos burdas a ese ideal constituido por los modelos posibles de la teoría.

Quiero llamar la atención sobre la diferencia entre los esquemas "A" y "B", y sobre el hecho siguiente: A pesar de que a primera vista, para la mayoría de la gente, puede resultar más comprensible la conclusión A. Considero que a largo plazo el esquema "B" tendrá cada vez más acogida, y solamente este esquema puede permitir desarrollos interesantes.

El esquema "A" está limitado por definición; pero el desarrollo de la I. A. se encargará de

desplazar la opinión general hacia el esquema "B".

Argumentos:

1. Ejemplos de otras disciplinas:

Si observamos el desarrollo de otras disciplinas como la electrónica, vemos que aunque los objetos de estudio de que se ocupó inicialmente (como bloques de materia, bobinas, etc.) tienen un comportamiento eléctrico complejo y limitado, luego se trasciende ese tipo de objetos y se llega a la concepción de modelos ideales como "resistencia", "inductancia", "capacitancia", "amplificador ideal", etc. y se pueden imaginar otros tipos de comportamientos eléctricos aunque no se conozcan objetos naturales que los posean. Una análoga dirección podrá tomar la I. A.

2. Dinamismo de la investigación científica:

El ímpetu propio de la investigación teórica trata siempre de superar las aparentes limitaciones de lo conocido, buscando nuevas generalizaciones y alternativas.

Así, la investigación en I. A. busca: caracterizar lo esencial del comportamiento inteligente, dilucidar los posibles mecanismos de dicho comportamiento, y reproducir esos mecanismos en máquinas. En este proceso, de manera natural, surge la concepción de otros mecanismos alternativos que, aunque parezcan no ocurrir en la naturaleza, se ven como teóricamente capaces de exhibir también un comportamiento inteligente, y, tal vez, de manera más eficiente que los sistemas "naturales".

3. Hipótesis sobre posibilidad de otros seres inteligentes:

Quienes aceptan como posible la existencia de seres extraterrestres inteligentes, deben admitir que aunque, en el mejor de los casos, esperamos tener con ellos elementos de similitud que nos permitan tanto clasificar a dichos seres en nuestra misma categoría como comunicarnos con ellos, también habrá diferencias cualitativas y cuantitativas entre la inteligencia de esos seres y la humana.

4. Antropocentrismo de la hipótesis y conclusión "A":

Examinando el asunto con más elementos de juicio, la hipótesis "A" se nos revela de motivación antropocéntrica, similar a la teoría que consideraba a la tierra como centro del universo.

Nuestros puntos de vista iniciales son a menudo condicionados por las circunstancias particulares en que nos toca vivir, por nuestro ancestro y cuerpo biológicos, y por nuestras experiencias, que en este caso se manifiestan así:

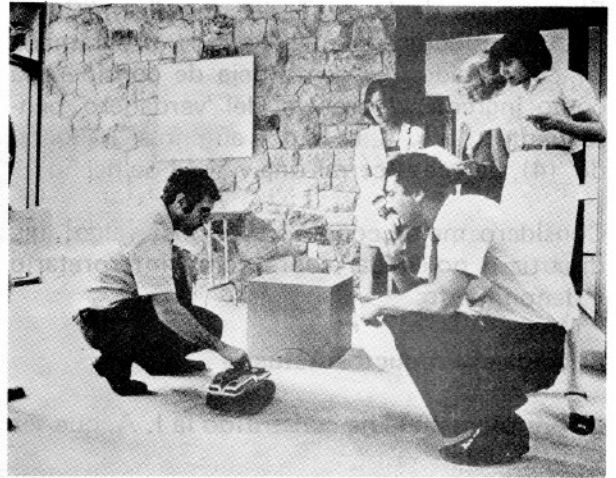
Los únicos seres de que tenemos conocimiento directo, conciencia, somos nosotros mismos, de manera individual y esencialmente incomunicable: Sólo uno mismo sabe cómo es eso de ser uno mismo(3). Así que asumir que otro ser pueda tener conciencia y otros sentimientos afines sólo lo admitimos por la semejanza entre esos seres y nosotros mismos, tanto en aspecto físico como en comportamiento, y mientras más cercanos en la escala biológica más concedemos a los animales sentimientos y funcionamiento mental o cerebral similares a los nuestros: imaginamos la tristeza de un perro e inclusive de un pajarito, pero nos cuesta más trabajo pensar que una cucaracha pueda sentir algo similar. Y todo esto es sumamente apriorístico, ya que aún no comprendemos de manera analítica, funcional, interna, qué significan los sentimientos, la conciencia o la inteligencia.

Con mayor razón no podemos admitir, de buenas a primeras, que una máquina pueda poseer algo siquiera semejante a cualidades humanas tan preciadas y tan "humanas" como: conciencia, sentimientos e inteligencia.

Sin embargo, siendo estrictamente consecuentes tampoco tenemos evidencia directa de la conciencia de otras personas, las únicas evidencias son indirectas: esas personas tienen un origen biológico igual al mío, son físicamente iguales, y su comportamiento también es igual. A pesar de lo indirecto de esas evidencias, muy pocos caen en el solipsismo de creerse los únicos existentes conscientes. Respecto a las máquinas, la prueba de Turing podrá aportar la evidencia behaviorista, las demás evidencias de inteligencia no parecen tan cruciales.

5. Limitaciones humanas:

El ejercicio de comportamientos inteligentes que puede tener un ente, está limitado por su



capacidad de interacción, que podemos considerar dividida en dos: Su Espacio Perceptivo, entendido como el conjunto de señales que puede percibir, y su Espacio Pragmático, o sea sus actuaciones posibles sobre el universo. Ahora bien, lo que esto implica es que el ser humano está limitado a percibir ciertos tipos de señales (y anchos de banda de éstas, como sonidos audibles, luz visible, etc.) que sabemos son bastante limitados. Para poder ampliar su espacio perceptivo debe utilizar transductores que conviertan desde el espacio ampliado a su espacio limitado (por ejemplo, de radiación infrarroja a imágenes visibles) y en sentido inverso en cuanto a su espacio pragmático. La máquina puede pues ejercer su comportamiento inteligente en espacios mucho mayores, de manera directa.

En resumen:

Vemos que podemos considerar la inteligencia como un cuerpo de comportamientos potenciales que cubren una amplia gama teórica, cualitativa y cuantitativa, y que son susceptibles de realizarse en seres concretos (animales, humanos, máquinas, extraterrestres, etc.) de manera exacta o aproximada.

Creo que no es muy productivo continuar la interminable discusión acerca de si pueden o no pensar las máquinas, pues esto ya ha sido bastante tratado, y al parecer, los argumentos en pro y en contra no son convincentes para uno u otro bando y se llega a posiciones juzgadas de fe por el bando opuesto.

El principio de Larry Tesler que dice: "Cuando una función mental ha sido programada en computador, la gente deja de considerarla como ingrediente esencial del 'verdadero' pensamiento: La I. A. es lo que aún no se ha logrado"(4), parece encerrar una verdad de hecho.

Considero mejor construir y utilizar algo, que discutir o polemizar sobre cómo interpretarlo y denominarlo.

Principales Productos Actuales

Ahora bien, ¿qué ha construido la I. A. que sea ya utilizable?

Se han logrado varios sistemas interesantes en el campo de investigación, aunque son menos los productos lanzados a la arena comercial e industrial.

Debemos reconocer que, así halla varios productos comerciales útiles atribuidos a la I. A., son los pioneros de este tipo de sistemas y quedan muchas mejoras por hacer.

En estado de primeras penetraciones en el mercado, podemos citar los siguientes sistemas, por los nombres más conocidos, aunque a veces se confunden sus aspectos en una sola implementación:

— **Interfaces de lenguaje natural:** Son subsistemas que permiten al usuario comunicarse con la máquina en un diálogo más natural o "humano" que los lenguajes restringidos y formalizados utilizados en los programas tradicionales.

Un módulo de éstos tiene aplicación como componente de otros sistemas de software: Puede ser simplemente el interface frontal (front end) que ve el usuario de una aplicación cuya tarea central puede ser en sí poco relacionada con el lenguaje como tal, por ejemplo: consultar información de una base de datos, resolver un problema matemático, diseñar un objeto, etc.; o la tarea del sistema puede centrarse en el lenguaje en sí, por ejemplo: producir resúmenes de narraciones, traducir idiomas, responder preguntas sobre una narración, etc. Como ejemplos concretos podemos mencionar el INTELLECT, interface que permite consultar bases de datos en un inglés natural, y el Unix Consultant que permite al usuario

un diálogo también natural con el sistema operacional Unix.

— **Sistemas llamados "expertos"** que, utilizando una base de conocimientos sobre un dominio limitado del saber, resuelven consultas siguiendo razonamientos similares a los que hace un experto en el tema.

Estos sistemas se han aplicado a muchas especialidades, como: diagnóstico médico, reparación de equipos, prospección minera, etc.

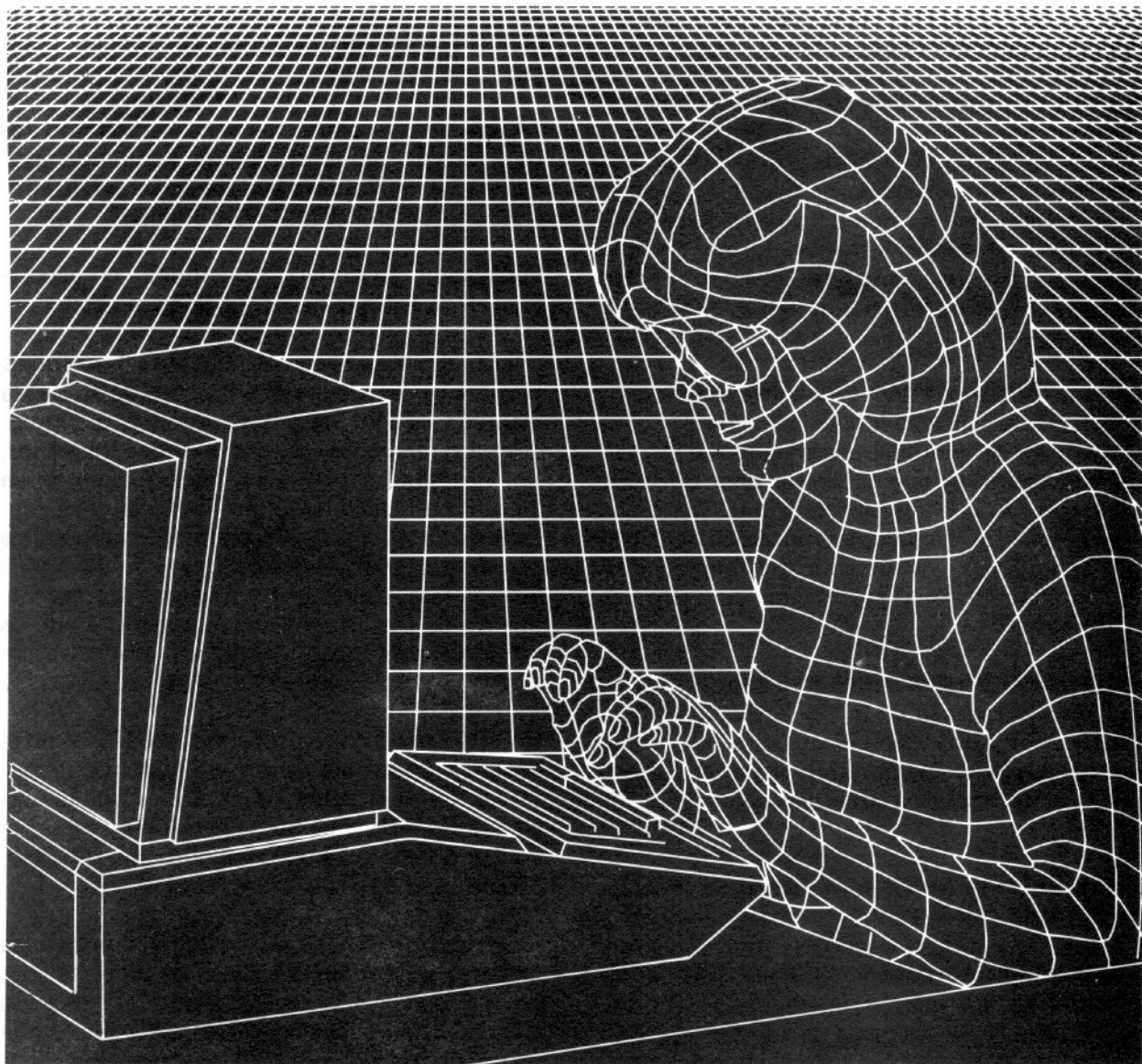
En este momento son ya muy numerosos y prefiero no mencionar algunos en particular para no ser inequitativo en su apreciación.

— **Sistemas basados en Conocimientos (Knowledge-based systems):** estrictamente esta denominación es bastante extensa y cubre prácticamente todo sistema que utilice una "base de conocimientos" como componente esencial para guiarse en la solución de determinada tarea. Sin pecar mucho se usa esta denominación como sinónimo de sistemas de I. A.

Existen en aplicaciones de diversos dominios, por ejemplo: Sistemas para el diseño de circuitos integrados VLSI conocidos como "compiladores para silicio", los cuales, a partir de una descripción lógica del funcionamiento deseado en un nuevo circuito, sintetizan el circuito agilizando inmensamente el paso de concepción a circuito real.

— **Herramientas de software para I. A.:** Los menciono aquí por estar relacionados con la I. A., aunque ellas en sí pueden no estar hechas con técnicas de I. A. Se trata de lenguajes y sistemas que facilitan la elaboración de aplicaciones basadas en conocimientos, tales como sistemas expertos, por ejemplo: editores para reglas de inferencia, sistemas que inducen reglas de a partir de una serie de ejemplos o casos particulares (como el sistema Radial de la compañía Radian), mecanismos de inferencia desprovistos de base de conocimientos, de propósito general (como el Emycin), etc.

Los sistemas Robot actuales, que alguien podría esperar en esta lista, dada su publicidad y penetración en algunas industrias, realmente no se deben clasificar por ahora dentro de los logros de la I. A. puesto que en general no poseen capacidades sensoriales ni de software



lo suficientemente sofisticadas, son tan solo manipuladores o herramientas flexibles, programables (y algunas de las cuales requieren supervisión constante de un operador humano); pero aún muy aisladas sensorialmente del medio, y poco o nada "inteligentes". Por lo tanto son más bien logros de la mecánica y los servomecanismos.

Investigaciones Genéricas Actuales:

También en cuanto a temas de investigación no se puede hacer una discriminación neta, ya que los diversos aspectos mencionados a continuación se entremezclan unos con otros — después de todo la inteligencia se manifiesta como una amalgama de capacidades:

Razonamiento Lógico: La mecanización del razonamiento monotónico, que consiste solamente en encadenamiento de silogismos o reglas de inferencia, —bien sea desde la verificación de premisas hacia la deducción de una conclusión, o desde la verificación de una conclusión hacia la inducción de unas premisas— ha sido exitosa, aunque hay que reconocer que este tipo de razonamiento tiene aplicaciones limitadas en la inteligencia humana, condicionadas a que haya tanto información como reglas de inferencia confiables y absolutas.

Otras complejidades del razonamiento, como la no-monotónica (cuando nueva información cuestiona la anterior), y el razonamiento con

información no confiable continúan siendo materia de investigación.

Muchos centros de investigación se ocupan de estos temas.

Heurística: La utilización de principios heurísticos para guiar el comportamiento ha sido fructífera en varios sistemas. Como ejemplo típico se pueden mencionar AM (automated mathematician) y Eurisko(5), desarrollados por D. Lenat.

Este mecanismo tiene gran potencialidad de aportar resultados prácticos sumamente útiles y por esto se continúa investigándolo. Un rasgo interesante es que se presta para un bucle recursivo, ya que los principios heurísticos no sólo guían, por ejemplo, la utilización de reglas, o la adquisición de conocimiento, sino que pueden guiar también la utilización de los principios heurísticos.

Estas investigaciones se llevan a cabo en varios centros, en especial en la Universidad de Stanford y en el MCC (Microelectronics and Computer technology Corporation, que de cierto modo es la respuesta estadounidense al ICOT japonés).

Representación del Conocimiento: Es un campo en donde se desean progresos, pues al parecer está un poco estancado. Se utilizan estructuras como los marcos (frames (6)), redes semánticas, etc., que en general datan del decenio 1970-1979. De esta área, así como en general de sistemas basados en conocimientos y sistemas expertos se ocupan muchos laboratorios.

Adquisición de Conocimientos: ¿Cómo dar al computador una buena y amplia base de conocimientos? en la hora actual, esto puede considerarse como uno de los mayores (¿menores?) cuellos de botella para disponer de sistemas potentes en cuanto a la amplitud, profundidad y flexibilidad de sus conocimientos.

Se trabaja tanto en la transferencia de conocimientos ya compilados y listos para ser utilizados por el sistema, como en el aprendizaje por parte del sistema, a partir de información bruta como: casos particulares y conocimientos análogos.

Otro subproblema interesante es la fusión de conocimientos provenientes de diversas fuentes.

Dignos de mencionar en este campo son los laboratorios de la compañía Schlumberger-Doll, la U. de California en los Angeles (UCLA), y las corporaciones MITRE y Radian.

Percepción sensorial: Constituye el acoplamiento de entrada de información al sistema.

Fuera de la utilización de los dispositivos periféricos particulares y ya clásicos de la máquina, como teclados, ratones, palancas, etc., los análogos de los sentidos humanos, tan necesitados por la robótica, todavía están en situación incipiente, no en la etapa de captación por medio de transductores, sino en las etapas de: transferencia de la información a una representación útil y eficiente, aprovechamiento de esa información, e integración de las informaciones provenientes de diversos sensores.

En este terreno hay problemas arduos como el de la visión en ambientes reales donde se desplazan tanto el autómatas como los objetos circundantes. Obviamente, la percepción sensorial corresponde a un tipo de adquisición de conocimientos a partir de señales brutas.

En el campo de máquinas videntes se destaca la investigación en la UCLA, las universidades de Maryland y Carnegie-Mellon, y la empresa General Electric. La comunicación hablada con las máquinas es objeto de investigación en el ICOT, la firma Fairchild, y el Air Force Institute of Technology (AFIT).



Arquitectura de máquina: El desarrollo de nuevas arquitecturas de computador, diseñadas específicamente para ser eficientes en el manejo de la información simbólica propia de las aplicaciones de I. A. es un tema que ocupa a los investigadores de varios centros en el mundo. Podemos mencionar al respecto las llamadas "máquinas Lisp" americanas, el proyecto francés MAIA (7), y las japonesas del proyecto de quinta generación (8). También hay desarrollos en la firma Fairchild, la corporación MITRE, la General Electric y el Jet Propulsion Laboratory (JPL).

Lenguaje Natural: Se investiga en especial en la UCLA, la U. de Stanford, la GTE y el AFIT (como interface para sistemas de pilotaje).

Programación automática, verificación automática de programas y temas afines, aplicados a la ingeniería del software, son asunto de investigación en la U. de Yale, la compañía Schlumberger-Doll y el Information Sciences Institute.

Algunas aplicaciones concretas investigadas

El diagnóstico médico con técnicas de I. A. continúa siendo objeto de exploraciones en varios lugares, como las universidades de Vanderbilt, Stanford y Maryland.

El diagnóstico de fallas en máquinas se investiga en especial en: la U. de Vanderbilt, el JPL, el AFIT, y el centro Langley de la NASA.

La prueba automática de teoremas se tiene como objeto de investigación en la U. de Texas.

El planeamiento de tareas o misiones se investiga en el JPL, el AFIT, la corp. MITRE, y los varios centros que se ocupan de la robótica, como el centro francés en Montpellier.

En la UCLA también es digno de mencionar el proyecto aplicado a la **jurisprudencia**.

BIBLIOGRAFIA

- (1) Schank, R. C. "The current state of A. I.: one man's opinion", revista "The A. I. magazine", vol. 4, Nro. 1, 1983, p. 3.
- (2) Charniak, E.; C. K. Riesbeck, y D. V. McDermott: "A. I. Programming", pag. ix, Lawrence Erlbaum Associates Publishers, 1980, Hillsdale, N. J., Estados Unidos.
- (3) Hofstadter, D. R. y D. C. Dennet: "The Mind's I", pag. 9, Penguin Books, Middlesex, Inglaterra, 1982.
- (4) Hofstadter, D. R.. "Gödel, Escher, Bach: an eternal golden braid", pag. 601, Vintage Books Div. of Random House, Nueva York, 1980.
- (5) Lenat, D. B. y J. S. Brown, "Why AM and Eurisko appear to work", revista "Artificial Intelligence", 23 (1984), págs. 269-294, Elsevier Publishing B. V., Holanda, 1984.
- (6) Minsky, M., "A framework for representing knowledge", en "The psychology of computer vision", ed. por P. H. Winston, pags. 211-276, McGraw-Hill Inc., Nueva York, 1975.
- (7) Bourgault, S. y J. P. Sansonnet, "MAIA: une machine pour les applications en intelligence artificielle", revista "L'écho des Recherches", no. 119/120, CNET, París, 1985.
- (8) ICOT, "Outline of research and development plans for fifth generation computer systems", Tokyo, abril 1984 (3a. edición).