

Colección La Empresa ®

TEMAS DE INTELIGENCIA ARTIFICIAL

TEMAS DE INTELIGENCIA ARTIFICIAL

por

Eduardo Daniel Sontag

con la colaboración de

José Luis Tesoro

Miembros del GEDECC

(Grupo de Estudio y Desarrollo en Ciencias de la Computación)

PROLAM S.R.L.

BUENOS AIRES / ARGENTINA / 1972

© 1972 by *Prolam S. R. L.*
Casilla de Correo 1495 - Buenos Aires - Argentina,
para su Colección La Empresa ®

Hecho el depósito que marca la ley 11.723
Reservados todos los derechos para el idioma español
Impreso y editado en la República Argentina
® *Marca Registrada*

INDICE

PROLOGO, 7

PREFACIO, 9

CAPITULO I: NOCIONES PRELIMINARES, 13

- 1.1. Sistemas, modelos y algoritmos, 13
- 1.2. Algunos modelos particulares; computadoras digitales, 19
 - 1.3. Los lenguajes y las máquinas, 23
 - 1.3.1. Estructura de frase; ambigüedades, 24
 - 1.3.2. Gramática transformacional, 29

CAPITULO II: SIMULACION DE PROCESOS INTELECTUALES HUMANOS, 31

- 2.1. Aplicaciones generales, 41
- 2.2. Aplicaciones en psiquiatría, 43
- 2.3. Automatas que sienten, 48

CAPITULO III: INTELIGENCIA ARTIFICIAL, 53

- 3.1. El método heurístico, 70
 - 3.1.1. Máquinas que juegan, 72
 - 3.1.2. Aprendizaje elemental de los juegos, 93
 - 3.1.3. Métodos heurísticos para temas "formales", 99
- 3.1.4. Un aprendizaje para la resolución "formal" de problemas, 105
 - 3.1.5. La solución de problemas, 107
- 3.2. "Pattern Recognition", 115
 - 3.2.1. Reconocimiento de caracteres alfabéticos y figuras geométricas, 123
- 3.2.2. El análisis de escenas y objetos tridimensionales, 128
 - 3.2.3. Conceptos y estructuras arborescentes, 135
- 3.3. Semántica y Representación, 137
 - 3.3.1. Lenguaje y Comunicación, 140
 - 3.3.2. Estructuras, 151
 - 3.3.3. Un modelo de aplicación, 156
- 3.4. Modelos Generales y Enfoques Formales, 159
 - 3.4.1. Ejemplos de la biónica. El Perceptrón, 165
 - 3.4.2. Estudios Probabilísticos y Estadísticos, 167
 - 3.4.3. Otras Formalizaciones, 172

CAPITULO IV: OBSERVACIONES FINALES, 177

- 4.1. ¿Y ahora qué?, 184
- 4.2. Conclusiones, 187

PROLOGO

El objetivo de la Inteligencia artificial es: "El estudio teórico y experimental de los diversos procesos intelectuales con el fin de conseguir una computadora que actúe, perciba y entienda en formas por el momento sólo posibles a los seres humanos", en el lenguaje más preciso de los autores.

Extraño animal el ser humano. Egocéntrico incurable —como se ve al estudiar a lo largo de la historia el concepto que tiene de su ubicación en el universo— sufriría terriblemente si se descubriera otra raza inteligente que le compitiera.

No obstante no le desagrada que sus creaciones puedan tener su imagen, porque inconscientemente percibe que lo que es su fruto puede limitarse o destruirse, como el niño percibe que su muñeco no es la competencia, sino el ser de carne y hueso, lejos de su alcance y voluntad, que viene al hogar a quitarle la posesión absoluta del centro.

Esa es la inteligencia artificial, apasionante disciplina que este libro trata por vez primera en nuestro medio en forma seria, metódica y documentada.

Es de aquellas disciplinas a las cuales el hombre se ve impulsado más que por una necesidad, por una avasallante curiosidad.

Es un poco como el viaje a la Luna, el hombre rompe la barrera de la gravedad sin tener claros sus objetivos.

Aquí el hombre juega con su creación favorita y no sabe tampoco bien porqué. Pero qué apasionante ver levantarse a esa organización de circuitos que es una computadora, y por ese mecanismo de adición de complejidades y circunstancias aleatorias hacer salir respuestas que los propios creadores de los algoritmos no pueden predecir.

Todo tendría sabor a juego realmente, si no pensara que la inteligencia artificial es una rama de la ciencia de muy importante porvenir. Una rama que entronca muy íntimamente

con el objetivo para el cual nacieron las computadoras: el uso humano del ser humano.

Trasladando cada vez más lejos la frontera de lo que una computadora puede hacer estamos precisando cada vez más qué es lo puramente del hombre, en el campo del pensamiento. ¿No sería una buena definición, que lo propio del ser humano en el campo del pensamiento, es lo que una computadora no puede hacer?

Frontera móvil que los estudios de inteligencia artificial buscan continuamente correr.

Ojalá que esta publicación origine un vigoroso movimiento en favor de una disciplina tan noble e importante, tal cual es la inteligencia artificial.

SIMÓN PRISTUPIN

PREFACIO

La Inteligencia Artificial es la disciplina que tiene por objeto el estudio de los métodos por los cuales mecanismos artificiales pueden ejecutar tareas consideradas como "inteligentes", que requieren racionalidad en la percepción, comprensión, decisión y acción; entre ellas la utilización del lenguaje natural, la percepción de configuraciones, la competición en diversos juegos y la demostración de teoremas matemáticos.

El estudio se realiza en los aspectos teóricos y experimentales, elaborando teorías y utilizando las fabulosas posibilidades de la computación para llevarlas a la práctica, comprobándolas y adecuándolas de acuerdo a los resultados. Se utilizan para ello los modelos obtenidos de la observación del comportamiento humano, como también los diseñados para resolver problemas en las áreas en que intervienen los procesos intelectuales en estudio.

Puede ubicarse a la inteligencia artificial en la intersección de varias disciplinas: la psicología, las matemáticas, la computación, la ingeniería, la biología, la lingüística y otras. Todas ellas, sin excepción, se benefician con los resultados aquí obtenidos; y recíprocamente, los adelantos en inteligencia artificial son resultado del estudio multidisciplinario y la interacción de los respectivos avances.

Es por ello que el presente libro es de utilidad para el estudioso en cualquiera de dichas áreas, así como para todos aquellos interesados en conocer los principios de este novedoso tema; más aún por el hecho de ser esta la primera obra sobre la materia en idioma castellano.

Debido a ello es que hemos intentado, sin sacrificar el nivel de exposición, hacer esta obra accesible incluso a quienes no han tenido un contacto previo con la computación, la psicología o las ciencias matemáticas; sin que resulte redundante para los expertos en las áreas mencionadas.

La idea de este trabajo surgió a partir de nuestra actuación en GEDECC (Grupo de Estudio y Desarrollo en Ciencias de la Computación), un equipo formado en la Facultad de Ingeniería de la Universidad de Buenos Aires, al vernos afectados por la falta de información en el área.

Nuestro objetivo ha sido el de brindar una visión clara, actualizada, objetiva y bien documentada sobre los temas tratados; a tal fin hemos debido abocarnos durante un largo período a la revisión del disperso material existente en publicaciones extranjeras, generalmente de difícil obtención, y a la comunicación con los principales centros de investigación del mundo, con el objeto de conseguir el material necesario. Debemos aclarar que no hemos encontrado ningún texto de este tipo, por lo que nuestro esfuerzo se ha visto acrecentado al no poder basarnos en otros esquemas de exposición.

En el primer capítulo repasamos brevemente algunas nociones importantes, especialmente para aquellos sin un contacto previo con la teoría de sistemas y la cibernética en general. En el segundo, analizamos someramente los esfuerzos realizados en el campo de la simulación de procesos intelectuales con fines psicológicos y psiquiátricos. El núcleo del libro lo constituye el capítulo 3, donde luego de analizar el área de la inteligencia artificial en general, pasamos a estudiar los diversos problemas que se plantean y las soluciones existentes. En particular revemos los programas capaces de competir en juegos, probar teoremas, conversar y el apasionante terreno de los robots.

No podemos dejar de señalar aquí la influencia que han tenido sobre la exposición los dos brillantes artículos del doctor Minsky, del M.I.T., publicados en 1961 y 1968 respectivamente. En el último capítulo, hacemos algunas observaciones finales y repasamos los últimos adelantos en el área, llegando a la Conferencia Internacional de septiembre de 1971.

En el afán de brindar un panorama amplio, hemos debido restringirnos en la exposición de los intentos particulares, resultando ello en una comprensión quizás excesiva de los resultados, por lo que remitimos al lector interesado a la bibliografía citada en cada caso. Por otra parte, deseamos desde ya excusarnos con todos aquellos investigadores a quienes hayamos podido malinterpretar en la exposición de sus resultados.

Para finalizar, queremos dejar constancia de nuestro agradecimiento al doctor Hugo Scolnik, director del Programa de Matemática de la Fundación Bariloche, por su colaboración en la etapa de documentación, a los miembros del GEDECC que nos motivaron a introducirnos en esta apasionante área (principalmente los ingenieros E. Di Tada y L. Trabb), y al personal

de la Biblioteca y Hemeroteca del Departamento de Matemáticas de la Facultad de Ciencias Exactas y Naturales de la U.B.A. También hacemos extensivo nuestro agradecimiento a los investigadores extranjeros que generosamente nos han provisto de información (entre ellos los doctores M. Minsky, del M.I.T., y R. Banerji, de la Case Western Reserve University).

Por último, nuestro especial agradecimiento es hacia quien, paciente y desinteresadamente, se dedicó a descifrar y pasar en limpio nuestro manuscrito, la señorita Patricia Gordon.

E. D. SONTAG

J. L. TESORO

Capítulo 1.

NOCIONES PRELIMINARES

En este capítulo nos referiremos a ciertos temas que resultan necesarios para el mejor aprovechamiento del resto del libro. No debe tomarse este resumen de definiciones y resultados como un trabajo en las áreas que se mencionen, sino simplemente como una guía para aquellos que conocen dichas materias, indicándoles las partes que se utilizarán. Respecto de quienes no han tenido un contacto previo con la teoría de sistemas, si bien sería útil la lectura de alguno de los usuales textos de cibernética, varios de los cuales se encuentran incluso traducidos al castellano, los autores confían en que aún sin hacerlo, los temas han sido tratados desde un nivel suficientemente elemental como para poder ser seguidos sin dificultad. Sobre la lectura de este capítulo en sí, nuestro consejo es una rápida mirada inicial, y el volver a ellos cada vez que sea necesario posteriormente. En cualquier caso, gran parte de los conceptos sólo se utilizarán en la sección 3.4 ("Métodos generales y enfoques formales"), la cual puede, a su vez, ser salteada sin pérdida de continuidad del texto.

Sobre la bibliografía nos es imposible dar referencias siquiera parcialmente completas, debido a la inmensidad de trabajos publicados. Sin embargo, en los libros que nombremos podrán a su vez encontrarse amplias bibliografías, las que se encuentran principalmente en inglés, francés, alemán y ruso, además de algunas traducciones a nuestra lengua. Antes de entrar en los temas, hagamos notar que debido a una falta de unificación en la traducción de ciertos términos, se vuelven en parte arbitrarios los que nosotros utilizemos, por lo que nos excusamos desde ahora ante quienes estén acostumbrados a distintos vocablos.

1.1. Sistemas, modelos y algoritmos.

Las diversas ciencias han tenido motivaciones, comienzos y desarrollos distintos, desde la antigüedad algunas y en los

últimos decenios otras. Quien compare superficialmente textos de biología, física y sociología, por ejemplo, difícilmente pueda encontrar una sola analogía, una sola idea en común. Sin embargo, despojando a éstas de su vocabulario especializado y concentrándose en los principios fundamentales, será posible advertir puntos de contacto mucho más significativos que simples coincidencias. Es entonces de suponer que un estudio de estos principios en sí mismos pueda llevar a una mayor comprensión de los diversos temas, al concentrarse en lo realmente fundamental y dejando de lado las complicaciones de detalle y forma no esenciales.

Ya durante la Segunda Guerra Mundial un grupo de eminentes científicos en los Estados Unidos, provenientes de las más diversas ciencias, notaron esta afinidad y decidieron el trabajo en conjunto concentrándose en particular en el problema del *control* en todo tipo de organismos, ya fuesen éstos biológicos o técnicos. En su exposición de los resultados obtenidos en aquella época, el líder de estos investigadores, el matemático Norbert Wiener, da una apasionante descripción de la forma en que se fueron encontrando similitudes entre las áreas más diversas; la introducción a su libro "Cybernetics" ("la ciencia del control y la comunicación en el animal y en la máquina")¹ es una lectura obligada para quien desee tener siquiera una idea de la significación de la gran revolución que está sufriendo en estos momentos la humanidad en todas las áreas, a la cual no es ajena esta nueva visión de los procesos, sean éstos biológicos, técnicos, psicológicos, sociales o económicos. Sobre gran parte de estas ideas es también interesante (además de no presuponer conocimientos matemáticos, como la obra de Wiener), un libro de W. Wieser², que en su traducción a nuestro idioma incluye una breve bibliografía de lo publicado en castellano. Más recientemente un libro, también traducido al castellano, incluye varios artículos introductorios³.

La noción básica es la de *sistema*. Un sistema es todo aquel conjunto de objetos dispuestos y *conectados* de algún modo. Es justamente esta propiedad de *conexión* la que permite diferenciar un sistema del simple agregado de sus componentes: el "todo" deja de ser "suma de las partes" debido a la presencia de un nuevo elemento: el intercambio de *información* que se

¹ Wiener, Norbert, *Cybernetics*, segunda edición, M.I.T. Press, 1961.

² Wieser, Wolfgang, *Organismos, estructuras, máquinas*, EUDEBA, 3ª edición, Buenos Aires, diciembre 1970.

³ Otto Walter Haseloff, *Cibernética hoy*, Editorial Tiempo Nuevo, Caracas, Venezuela, marzo 1971.

lleva a cabo. Al estudiarlos deben trazarse bien claramente los límites del sistema en consideración; así, una computadora digital es un sistema, pero la misma máquina en conjunto con un ser humano encargado de su operación forman uno distinto, compuesto a su vez por dos sistemas que actúan como elementos del que los contiene (*subsistemas*).

Los sistemas pueden clasificarse de diversos modos, de acuerdo a su complejidad, a la forma de las conexiones entre los elementos que lo constituyen o a su comportamiento. Es útil su representación mediante diagramas en los cuales los elementos se indican mediante circunferencias o cuadrados (*nodos*) y sus conexiones con flechas; matemáticamente tales objetos se denominan "grafos" (orientados), y podemos distinguir entre aquellos que poseen o no "rótulos" o indicadores del *tipo* de conexión adjuntada a la flecha. Como ejemplo, en la Fig. 1 observamos (i) cómo se podría indicar que un elemento "A" actúa sobre otro "B"; (ii) el hecho de que el segundo a su vez actúa sobre el primero (retroalimentación o "feedback") o (iii) otro en el que dos rótulos indican el camino a recorrer de acuerdo a

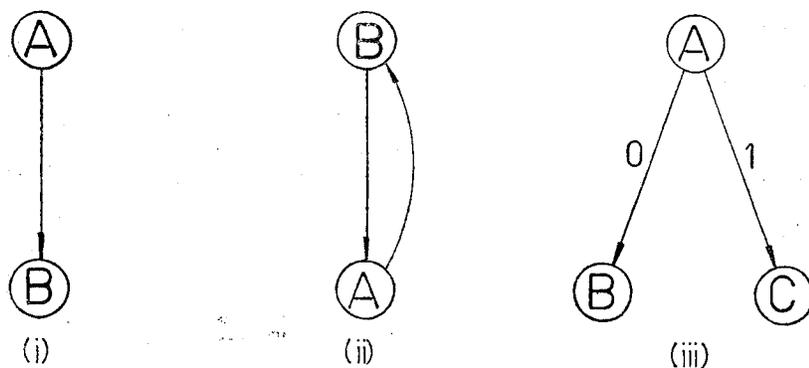


Figura 1.

que un cierto valor sea "0" ó "1". Cuando en un grafo no hay "ciclos" (no se puede volver a un nodo partiendo de él y siguiendo la dirección de las flechas), se acostumbra a denominarlo "árbol" si además es tal que haya un único nodo *inicial* al cual no lleguen flechas y para cualquier otro nodo hay exactamente una flecha. Un árbol *binario* será aquel en que de cada nodo salgan dos conexiones.

Para una interesante lectura sobre los principios de ciber-

nética y sistemas en general, aconsejamos el trabajo de S. Beer ⁴ y el pequeño libro de Greniewsky ⁵, aunque para un estudio algo más formal y matemático, es inmejorable la excelente obra de Ashby ⁶.

En el estudio de un sistema muy complejo, podemos limitarnos a su enfoque como una "caja negra", analizando sus *respuestas* (o *salidas*) ante ciertos *estímulos* (o *entradas*), sin preocuparnos por su constitución interior, pero obteniendo resultados válidos acerca de su comportamiento. Así podemos distinguir entre los sistemas "determinados", en los que a una cierta entrada corresponde siempre la misma salida (el interruptor de un aparato eléctrico, por ejemplo) o no determinados ("probabilísticos", etcétera), donde si bien no es posible predecir la respuesta con seguridad, en ocasiones se puede dar una *distribución de la probabilidad* para éstas (en el caso del sistema "hombre + moneda", cada vez que el sujeto arroje la moneda no sabrá con seguridad si caerá "cara" o "ceca", pero si la moneda está bien equilibrada, la posibilidad será de "1/2" para cada una, lo que significa que si la experiencia se repite gran cantidad de veces, la frecuencia de aparición de cada lado tenderá hacia este valor).

Es claro que no nos interesa, al hablar de un sistema, la forma física de éste, sino sólo su aspecto funcional. Cuando entre dos sistemas existe una correspondencia unívoca entre sus elementos, de modo tal que si una relación se verifica entre dos componentes de uno, entonces también se la encuentra entre los correspondientes en el otro, se dice que existe un "isomorfismo" entre ambos, o que son "isomorfos". Por ejemplo, serían isomorfos un avión supersónico, una maqueta en madera suficientemente detallada como las que se utilizan en las pruebas de diseño, y los planos y ecuaciones matemáticas que lo describen.

Esta correspondencia puede hacerse menos fuerte, pidiendo ahora no un paralelo exacto de ambos sistemas, sino sólo el que uno de ellos sea el resultado de *simplificar* el primero y se dice en este caso que se tiene un *homomorfismo*; se debe notar que en el segundo no pueden aparecer *relaciones nuevas* entre sus elementos, sino que sólo se producirá un fenómeno similar a la observación de un conjunto de piezas (un *subsistema*) como un único bloque o elemento en el nuevo sistema. Por ejemplo, el

⁴ Beer, Stafford, *Cibernética y administración*, Compañía Editorial Continental S.A., México, 1965.

⁵ Greniewsky, H., *Cibernética sin matemáticas*, Fondo de Cultura Económica, México, 1965.

⁶ Ashby, W. Ross, *Introducción a la cibernética*, Editorial Nueva Visión, Buenos Aires, 1960.

segundo sistema de la Fig. 2 es resultado de aplicar un homomorfismo al primero, ya que el elemento "I" consiste en la identificación del subsistema "A, B, C", conservando sus relaciones con el resto. El concepto de homomorfismo se puede aplicar, alternativamente, sólo al comportamiento del sistema en

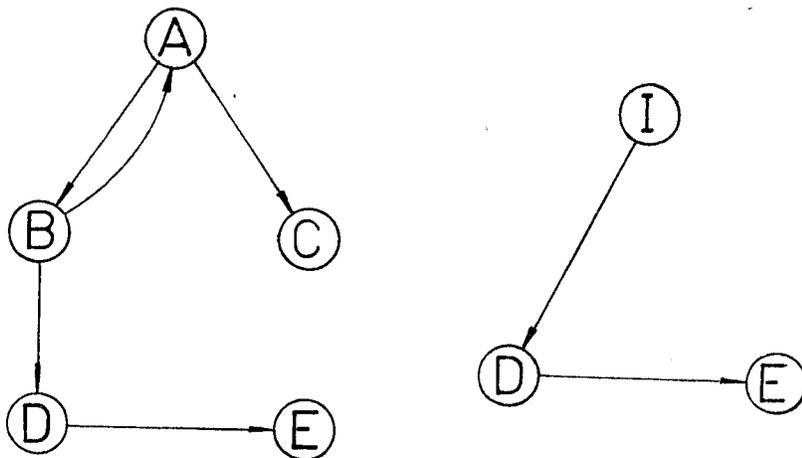


Figura 2.

lo que hace a la "E/S" (entrada-salida), pero no es éste el lugar para ahondar en estos temas, que pueden encontrarse en la literatura citada.

Un *modelo* es una simplificación, en el sentido del párrafo anterior, de un cierto sistema. Así, un modelo podrá darnos información sobre propiedades que nos interesan de un sistema muy complicado, evitándonos la necesidad de una experimentación directa. Un *modelo matemático* consiste en un conjunto de *axiomas* y *definiciones* a partir de los cuales se obtienen deducciones lógicas; en estos axiomas es donde debemos representar las propiedades que nos interesan de un sistema dado, y los *teoremas* que se obtengan tendrán validez siempre que dicha representación inicial haya sido correcta. Las estructuras matemáticas usuales (semigrupos, grupos, cuerpos, etcétera) se definieron justamente como abstracción de las propiedades de modelos en diversas áreas.

Ciertos sistemas se pueden representar con algunos modelos matemáticos muy simples en lo que se refiere a su E/S. Ante todo, se debe notar que un sistema determinado actúa como una *función* entre el conjunto de entradas y el de salida (o sea, re-

presenta una correspondencia tal que a cada elemento del primer conjunto asigna unívocamente un elemento del segundo). Cuando tanto los datos de entrada como los de salida pueden tomar sólo dos valores, que simbolizamos "0" y "1" (o, equivalentemente, "no" y "sí", "negativo" y "positivo", "no excitación" y "excitación", "falso" y "verdadero", etcétera), diremos que se trata de un sistema "binario". Los datos se pueden combinar en "vectores" o listas de números separados con comas y entre paréntesis; así por ejemplo una serie de informaciones sobre mediciones de dos aparatos podría presentarse como vectores del tipo "(10, 275)", donde la primera posición, o *coordenada*, indica uno de los instrumentos y la segunda el otro.

Entre los sistemas binarios más utilizados se encuentran los de negación, representados por la función que a un "1" asigna un "0" y viceversa, de conjunción o producto lógico, de dis-



Figura 3.

yuncción inclusiva, etcétera. Los dos últimos actúan sobre una entrada compuesta por dos elementos, dando una salida de acuerdo a las tablas de la Fig. 3, y se acostumbra a diagramarlos como en los dibujos respectivos. En cambio tiene una entrada y dos salidas (o una única en forma de *vector* de dos componentes) el *duplicador*, que simplemente repite la señal recibida, enviándola en dos direcciones distintas.

Generalmente interesa un estudio *dinámico* del sistema dividiendo el tiempo en intervalos, siendo entonces importantes los *retardadores*, que emiten en un instante dado la señal recibida en el anterior. Tales sistemas se pueden describir con *ecuaciones en diferencias* como $A_t = B_{t-1} \cdot C_t$, que indica que "el valor del elemento A en el instante 't' se puede obtener como el producto lógico del de B en el instante anterior con el de C en el actual", ecuación que es modelo matemático del sistema diagramado en la Fig. 4.

Una función que podamos describir de un modo efectivo será un *algoritmo*. Intuitivamente, un algoritmo es un conjunto de instrucciones que permitan obtener una solución a un cierto problema, dados ciertos datos de entrada, y en un número finito

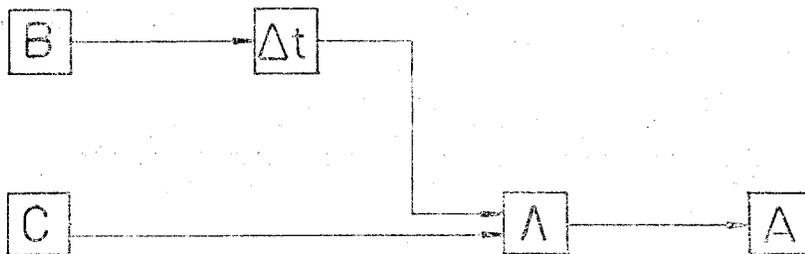


Figura 4.

de pasos. Una de las áreas más apasionantes de la lógica matemática actual está en el estudio de la *existencia* de algoritmos, y mediante una *definición* adecuada de lo que se entiende por "conjunto de instrucciones que permitan..." etcétera, se puede formar un *modelo matemático* del concepto de algoritmo, llegándose a demostrar como teoremas la *imposibilidad* de obtener algoritmos para la resolución de ciertos problemas.

Entre los modelos propuestos para formalizar esta noción se encuentran los esquemas de Church, Post, Turing, y otros; en particular las *funciones recursivas* y la máquina de Turing, a la cual volveremos en la sección siguiente. Se ha demostrado la *equivalencia* entre estos modelos (en el sentido de que un problema es "decidible" en uno si y solo si lo es en el otro), y la evidencia ha llevado a postular el principio de que el concepto intuitivo de algoritmo es equivalente a cualquiera de éstos. Se puede encontrar una explicación bastante digerible de estos temas en el primer capítulo del trabajo de Glushkov⁷, y un estudio muy completo, pero más complejo, en el conocido libro de Davis⁸.

1.2. Algunos modelos particulares. Computadoras digitales.

En la sección anterior nombramos a la máquina de Turing como uno de los modelos que se han propuesto con el fin de formalizar la noción intuitiva de algoritmo. Este modelo puede interpretarse como constituido por una "máquina" que puede inspeccionar una cinta dividida en "celdas", cada una de las

⁷ Glushkov, Viktor M., *Introduction to cybernetics*, Academic Press, 2ª edición, 1968.

⁸ Davis, Martin, *Computability and Unsolvability*, McGraw Hill, 1958.

cuales se encuentra en blanco o contiene un cierto signo. Se supone que esta máquina puede encontrarse en uno de varios "estados", y su comportamiento en un instante dado está determinado por su estado actual y por el signo que se encuentra en una posición dada de cinta (la que se está "observando"); sus acciones pueden consistir en borrar o escribir un símbolo en dicha posición, cambiar de estado o "mover" la cinta en alguna de las dos direcciones con el fin de observar otra celda. Se supone finito el número de estados posibles y la cantidad de casilleros o celdas que contienen algún símbolo en cada instante. Entre los estados posibles podemos suponer aquel de "detenido" en el que terminó su trabajo.

Formalmente, se describe la máquina mediante una función que indica las acciones a tomar ante cada estado y cada símbolo observado. Recordando lo enunciado sobre algoritmos, digamos que la máquina de Turing, aun con su descripción tan sencilla, es capaz de realizar cualquier operación definible. Inclusive es, en principio, más potente que cualquier máquina existente, debido a que es un dispositivo *infinito*: aunque en cada instante la cinta tiene una longitud "ocupada" finita, se supone que se extiende (incluyendo las celdas en blanco) en las dos direcciones hasta el infinito. Debemos siempre recordar que la máquina de Turing es un modelo matemático, y no un aparato electrónico; sin embargo, es útil estudiar muchas de las propiedades de estos últimos mediante las primeras, en razón de que el modelo es extremadamente simple y se presta a un adecuado estudio teórico. Por otra parte, en virtud del principio antes señalado, todo cálculo que no pueda hacer ninguna máquina de Turing (dada una *codificación* adecuada en términos de los símbolos en la cinta), menos aún podrá hacerse con otro método cualquiera.

Para modelar los diversos sistemas discretos se acostumbra a introducir los "autómatas finitos", "máquinas secuenciales", etcétera, y en general las "máquinas de finitos estados"^{1, 2, 3, 4}. Se trata de suponer máquinas en las que se reconocen un cierto número de estados y que pueden ser definidas por dos funciones, una de las cuales ("función de transición") indica el estado al que se pasará con una entrada dada y la otra (de "salida") el

¹ Gill, Arthur, *Introduction to the theory of Finite-State Machines*, McGraw Hill, 1962.

² Ginsburg, Seymour, *An Introduction to Mathematical Machine Theory*, Addison-Wesley, 1962.

³ Harrison, M. A., *Introduction to Switching and Automata Theory*, McGraw Hill, 1965.

⁴ Di Tada, Esteban, *Introducción a la teoría de autómatas y máquinas secuenciales*, Apuntes de Curso, Universidad de Buenos Aires, 1971.

símbolo a emitirse como respuesta (máquinas secuenciales), o que sólo poseen la función de transición (autómatas finitos). En el último caso se supone que se tienen ciertos estados distinguidos o "finales", y un estado "inicial"; si una entrada lleva a un estado final partiendo del inicial, se dice que el autómata la "aceptó".

Por ejemplo, el grafo de la Fig. 5 representa un autómata de cuatro estados (A, B, C, D), donde las transiciones están indicadas por los rótulos de las flechas (así, estando en B, si recibe un "0" pasa a C, y si recibe un "1" a D). Si suponemos que el estado inicial es A y que el único final es C, el lector podrá verificar que las entradas "aceptadas" son exactamente aquellas con una

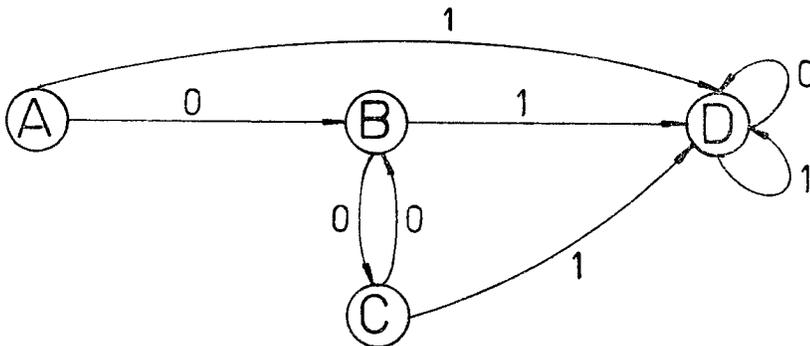


Figura 5.

cantidad par de ceros y ningún "1". Por ejemplo, "0000" produce la siguiente secuencia de estados a partir de A: B, C, B, C; en cambio, "001" da: B, C, D, no siendo D final, por lo que no es aceptada.

La ventaja del trabajo con "máquinas de finitos estados" está en la existencia de numerosos algoritmos de simplificación. Uno de ellos nos diría, por ejemplo, que el mismo trabajo del autómata anterior lo puede realizar el de la Fig. 6, que posee

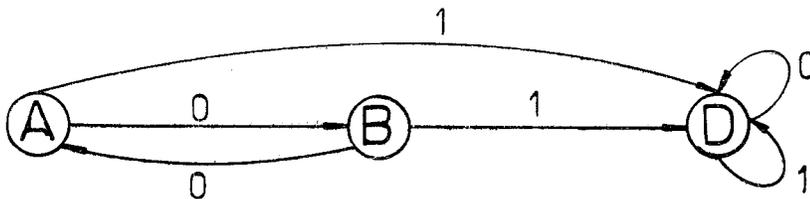


Figura 6.

un estado menos (comprobarlo con algunos ejemplos). Esto es particularmente importante en el diseño de sistemas en los que cada componente adicional aumenta el costo.

Cuando un modelo de este tipo representa un programa de computadora, o sencillamente un proceso cualquiera a seguir sobre un cierto alfabeto de entrada, se pueden obtener interesantes resultados sobre la posibilidad de análisis de un lenguaje con un método dado. Para este fin se han creado otros tipos de autómatas, como los "push-down", en los que se supone una "memoria auxiliar", definiéndose adecuadamente los conceptos de "aceptación", etcétera ⁵.

Al final de la sección anterior habíamos visto sistemas capaces de realizar operaciones lógicas simples. Estos dispositivos pueden ser fácilmente implementados en circuitos electrónicos, y su adecuada interconexión permite efectuar cálculos mucho más complicados. Junto con adecuados sistemas físicos de *almacenamiento* de la información, tenemos los medios técnicos necesarios para la construcción de una computadora digital.

Una *computadora digital* es un dispositivo capaz de ejecutar cualquier proceso definible mediante una secuencia precisa de instrucciones básicas (*programa*), utilizando, si es necesario, cierta adicional "información de entrada", y produciendo una salida adecuada. Se compone básicamente de una unidad de control, encargada de extraer las instrucciones e información auxiliar (*datos*) de un almacenamiento (*memoria*), de entender o *decodificar* las instrucciones y de ordenar su ejecución a las restantes unidades (de E/S y aritmético-lógicas). Las computadoras actuales tienen tanto los dispositivos de control como de memoria, electrónicos.

En particular, se puede programar una computadora de modo tal de comportarse como cualquier autómata finito, y aun en la práctica, como una máquina de Turing, ya que con el agregado de memorias auxiliares se puede simular la existencia de una "cinta infinita". De lo anterior deducimos que, en principio, sería posible ejecutar cualquier algoritmo.

Ante la complejidad de la programación en el lenguaje propio de la máquina, se han diseñado diversos "lenguajes de alto nivel" orientados hacia diversos problemas, los cuales son "traducidos" por la computadora a su propio código para luego ejecutarlos (*compilación*). En particular, los investigadores en *inteligencia artificial* utilizan usualmente el LISP u otros "lenguajes procesadores de listas". Otro notable aumento de efi-

⁵ Al respecto, consultar en general los diversos números de la publicación *Information and Control*, y Bibliografía sobre Teoría de lenguajes.

ciencia se logra al dividir un programa en uno principal y varias *subrutinas*, que son llamadas en el momento en que se deba realizar un cierto cálculo particular.

1.3. Los lenguajes y las máquinas.

La lingüística, la “ciencia del lenguaje”, es en sí anterior a los adelantos científicos de las últimas décadas. El estudio filológico e histórico del lenguaje tuvo sus comienzos y desarrollo junto con otras ciencias tradicionales, y es así como la filología y la etimología fueron materia de estudio.

Pero la gran revolución en los conocimientos no podía dejar de alcanzar a la lingüística. La traducción automática, los grandes bancos de datos para centralizar y distribuir la increíble avalancha de información científico-técnica que ya comenzamos a recibir, la necesidad imperiosa de medios más eficientes de comunicación entre el hombre y la máquina y, en fin, una innumerable cantidad de interrogantes planteados por la “Era de la Información”, hicieron necesario un estudio formal de los lenguajes, tanto los *naturales* (castellano, inglés, etcétera), como los artificiales, usados en computación, y los lenguajes de la lógica matemática contemporánea.

Del mismo modo como se planteó la necesidad de modelos teóricos de los sistemas físicos, lo cual dio lugar a las teorías de autómatas finitos, “push-down”, etcétera, así también para responder a preguntas tales como “¿qué es y cómo se genera un lenguaje?”, “¿qué características debe tener un autómata (o, lo que es *equivalente*, un *programa* en una cierta computadora) para poder establecer una comunicación efectiva en un tipo dado de lenguajes?”, o más generalmente, “¿qué relaciones hay entre las diversas categorías de lenguajes y los modelos de autómatas?”, obligó a un estudio matemático-lógico de las estructuras en cuestión. El complemento del análisis experimental y estadístico fue por supuesto necesario al ocuparse de los lenguajes naturales.

Formalmente, se define un “lenguaje” como una parte (un *subconjunto*) de todas las posibles *cadena*s finitas de símbolos (*palabras*) formadas a partir de un conjunto de signos o “letras” (*alfabeto*). Por ejemplo, si el alfabeto es $A = [a, b, c]$, serán palabras “aba”, “ba”, etcétera, y podríamos definir un lenguaje como el consistente sólo de la palabra “aac”, o como el formado por todas las palabras que consistan de una cadena de “a’s” seguida por otra de igual longitud formada por “b’s”: $[a^n b^n, n \in \mathbb{N}]$; finalmente, podemos definir el “español” como el subconjunto de las palabras formadas sobre el alfabeto latino constituido por

las que aparecen en el último Diccionario de la Real Academia. Por razones de conveniencia se considera en todo lenguaje la existencia de una "palabra vacía" (tal que si la notamos " Λ ", y si " x " es una letra, son equivalentes " $x\Lambda$ ", " Λx ", o " x "), resultando que con la operación de *concatenación* (colocar una letra al lado de otra), el conjunto de palabras sobre el alfabeto dado es el *semigrupo libre* generado por éste.

Los *lenguajes regulares* (o "conjuntos regulares") son los que pueden ser "aceptados" o reconocidos por un autómata finito, y han sido muy estudiados en cuanto a sus propiedades matemáticas. Son, sin embargo, demasiado simples para poder ser aplicados sus resultados a los lenguajes naturales y a los complejos lenguajes de alto nivel que utilizan las computadoras actuales.

Debemos entonces, por una parte, limitar los modelos de tal modo que permitan obtener conclusiones particulares, pero por otra permitirles ser suficientemente generales como para abarcar los diversos casos. Antes de continuar hagamos notar que simplemente denominando "letras" a las palabras del castellano, y agregando el espacio y los signos de puntuación también como letras, el conjunto de todas las posibles sentencias podrá ser descrito como un lenguaje; así podríamos decir que "El estudia lingüística" es una palabra formada por las "letras" *él*, *estudia*, *lingüística* y el espacio en blanco, sin preocuparnos por su estructura interna (del mismo modo, "portaequipaje" será un símbolo distinto y no necesariamente relacionado en modo alguno con "porta" o con "equipaje").

1.3.1. Estructura de frase; ambigüedades.

Una *gramática* es una descripción de la estructura de un lenguaje. El hecho de que ésta deba ser finita no implica necesariamente que el lenguaje obtenido debe estar así restringido: con las 7 palabras "todas las cadenas de ceros y unos" estamos *indicando* una cantidad infinita de palabras. Nos interesan fundamentalmente las *gramáticas generativas* en las que se parte de un "símbolo inicial" y mediante la aplicación de ciertas reglas o "producciones" se obtienen todas las posibles expresiones *sintácticamente válidas*. De este modo llegamos a las importantes gramáticas con *estructura de frase*, estudiadas en un principio por Chomsky¹, basadas en la idea de que toda sentencia,

¹ Chomsky, N., *On certain formal properties of grammars*, *Information and Control*, 1959.

aún las más complicadas, puede ser analizada en función de sus "frases" constituyentes.

Las "producciones" indican los "reemplazos" permitidos; se trata de reglas del tipo "libres del contexto":

$$a \rightarrow b$$

que indica que cada vez que se tenga el símbolo "a", éste puede ser reemplazado por "b", o del tipo "sensitivas al contexto":

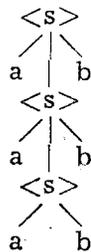
$$x a y \rightarrow x b y$$

señalando que "a" se puede cambiar por "b" sólo si se encuentra entre "x" e "y". Algunos símbolos (los que no pertenecen al lenguaje, sino que cumplen exclusivamente funciones auxiliares) son denominados "no terminales", en contraposición con los símbolos *terminales*, los cuales no son reemplazados una vez que aparecen. Para distinguirlos, se acostumbra a encerrar los símbolos no terminales entre los signos "<" y ">"; el símbolo *inicial*, que denotaremos "S" (*sentencia*) es en particular no terminal. Cada regla se puede aplicar opcionalmente cada vez que sea permitido, y en el orden en que se desee; cuando sólo queden símbolos terminales, termina el proceso de *derivación* de la *sentencia*.

Por ejemplo, si el alfabeto es [a, b], la gramática:

1. $\langle S \rangle \rightarrow a \langle S \rangle b$
2. $\langle S \rangle \rightarrow ab$

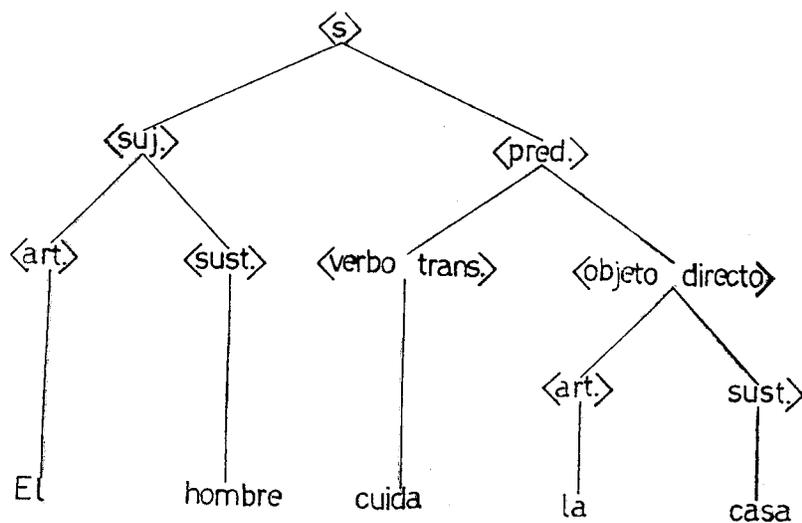
genera el conjunto de todas las cadenas de "a's" seguidas por igual cantidad de "b's"; si cada paso de la derivación lo indicamos con " \Rightarrow ", podemos indicar el modo de obtener "aaabbb" mediante la secuencia de derivación: " $\langle S \rangle \Rightarrow a \langle s \rangle b \Rightarrow aa \langle s \rangle bb \Rightarrow aaabbb$ ", donde hemos aplicado dos veces la regla (1) y una vez la (2). Otro modo de indicar cómo se obtuvo la *sentencia* es mediante el uso de paréntesis: (a(a(ab)b)b) o un *árbol*:



El siguiente ejemplo de gramática es quizá más conocido:

1. <s> → <suje<to>> <predicado>
2. <suje<to>> → <artículo> <sustantivo>
3. <predicado> → <verbo intransitivo>
4. <predicado> → <verbo transitivo> <objeto directo>
5. <objeto directo> → <artículo> <sustantivo>
6. <artículo> → el
7. <artículo> → la
8. <sustantivo> → hombre
9. <sustantivo> → perro
10. <sustantivo> → casa
11. <verbo intransitivo> → vive
12. <verbo transitivo> → cuida

de aquí podemos obtener “El hombre cuida la casa” aplicando (1), (2), (6), (8), (4), (5), (12), (7) y (10), en ese orden, lo cual podemos simbolizar así:



Derivación de “El hombre cuida la casa”.

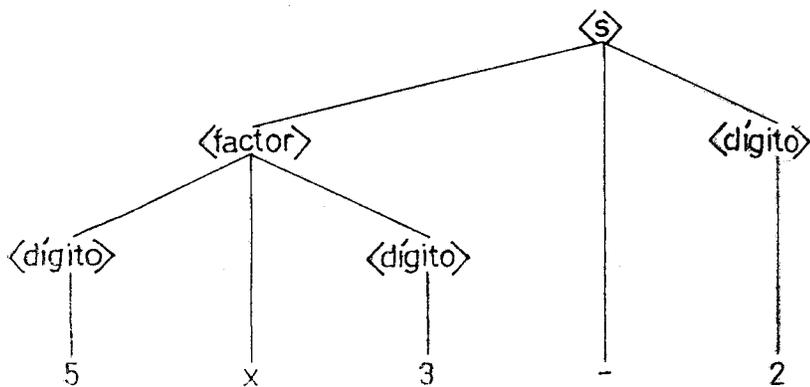
Claro está que también se pueden obtener "la casa vive", "la casa cuida la casa" y otras sentencias que pudieran no tener sentido, pero lo que nos interesa por el momento es la construcción de sentencias *sintácticamente* bien formadas, sin preocuparnos por su significado; para evitar esas contradicciones es que se utilizan las gramáticas dependientes o "sensitivas" al contexto; para ello deberá existir una regla, por ejemplo, que indique que "vive" sólo se puede usar si el sustantivo es "hombre" o "perro", etcétera.

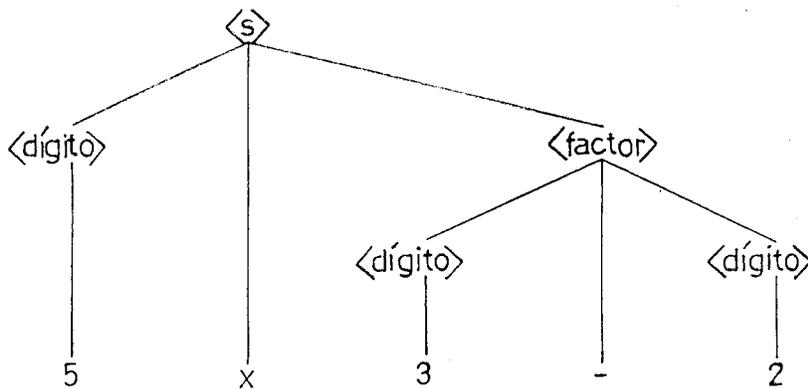
Para poder comprender una oración se deberá entonces poder realizar de algún modo el "parsing" o *análisis sintáctico*, que nos permita percibir su sentido, determinando el sujeto, la acción que lleva a cabo y el objeto de dicha acción. En otros casos, el análisis sintáctico nos permite decidir las operaciones a efectuar sobre una entrada; por ejemplo, si tenemos: " $5 \times 3 - 2$ ", un análisis que indique que la forma como fue obtenida esta "sentencia" es la siguiente:

$\langle s \rangle \Rightarrow \langle \text{factor} \rangle - \langle \text{dígito} \rangle \Rightarrow \langle \text{factor} \rangle - 2 \Rightarrow$
 $\langle s \rangle \Rightarrow \langle \text{dígito} \rangle \times \langle \text{dígito} \rangle - 2 \Rightarrow$
 $\Rightarrow \quad 5 \quad \times \quad 3 \quad - 2$

indicado también así:

significará que se debe primero multiplicar 5×3 y luego restar 2 al resultado; en cambio





indicará $5 \times (3 - 2) = 5 \times 1 = 5$. Mientras que acá se elimina la segunda interpretación mediante las “reglas de precedencia” de los operadores aritméticos, fenómenos análogos de *ambigüedad* se verifican en los lenguajes naturales. Estos se pueden deber a ambigüedades sintácticas (dos análisis distintos posibles) o *semánticas*, cuando una palabra tiene varios significados posibles.

El estudio de los métodos de análisis sintáctico es una de las áreas más importantes de la teoría de lenguajes y de las más atacadas en la actualidad por los teóricos de la Ciencia de la Computación; los problemas de ambigüedades los analizaremos en parte, relacionándolos con las máquinas que se comunican en lenguajes naturales, uno de los temas de mayor preponderancia dentro del campo de la inteligencia artificial.

Se demuestra matemáticamente que un lenguaje es independiente del contexto (definiendo este concepto formalmente, cosa que *no* hemos hecho) si y solo si existe un autómata “push-down” que lo “acepte”. Hay gran cantidad de caracterizaciones semejantes para los diversos tipos de gramáticas y autómatas². El “significado” de una sentencia se puede definir como una *función* que aplica las palabras en las “proposiciones” lógicas o “conceptos”³; aunque ésta no es en principio una definición muy útil operativamente, veremos más adelante cómo se la puede llegar a aproximar en la práctica.

² Ver, por ejemplo, Harrison, M. A., *On the relation between grammars and automata*, Berkeley, 1970; también los diversos números de *Information and Control* y otras publicaciones especializadas.

³ Este y otros temas como los anteriores, por ejemplo, en: Solomon Marcus, *Algebraic Linguistics, analytical models* (Rumania), publicado por Academic Press en 1967, así como en la bibliografía citada.

El análisis sintáctico se puede realizar simplemente como un proceso inverso al de la generación de las sentencias, identificando las palabras o grupos de palabras que aparezcan también a la *derecha* de las producciones y reemplazándolas por el símbolo de la izquierda, repitiendo el proceso hasta llegar al símbolo inicial "S", momento en el cual ya tendremos formado un árbol que nos indicará la estructura de la sentencia en cuestión. Claro que acá tendremos en muchos casos el problema ya mencionado de las ambigüedades: en "yo quiero una vaca que pueda comer" no tenemos modo de saber en principio si lo deseado es una vaca hambrienta o tierna (si "que pueda comer" es modificador directo de "vaca", o se refiere al sujeto "yo"), a menos que estudiemos la situación en que fue emitido el juicio. . Esta situación se presenta a menudo en presencia de pronombres y/o sujetos tácitos. Volveremos a estos problemas en 3.3.

1.3.2. Gramática transformacional.

Las gramáticas con estructura de frase, especialmente las independientes del contexto, han sido muy estudiadas en relación con los lenguajes de programación, desarrollándose algoritmos óptimos de análisis sintáctico, etcétera. Para reducir la posibilidad de ambigüedad, se efectúa el "parsing" en un sentido determinado (de izquierda a derecha de la oración, por ejemplo) lo cual también se corresponde en cierto modo con la manera en que los seres humanos percibimos el lenguaje: si leemos o escuchamos "el", por ejemplo, esperamos recibir un sustantivo, el cual, una vez aparecido, puede completar un "sujeto", con lo que interpretamos lo siguiente como un predicado; si en algún momento se produce alguna contradicción, intentamos "entender" la oración de algún modo alternativo.

Una idea de la importancia de este tipo de gramáticas y lenguajes está dado por el lenguaje algorítmico internacional ALGOL-60 que, descrito por una gramática libre del contexto y sin ambigüedades *, puede ser compilado por las computadoras a una gran velocidad mediante un rápido análisis sintáctico, lo cual se traduce en una optimización en el uso de los sistemas de procesamiento.

Sin embargo, para el estudio de los lenguajes naturales, la estructura de frase resulta inadecuada, debido a la complejidad y a las irregularidades que éstos presentan. Es imposible obtener una gramática del tipo indicado para dichos lenguajes, de-

* Salvo contadas excepciones, modificadas en el ALGOL-68.

bido a las características señaladas, lo cual obliga a buscar otras formas de explicitar su estructura.

El medio aparentemente más útil para lograrlo es mediante la introducción de las "transformaciones", como *complemento* de la acción de una gramática de las anteriores. Estas transformaciones ("obligatorias" u "optativas") son del tipo "A.B. → A y B." (para formar oraciones compuestas), "A escribe B → B es escrito por A" (para pasar a voz pasiva), etcétera.

De este modo, se admite la existencia de dos estructuras: la "profunda", que es un lenguaje generado por una gramática con estructura de frase, y la "superficial", que consiste de todas las variaciones y giros idiomáticos posibles, obtenidos a partir de las anteriores. O sea: se producen las sentencias "núcleo" (*kernel*), que son en general de un tipo sencillo, y luego son modificadas. Con esta teoría, el "entendimiento" del lenguaje se operaría mediante la aplicación de las "transformaciones inversas", realizando luego un análisis sintáctico sobre un lenguaje mucho más simple.

En cierto modo casi todos los programas efectivos de comprensión del lenguaje natural se basan en esquemas similares, y el "PROGRAMMAR" de T. Winograd (ver cap. 3), está encajado en la gramática "sistemática"⁴, una variación que tiene en cuenta "características" útiles para guiar el procedimiento de análisis.

El agregado del "segundo nivel lingüístico" constituido por las transformaciones elimina muchas de las dificultades que se presentan con el único nivel de las gramáticas de frases, como la imposibilidad de derivar *constituyentes discontinuos* (inversión del sujeto y verbo en interrogaciones, por ejemplo), duplicación excesiva de reglas (unas para las sentencias con sujeto en singular, otras en plural, etc.), dificultad para formar sentencias complejas y no demostrar explícitamente la relación entre sentencias con igual significado. Puede suponerse también un último nivel, el "fonético", diferenciado del segundo, del que se obtienen los "fonemas" a emitir.

Todas estas teorías están en un estado experimental, pero ya se han obtenido resultados relativamente buenos con los esquemas señalados. En 3.3 los estudiaremos en relación con la inteligencia artificial y como ejemplos de la utilidad de buenas estructuras (modelos) interiores en la comprensión del lenguaje.

⁴ En inglés, "systemic"; podría quizá traducirse más correctamente como "sistemática".

Capítulo 2. SIMULACION DE PROCESOS INTELECTUALES HUMANOS

La simulación de procesos intelectuales humanos por computador constituye una de las áreas de mayor espectacularidad y generadora de complejas y problemáticas discusiones relacionadas con la lógica, epistemología, ética y psicología.

Siempre se ha creído y aún se cree, en parte con razón, que la inmensa complejidad de los procesos intelectuales limita las posibilidades de estudio analítico y los hace inaccesibles a todo intento de cuantificación, estudio sistemático y representación analógica mediante modelos físicos o formales y su simulación en mecanismos automáticos deterministas.

El problema reside en que aunque tal complejidad resulta evidente, y que aumenta a medida que aumenta el grado de penetración, nadie puede probar que el diseño y operación de modelos representativos de tales procesos sea una tarea imposible de llevar a cabo.

La auténtica esencia de los procesos intelectuales se oculta bajo un conjunto de apariencias; el proceso pensante permanece aún subyacente bajo la masa de nociones psicológicas referentes a las actividades intelectuales. A la pregunta: ¿Qué es el pensamiento? se responde describiendo un conjunto de mecanismos psíquicos que tienen lugar cuando el hombre se ocupa en la compleja función pensante. Estos mecanismos psíquicos o subfunciones asociadas al proceso pensante son la percepción, la abstracción, la comparación, el juicio, la generalización, la inferencia, etcétera; indudablemente estos mecanismos interactúan recursivamente en el proceso global; sin estas capacidades el hombre no podría ocuparse en ningún tipo de actividad intelectual, pero este conjunto de mecanismos interactuantes no constituyen el pensamiento sino que son complementarios de la actividad central.

El pensamiento es un mecanismo que el hombre ejerce por algún motivo y para el logro de algún objetivo; no es algo que en él *acontece* sino que para ejercerlo moviliza todas sus actividades intelectuales para llegar a un resultado, a la finalidad per-

seguida. Esta finalidad propuesta define más rigurosamente el pensamiento que el conjunto de mecanismos coordinados que actúan para su consecución.

Mediante una noción vaga y superficial de la esencia del pensamiento la psicología ha podido llegar a clasificar ciertos fenómenos psíquicos como predominantemente *intelectuales*; esta denominación común se debe a que estos mecanismos *intervienen* en el pensamiento; pero en realidad no constituyen la esencia del proceso pensante, sino que brindan la mera posibilidad instrumental para que el proceso central pueda ejercerse. Allí reside la principal restricción en el aporte de la psicología al conocimiento de la naturaleza del proceso pensante.

La lógica también efectúa su aporte fragmentario destacando y sistematizando sólo ciertos y determinados esquemas formales del pensar, ciertas estructuras de razonamiento *válidas* que son las que ostentan los presuntos atributos *lógicos*; constituyendo otra ocultación en la búsqueda de la naturaleza esencial del proceso pensante. La inducción y la deducción son métodos indirectos para obtener proposiciones verdaderas, pero mediante su aplicación se obtienen verdades *deducidas* o *inducidas* cuya validez se cimenta en leyes lógicas.

Estos aportes fragmentarios han enriquecido la construcción teórica e incentivado la investigación; todo lo verdaderamente estudiado permanece equívoco y esto no constituye una deficiencia de la univocidad lógico formal a la que debería tenderse; sino que la equivocidad es el elemento en el que necesita moverse la investigación. Por esto es necesario analizar el proceso pensante en medio de su equivocidad sin que sea necesario arribar a un resultado palpable y terminante, estamos en el dominio de lo inagotable por su esencia, que es la característica de lo creativo. Estamos adjudicando aspectos y apariencias para profundizar el análisis, los vacíos contenidos en los aportes parciales son el material de trabajo, lo no analizado es el aporte más fructífero que un análisis tiene para ofrecer.

¿Qué significa pensar? La pregunta aparece como unívoca pero resulta ser equívoca, y esto frustra toda tentativa de dirigirse en línea recta y sin más hacia una respuesta. Es necesario esclarecer el múltiple significado de la pregunta ya que detrás de ella se ocultan varias vías por las que puede discutirse. Vamos a enumerarlas: a) ¿qué significa la palabra "pensar"? ¿qué se designa con el nombre de "pensar"?; b) ¿Cómo se concibe y define el mencionado pensar en la doctrina tradicional sobre el pensar? ¿En qué se hace consistir desde hace dos milenios y medio, el rasgo fundamental del pensar? ¿Por qué la doctrina tradicional del pensar lleva el nombre de *lógica*?; c) ¿Qué se

requiere para estar nosotros en condiciones de pensar de manera correcta? ¿Qué se exige para llevar a cabo en cada caso el proceso pensante?; d) ¿Qué es lo que nos *ordena* pensar? ¿Qué es lo que nos convoca a pensar?

Podemos formular la pregunta según estas cuatro maneras, tratando de aproximarnos a una respuesta mediante las discusiones correspondientes; estas maneras de formularla están conectadas entre sí, lo inquietante de la pregunta no radica tanto en la equivocidad cuanto en la univocidad que están designando esas cuatro maneras por igual. Se trata de reflexionar sobre si una sola de las cuatro es válida, resultando las restantes como inconsistentes y posteriormente añadidas, o si todas son igualmente necesarias por estar unificadas entre sí. Quizá la unificación, el nexo entre las distintas maneras de formular la pregunta se agrega como un quinto factor a la pluralidad de las cuatro maneras de preguntar. ¿Hay entre las cuatro una que por su jerarquía domina a las restantes?, ¿existe una estructura jerárquica conforme a la cual las cuatro maneras de preguntar se ensambлан subordinándose a aquella primaria que da la pauta?

Las cuatro maneras enumeradas de formular la pregunta “¿qué significa pensar?” no están separadas y unidas por mera yuxtaposición, sino que guardan conexión entre sí en virtud de una unidad cuya estructura procede de una de estas cuatro maneras: ¿qué es aquello que nos dirige hacia el pensar y determina el proceso? La pregunta va en busca de *aquello* que nos imprime cada vez el impulso a pensar, dada la ocasión y en orden a un objeto determinado.

El pensar aparece como un tema que puede tratarse como tantos otros, se hace un objeto de la investigación, centrado en un proceso que tiene lugar en el hombre. El hombre tiene una participación especial en el proceso de pensar por cuanto es el que lo ejecuta, pero este hecho no importa mayormente a la misma investigación del pensar. Es un hecho que se sobrentiende y que puede permanecer al margen de la consideración sobre el pensar como algo indiferente; más aún, es necesario que así sea, porque las leyes a que obedece el pensar son independientes del hombre que ejecuta en cada caso el proceso pensante.

Si por el contrario la pregunta “¿Qué significa pensar?” nos dirige hacia el pensar, estamos preguntando por algo que nos toca a nosotros mismos por cuanto apunta a nuestra esencia poniéndonos a nosotros mismos en tela de juicio, no ataca solamente un objeto, como ocurre tratándose de un problema científico.

La pregunta expresada en cuarto lugar es la que da la pauta cuando el pensar se vuelve cuestionado en su esencia; esto no

quiere decir que las primeras sean inconexas fuera del ámbito de la cuarta; sino que se ensamblan en la cuarta, la cual determina por su parte la conexión que une las cuatro maneras de preguntar.

En lo que respecta al pensar nos movemos en el ámbito de una tradición que abarca dos milenios y medio; hemos tomado la cuarta forma de preguntar como la directriz, pero queda expuesta la indeterminación científico-filosófica sobre la esencia de los procesos intelectuales humanos. Ahora pasaremos a analizar la posibilidad de construir modelos sobre los mecanismos instrumentales del proceso pensante y la de asociar los resultados del modelo con los reales.

Debemos analizar la secuencia de procesos que comienza con la percepción de un cierto estado de configuración externa observable (entrada) y se manifiesta mediante algún tipo de comportamiento (salida). En el ser humano pueden distinguirse ciertos segmentos instrumentales de la actividad intelectual: a) "cerebro interno", que maneja los procesos de la vida instintiva y visceral (paleoencéfalo) y b) "cerebro externo" (neocortex o corteza), del que dependen principalmente las funciones cognitivas e intelectuales más diferenciadas, que permiten percibir la realidad y manejarla mediante conceptos abstractos organizados en estructuras dinámicas.

Los niveles vitales aristotélicos; vegetativo, sensitivo e intelectual, representan jerarquías diferenciadas que requieren estructuras biológicas de diverso grado de complejidad. En la vida comportamental humana pueden distinguirse dos dimensiones: a) dimensión exterior, observable por otros (conducta) y b) dimensión interior, accesible sólo al protagonista mediante un análisis introspectivo. Ambas dimensiones constituyen aspectos esenciales del comportamiento. El "sistema nervioso central" recibe los impulsos provenientes de los sentidos, transformándolos e integrándolos de acuerdo a la complejidad de la respuesta requerida; de estas integraciones centrales surgen los imperativos para la acción en forma de impulsos eferentes dirigidos a los músculos y glándulas que deben intervenir en la conducta pertinente. La capacidad del hombre para abstraer de su percepción un modelo simbólico que representa las configuraciones externas a que se halla sometido, constituye la sustancia de los procesos intelectuales. El hombre logra abstraer de las configuraciones que percibe núcleos significativos que las representan en forma no sensorial, interiorizándolas en forma de conceptos susceptibles de manipulación simbólica.

El hombre consigue apartarse y superar la percepción meramente reproductiva de las cosas, para analizar y reconstruir

su estructura; para establecer nuevas relaciones entre ellas y reconstruir el mundo en una prefiguración virtual del mismo que se opera dentro de la mente, pero que puede plasmarse en realizaciones materiales concretas.

Ante la dificultad de penetrar rigurosamente en la esencia de los procesos intelectuales, el psicólogo ha desarrollado un campo de descripciones operativas, abandonando provisoriamente el de las definiciones ambiciosas.

En la búsqueda y formulación de tales descripciones operativas se han empleado distintas metodologías como la introspección, la representación mediante las operaciones lógicas clásicas, observación de comportamiento o recopilación de declaraciones verbales. Respecto a la metodología de búsqueda de información se han planteado interesantes discusiones, como las provocadas por la escuela conductista.

El conductista tiene como metodología describir el comportamiento humano en términos de estímulo-respuesta, y considera el comportamiento observable ante un estímulo dado como única posibilidad experimental para obtener información básica. Así para la escuela conductista el problema se centra en "dado un estímulo (E), determinar la respuesta (R)", o "dada una respuesta (R) determinar el estímulo (E) que la provocó", y se resuelve cuando E y R han sido determinadas.

Al considerar el comportamiento observable ante un estímulo como única posibilidad de obtención de información básica, el conductismo renuncia deliberadamente a una gama muy amplia de potencialidades de obtención de información.

Entendemos que si tratamos de construir modelos que simulen pautas de comportamiento humano necesitamos un cúmulo de información básica experimental, para lo que resulta conveniente utilizar toda fuente de información posible como ser la propia introspección, la introspección de sujetos experimentales, observación de conductas de sujetos experimentales, entrevistas, etc.

En el afán de construir modelos de procesos intelectuales humanos se ha caído frecuentemente en hipótesis simplistas como la de que toda actividad humana está estrictamente determinada por procesos anteriores, y por lo tanto, en principio puede ser prevista con absoluta exactitud. Nos preguntamos si este tipo de premisas brinda alguna utilidad, si constituye una buena hipótesis de trabajo; vale decir, si sirve para encauzar provechosamente la investigación. Vemos que nada de esto se cumple; en la esfera del estudio del hombre dicha premisa jamás demostró tener valor como hipótesis de trabajo. Esta teoría se contradice con realidades esenciales del comportamiento hu-

mano, el hombre es una criatura que lucha por la consecución de fines, piensa, investiga e intenta lograr lo que no existe, lo que él aspira.

Una premisa de tipo simplista es aceptable cuando constituye una buena hipótesis de trabajo, un punto de partida adecuado para investigaciones más profundas; pero el método científico no permite que una hipótesis de este tipo cristalice en dogma. Este peligro siempre existe porque la mente del investigador propende a resultar dominada en lugar de inspirada por ciertas hipótesis simplistas, quizá en su sano afán de avanzar.

Para alcanzar sus fines, el método científico debe tomar los recaudos necesarios para reemplazar las estructuras provisionales iniciales a medida que se desarrollan otras más elaboradas. Toda investigación necesita "fórmulas previas" para afirmarse y definirse, pero éstas deben abandonarse al avanzar.

El investigador en el área de simulación de procesos intelectuales humanos está interesado en desarrollar modelos que operen en forma análoga al sistema real en los aspectos bajo estudio. El intento resultará exitoso en la medida en que estos modelos brinden resultados análogos a los del proceso que se desarrolla en el ser humano simulado, independientemente de que este proceso real sea o no satisfactorio. El modelo de simulación incorpora el conjunto de hipótesis del investigador respecto al sistema de procesamiento de información subyacente al comportamiento que tiene lugar en el sujeto simulado.

El investigador que se encamina en esta dirección debe tomar varias decisiones: a) selección del comportamiento a predecir: del repertorio de comportamientos posibles, debe definirse perfectamente aquel que se intenta estudiar; b) medio de objetivación del modelo: deberá decidir si su modelo será de tipo verbal, gráfico, de tipo analógico físico o programa de computador digital; c) nivel descriptivo-explicativo del modelo: especialmente respecto al nivel de detalle o de jerarquía; en términos de procesos neuronales, de procesos lógicos de manipulación de información o a nivel integrador mediante la combinación de distintos niveles. Téngase en cuenta la complejidad y variedad casi infinita de estructuraciones neuronales que pueden formarse con los diez mil millones de células de la corteza cerebral interrelacionadas de acuerdo a una variedad y flexibilidad extraordinarias.

Encarado el estudio en un nivel superior, analizando el comportamiento humano y los procesos intelectuales subyacentes en términos de sistemas de información, entramos en el dominio de disciplinas como la psicología, la sociología y la antropología y es necesario asumir un criterio integrador que

tenga en cuenta todas las variables que confluyen en la personalidad de cada individuo.

De acuerdo a la concepción clásica la psicología es la ciencia que estudia los fenómenos u operaciones psíquicas del hombre; esta concepción estaba basada en el supuesto de la existencia de un *hombre abstracto* que reunía en sí todos los rasgos característicos de la naturaleza humana. Se analizaba al sujeto supuestamente aislado que desarrollaba su personalidad sin las influencias de los grupos o de la sociedad. Esta concepción tenía como objetivo llegar a establecer leyes de validez general para todos los hombres y cualquier desviación estaría determinada por causas accidentales o a las influencias sociales, que se consideraban contrarias a la naturaleza humana. El concepto de *naturaleza* humana constituye otra hipótesis simplista que la evolución científica fue descartando; así es como la psicología actual está caracterizada por una concepción del hombre muy distinta. Si bien acepta que el hombre nace con un determinado nivel de *inteligencia*, cierto *temperamento* y varios *impulsos* orgánicos como el hambre, la sed, el sexo, etc.; considera al hombre como ser típicamente social e histórico, y toda la génesis de su personalidad desde el momento de su nacimiento hasta la adultez transcurre dentro de un marco de complejas relaciones interpersonales.

Para nuestros fines usaremos dos tipos básicos de definiciones de la psicología: a) *ciencia del yo o de la persona individual*, que históricamente abarca dos orientaciones: la *racionalista*, que estudia los procesos psíquicos mediante la lógica y supuestos apriorísticos y la *empírica* que desarrolla sus estudios mediante la observación y la experimentación. b) *investigación sistemática de la conducta de los organismos*, conceptualización típica de la escuela conductista en la que existe una gama de posiciones respecto a la pertenencia de los fenómenos de la conciencia al campo de la investigación científica, y hasta sobre su existencia.

Los métodos más usuales en la investigación psicológica son:

- a) *método experimental*: consiste en la observación desarrollada bajo condiciones artificiales en las que el investigador reproduce a voluntad el fenómeno bajo análisis manteniendo bajo su control los factores que intervienen en la situación y tratando de reducir las variables al mínimo posible.
- b) *método clínico*: se caracteriza por el contacto directo y personal del investigador con el sujeto bajo estudio y la persecución de objetivos prácticos mediante el estu-

dio detallado y profundo basado en la observación directa con enfoques global y unitario.

- c) *introspección*: consiste en la contemplación o reflexión por parte del sujeto respecto a una experiencia, ya sea durante ésta o por medio de la memoria. Resulta en una descripción de tal experiencia en términos de percepciones, elementos y actitudes.
- d) *observación*: es el examen analítico de fenómenos o sucesos como parte de la investigación científica en el que las condiciones no están bajo control del investigador.

En un sentido sumamente amplio y generalista puede decirse que la psicología estudia la *personalidad* de los seres humanos, siendo la *personalidad* la integración organizada de todas las características cognoscitivas, afectivas, volitivas y físicas de un individuo, tal como se manifiestan en él diferenciándolo de otros individuos. El fenómeno que nos permite estudiar esta personalidad es la conducta manifiesta y exteriorizada por el individuo; sin embargo la personalidad es mucho más que el repertorio de las conductas posibles ya que comprende las motivaciones profundas, las actitudes, las creencias y otros fenómenos subyacentes. Los fenómenos psíquicos constituyen un campo de tal extensión que la ciencia general se ve forzada a desmembrarse en disciplinas específicas; además cada una de las grandes corrientes de pensamiento psicológico enfatiza en un determinado aspecto de la realidad psicológica, desarrollando sus propios métodos de investigación y trabajando con unidades de análisis más específicas de acuerdo a sus finalidades, y hasta llegan a ser exclusivas.

Entre las corrientes psicológicas más significativas merecen destacarse el psicoanálisis, el Gestalt y el conductismo; que han significado la apertura de nuevas direcciones que en su evolución han originado otras subcorrientes que constituyen en gran parte la estructura central de la construcción teórica de la psicología contemporánea.

El psicoanálisis (S. Freud) se basa en la conducta instintiva del ser humano partiendo de una estructura tripartita de la personalidad compuesta por el *yo*, el *superyó* y el *ello*. Utiliza como unidades de análisis los elementos básicos de vida y de muerte y estudia la génesis de tensiones y conflictos desde la primera infancia, las formas de superarlos y los mecanismos de defensa. Esta escuela ha desarrollado algunas técnicas específicas como por ejemplo la "libre asociación".

El Gestalt (Wertheimer, Köhler y Koffka) estudia los fenómenos de la percepción, la memoria y el proceso del aprendizaje mediante la introspección y la observación; las unidades

de análisis con las que trabaja son las estructuras y muestra a través de las leyes de la percepción la forma en que ésta varía en cada individuo, influenciada por factores personales y sociales.

El Conductismo ha dado lugar a una gama de subcorrientes cuyos extremos contienen grandes diferencias, para nuestro análisis tomaremos la línea teórica más comprensiva (Mead), que se dedica a profundizar el análisis de todas las manifestaciones, exteriorizadas o no, de la conducta del hombre en situación. Es decir que incluye la formación de hábitos, los conflictos, el aprendizaje, la frustración, etc. Adopta como unidad de análisis la "conducta molar", que de acuerdo a Bleger es una "totalidad organizada formando una unidad de experiencia con una unidad de significado". La conducta molar es un segmento de la conducta continua caracterizado por una finalidad propia, vinculado a determinado objeto y respondiendo a motivaciones específicas. Conociendo el repertorio de las modalidades de conducta posibles en un sujeto, podría predecirse su reacción frente a estímulos internos o externos; ya que aún conociendo la gran variedad de conductas posibles ante hechos semejantes, sabemos que cada cultura restringe esta variedad de acuerdo a la tendencia generalizada a nivel de individuo, grupo y sociedad a canalizar y organizar la conducta en forma rutinaria, formando estructuras predominantes que a nivel individual constituyen la "personalidad" y a nivel de agregado social la "personalidad básica".

La psicología social es la rama de la psicología que estudia la conducta del individuo en el aspecto en que ésta estimula a otros individuos, o es en sí una reacción a su conducta; describe la conciencia del individuo en cuanto es conciencia de objetos y relaciones sociales; comprende las características y fases del proceso de aprendizaje, las variables que intervienen en la conducta social, el proceso de socialización del individuo y el análisis de la persona como organización de actitudes, los procesos perceptuales y su relación con factores socioculturales, análisis de fenómenos psicoculturales como el prejuicio, o socioculturales como la enajenación y la alienación.

Para nuestros fines resulta útil además introducir algunos conceptos elementales sobre sociología, dado el papel fundamental que desempeñan los aspectos sociales del individuo en sus procesos intelectuales.

El área de conocimiento que estudia la sociología es irreductible a cualquier otra esfera de la realidad, ya que trasciende la existencia individual y toma una orientación totalista. Gurvitch analiza y resume el pensamiento de Dürkheim, uno

de los más destacados teóricos sociales y define la sociología como la ciencia que estudia con visión de conjunto y de manera tipológica y explicativa, los diferentes grados de cristalización de la vida social, cuya base se encuentra en los estados de conciencia colectiva, irreductibles y opacos a las conciencias individuales. George Mead considera que la culminación del proceso de socialización reside en la transformación del ser humano en persona, tomando la comunicación como proceso esencial.

La comunicación consiste en el intercambio de los roles sociales asumidos por el individuo mediante la internalización de los roles de los otros. Una vez que el niño llega a diferenciar su propio ser del mundo que lo circunda, comienza a internalizar distintos roles mediante el juego (madre, padre, maestro, médico, etc.), comenzando así un proceso de socialización que se continuará a medida que el individuo aprenda a tener en cada instante la función que está cumpliendo dentro de cada estructura, relacionando su actividad con los roles de los demás.

Nuestro estudio está también vinculado a la antropología cultural, ya que la cultura en que está inmerso constituye un rasgo fundamental en el ejercicio de los procesos intelectuales del individuo. La cultura constituye la base de la personalidad social; se manifiesta a través de la conducta, creaciones y actitudes individuales, pero su carácter es supraindividual ya que ningún individuo puede abarcar la totalidad de su cultura. Los elementos que constituyen la cultura se transmiten y conservan a través de la comunicación y el aprendizaje; el hombre es el único ser viviente que ha llegado a poseer una cultura por haber podido desarrollar un sistema de comunicación abstracto y simbólico que es el lenguaje; el animal puede llegar a establecer una relación de tipo semiótico mediante el aprendizaje, pero no tiene la capacidad de abstracción y generalización que le permitirían trascender los límites de la experiencia directa. El símbolo se usa en el ser humano para transmitir el saber acumulado y para posibilitar la evocación abstracta de un fenómeno y su presencia continua independientemente del tiempo y del espacio. Así la diferencia entre el hombre y los demás seres vivientes reside no en el grado sino en la especie.

La cultura de un cierto grupo humano se basa en su visión particular del mundo de acuerdo a una integración determinada por sus postulados existenciales (naturaleza humana) y normativos (sistema de valores). La integración es una característica de la cultura; nunca podrá analizarse un rasgo específico en su significado y función sin establecer la relación de éste con la totalidad cultural a la que pertenece. Así es como cada elemento de una cultura guarda una relación con el todo;

la contribución de cada parte al sistema cultural constituye su *función* y la forma en que todos los elementos se relacionan entre sí constituye la *estructura* de la cultura. La personalidad es la suma total de modos ostensibles de conducta de una persona, o al menos aquellas manifestaciones parcialmente integradas y congruentes; también puede hablarse de carácter mental básico subyacente, que produce una "personalidad social" característica en individuos provenientes de una misma sociedad con una determinada cultura.

Al estudiar los fundamentos de una determinada cultura deben usarse técnicas de psicología experimental y clínica para analizar las percepciones, motivaciones, aprendizaje, el desarrollo de la personalidad de sus miembros; así podrá conocerse también el "sistema de ideas", que se compone de aquellas estructuras cognoscitivas subyacentes en todo individuo independientemente de su manifestación.

Tiene también importancia fundamental conocer las actitudes, que constituyen estados mentales de disposición organizados a través de la experiencia y que ejercen una influencia directriz sobre las respuestas del individuo ante todas las situaciones que enfrenta.

2.1. Aplicaciones generales.

En nuestro estudio de simulación de procesos intelectuales, al encararlos en toda su complejidad estamos reuniendo y produciendo material de valor inestimable para las ciencias que estudian al hombre, ya que podrán ver los problemas que les son propios en una forma más concreta; a su vez nuestra área es receptiva a los aportes de todos los campos del conocimiento, ya que constituye una cuestión de tal complejidad que debe tratarse desde todos los aspectos posibles.

Este enfoque pluridisciplinario brinda la posibilidad de que una integración de los resultados de la investigación en las distintas ciencias converja hacia el descubrimiento de la esencia de los procesos bajo estudio.

Ultimamente se han producido muchos libros e informes sobre diversas aplicaciones de la computación electrónica y la teoría asociada a la investigación en las ciencias del comportamiento humano. En lo que resta del capítulo describiremos algunas de las aplicaciones más significativas y citaremos algunos ejemplos concretos.

La aplicación básica de la computación electrónica en ciencias del comportamiento consistió en análisis estadístico de

información con el objeto de hallar relaciones funcionales entre las variables del comportamiento. Este tipo de aplicación aprovecha la potencialidad de las máquinas para desarrollar procesos de cálculo; así es como la investigación cuantitativa en el ámbito del comportamiento ha manifestado un avance sin precedentes gracias a la posibilidad de ejecutar procesos de cálculo en forma rápida y exacta.

El análisis de datos verbales es una aplicación que se desarrolló al descubrirse que las computadoras son, más que calculadores rápidos, sistemas procesadores universales de símbolos; que podían aplicarse a la evaluación de la personalidad. La descripción de la personalidad y predicción de comportamientos constituye un área que ha sido tratada por muchos psicólogos. Evidentemente el método de estudio normativo es ideal para la aplicación de computadora; pudiendo ejecutarse automáticamente todos los procesos que van desde la presentación del estímulo hasta el registro de la respuesta.

Existen varios desarrollos significativos en la aplicación de sistemas interpretadores de cuestionarios psicológicos usuales, como el Minnesota Multiphasic Personality Inventory (Swenson y otros, 1965), como así también en evaluación de tests basados en imágenes o manchas de tinta, como el "Holtzman Ink-blot Technique" (Moseley y otros, 1963) y el test de Rorschach (Piotrowski, 1964); con estos programas evaluadores se han obtenido resultados muy similares a los que brinda la evaluación convencional humana. El beneficio que han brindado estos sistemas reside especialmente en que su elaboración requiere que las reglas de interpretación y evaluación sean precisas y las definiciones operacionales rigurosas.

Se han desarrollado también simulaciones de comportamientos individuales de expertos en interpretación de tests, como los sistemas de Kleinmuntz (1963) y de Cogswell y Estavan (1965) en los que se estudia y describe el comportamiento interpretativo real de un experto y se lo traslada a programas de computador, que posteriormente son utilizados para interpretar evaluaciones de otros individuos. Cogswell y Estavan programaron una computadora en time-sharing para conducir entrevistas de tipo vocacional con estudiantes, utilizando los resultados provenientes de la rutina de interpretación. Starkweather (1965) extendió estas aplicaciones a entrevistas de diagnóstico psiquiátrico.

Aparentemente las aplicaciones más promisorias de las técnicas basadas en el uso de computadoras en el campo de la evaluación y predicción del comportamiento individual son aquellas que tienden a combinar el análisis normativo computeriza-

do con la sensibilidad y sentido común de los especialistas para el desarrollo de la entrevista personal; estas aplicaciones resultarán cada vez más factibles a medida que se generalice el uso de las técnicas de tiempo compartido (time-sharing).

Respecto a las técnicas de simulación se han desarrollado varias investigaciones. Newell y Simon (1961) elaboraron modelos de simulación de procesos cognoscitivos y juntamente con J. C. Shaw desarrollaron el "General Problem Solver" (1959), un programa que emplea procesos heurísticos aplicables a una amplia gama de actividades.

Uno de los programas más completos en el que se empleó el computador para procesar datos verbales fue el "General Inquirer", de la Harvard University (Stone y otros, 1962), que permite el análisis automático del contenido de cualquier tipo de material verbal. Este análisis se desarrolla con referencia a un "diccionario" de términos organizado de acuerdo con un sistema predeterminado de conceptos, pudiendo compilarse diferentes diccionarios para distintos propósitos. El resultado consiste en un conjunto de índices cuantitativos sobre el uso de los conceptos. En la Universidad de California, J. A. Starkweather (1964) ha usado un computador para compilar las características del uso de palabras durante entrevistas, antes de proceder a un análisis intensivo del contenido. Un informe posterior se refiere al uso del computador como entrevistador (1965).

Las ventajas que ofrecen los sistemas automáticos de evaluación de tests cuando son administrados masivamente residen en la rapidez del procesamiento y su rigurosa objetividad; estos sistemas "deshumanizados" no pueden cubrir el rango de posibilidades de un experto, pero estimamos que lo promisorio de estas aplicaciones está en la posibilidad de elaborar sistemas interactivos hombre-computador.

En la Universidad de Texas, R. F. Peck, Donald J. Veldman y Shirley Menaker están trabajando en el desarrollo de un sistema de análisis de la personalidad usando un sistema de completar blancos en frases (Peck y Veldman, 1962). Las respuestas, consistentes en palabras, pueden ser compiladas y reducidas a "palabras raíces" y pueden usarse para predecir el comportamiento de un sujeto ante determinada situación o para elaborar un análisis descriptivo del comportamiento verbal manifiesto en términos normativos.

2.2. Aplicaciones en psiquiatría.

Kenneth Mark Colby, M. D., especialista en psiquiatría, trabaja en el Departamento de Ciencias de la Computación de la

Stanford University dedicándose a investigaciones en el área de simulación por computadora. Los problemas clínicos de psicoterapia requieren un gran esfuerzo de investigación, ya que se conoce muy poco aún sobre los procesos esencialmente cognoscitivos y afectivos.

Los problemas clínicos se conceptualizan como procesos de información que incluyen funciones mentales de gran complejidad; así la estructura conceptual y las técnicas empleadas en ciencias de la computación, resultan adecuadas para su tratamiento.

El objetivo perseguido es el desarrollo de teorías y modelos más satisfactorios respecto de procesos psicopatológicos y ampliar las posibilidades de experimentación sobre modelos, cuyos resultados se incorporarán al cuerpo teórico.

Hasta la actualidad se han desarrollado dos tipos de programas: a) modelos de sistemas individuales normales y anormales, y b) programas entrevistadores, capaces de conducir diálogos con individuos reales para recopilar datos respecto de las relaciones interpersonales. El modelo de un sistema individual consiste en una gran estructura de datos sobre conceptualizaciones del individuo y una secuencia de instrucciones para procesar dicha información. La información básica se recopila mediante entrevistas con el sujeto a simular, posteriormente se elabora la secuencia de instrucciones, se opera el modelo, se obtienen los resultados y nuevamente se emplea la entrevista para que el paciente verifique o no dichos resultados en el proceso de validación.

La conveniencia de elaborar programas entrevistadores reside especialmente en que la subjetividad de los entrevistadores humanos introduce cierta variación y sesgo, de manera que no se presenta generalmente la consistencia que sería deseable. Estos inconvenientes pueden superarse mediante sistemas entrevistadores hombre-máquina en línea.

Más adelante nos referiremos a sistemas específicamente entrevistadores; la finalidad principal de las experiencias del doctor Colby reside en lograr que la máquina aprenda a "simular" ciertas convicciones y complejos del individuo entrevistado.

Veamos por ejemplo en qué consiste el programa "modelo de personalidad neurótica". Este programa contiene dos niveles distintos de memoria artificial: el nivel inferior almacena ciertos conceptos básicos tales como "yo", "padre", "madre", "hermano", "amor", etcétera; el nivel superior almacena relaciones entre estos conceptos representando convicciones y creencias. Estas convicciones pueden ser compatibles, independientes o contradictorias.

El operador induce al computador a relacionar algunas de las convicciones; en caso de que no existan contradicciones graves, la computadora las expresa mediante la tipeadora, si existe conflicto una de las convicciones se modifica para hacerla compatible con las demás por un proceso análogo al mecanismo de defensa psicoanalítico. Cuando después de sucesivos intentos no se logra establecer la transformación necesaria para compatibilizar las convicciones, aumentará el valor de una variable que representa la "ansiedad"; que llegando a ciertos valores hace posible que la máquina rehúse continuar pensando en el argumento generador del conflicto.

Además de la utilidad de estos modelos para la investigación, son sumamente adecuados para capacitación de psicoterapeutas, ya que los alumnos pueden someter a las máquinas a sesiones de psicoanálisis, aprendiendo a inducirla a tratar argumentos que no provocan conflicto, para después alentarla gradualmente a abordar temas más rechazados, o inducirla a examinar situaciones de conflicto que llegarán a debilitarse si carecen de apoyo en otras convicciones.

John C. Loehlin, de la Universidad de Texas (Austin), construyó el programa "Aldous", un modelo de la personalidad humana dotado de las siguientes facultades: a) reconocer la configuración que se le presenta (simbolizada por un conjunto de números); b) reaccionar emotivamente frente a la configuración; c) seleccionar acciones a ejecutar ante la configuración; d) capacidad de introspección, y e) capacidad de aprendizaje. Ante la configuración de entrada, el Aldous posee un repertorio de tres reacciones emotivas: i) temor; ii) cólera; iii) atracción. Actuará con una actitud de repugnancia, agresión, acercamiento, conflicto o indiferencia. Al cabo de cierto número de pruebas aprenderá a asumir actitudes generales o específicas ante los objetos y podrá reaccionar con rapidez o con vacilaciones, de acuerdo con sus experiencias precedentes y su estado en el instante considerado.

Las configuraciones presentadas a Aldous consisten en un conjunto de números (tres o más) que simbolizan determinadas características de la configuración. Las actitudes hacia las configuraciones se van modificando paulatinamente de acuerdo con el éxito o el fracaso resultante de los sucesivos enfrentamientos con la configuración, ya que posee capacidad de aprendizaje basada en la posibilidad de reconocer las consecuencias de sus actos.

También es posible desarrollar por simulación el enfrentamiento entre dos programas Aldous (A1) y (A2). Si la repugnancia de (A1) en respuesta a la agresión de (A2) es menos

vigorosa que la agresión, (A1) registra un fracaso con la vergüenza asociada, mientras que (A2) asumirá un "sentimiento" de satisfacción. Una experiencia con perspectivas terapéuticas consiste en intentar modificar la actitud agresiva negativa de uno por medio de la acción del otro. Los resultados más interesantes se obtienen con un modelo de "psicoterapeuta" capaz de evitar las agresiones fuertes, contrarrestar los ataques débiles y favorecer una retroacción moderada del modelo con personalidad agresiva. El resultado es una "transferencia" y un fuerte "sentimiento positivo" hacia el psicoterapeuta.

En 1969 el centro de psiquiatría Douglas Singer Zone Center de Rockford (Illinois-EUA) comenzó a utilizar el computador para aumentar la precisión y seguridad de los diagnósticos de enfermedades mentales, mejorar el control de los pacientes y seleccionar el tratamiento más adecuado para cada caso.

En una enorme mayoría de los casos se determinó que los diagnósticos generados por el computador guardan un paralelismo casi perfecto con los formulados por los psiquiatras.

El mecanismo es el siguiente: cuando un paciente llega al centro es examinado profundamente por un psiquiatra, quien evalúa detalladamente los síntomas y la gravedad del paciente. Estos datos se introducen en el computador y éstos los confronta con las 38 (treinta y ocho) "características" de síntomas almacenados en la memoria, reconocidos por la American Psychiatric Association. Una vez completado el análisis el computador imprime un informe donde aparecen los tres diagnósticos más probables y las probabilidades de cada uno.

La decisión final corresponde siempre al profesional, pero la ventaja del proceso computerizado reside en que éste puede confirmar sus juicios, evaluar nuevas hipótesis o estudiar más profundamente el caso.

En el Institute of Living (Connecticut), se está tratando de desarrollar sistemas y técnicas computerizadas para automatizar todo el proceso de atención del paciente psiquiátrico.

En 1969 se programaron los computadores para que trazaran gráficos sobre los cambios en el comportamiento de los pacientes durante un período de tiempo; para ello se representaban cuantitativamente la ansiedad, depresión, comportamiento adaptativo y otros aspectos de la conducta que constituyen los datos básicos del paciente psiquiátrico y que son indicadores muy sensibles de las variaciones. Estos datos se obtienen, por ejemplo, de los registros diarios de las enfermeras, una lista de 120 observaciones comunes del comportamiento, incluyendo hábitos personales, conducta social, comportamiento durante las comidas y du-

rante el sueño, reacciones a los tratamientos y manifestaciones generales.

La función de la enfermera se reduce simplemente a hacer una marca con lápiz junto a las observaciones que se aproximen más a la descripción de la conducta del paciente, también puede incluir comentarios escritos. Se introducen estas observaciones en el computador y éste entrega un informe de media página, calcula el grado de ansiedad, depresión, etcétera, mostrados por el paciente.

Algunas otras pruebas computerizadas usadas en el Institute of Living son el Minnesota-Hartford Personality Assay y el Minnesota Multiphasic Personality Inventory.

El MHPA es un sistema desarrollado para proveer medios sistemáticos de obtención y registro de juicios clínicos, para medir el nivel de adaptación del paciente y los cambios resultantes de la intervención terapéutica. El sistema computerizado elaborado por el IL se usa en otras muchas instituciones.

El computador ejecuta toda la secuencia de operaciones de administración, cuantificación y evaluación del estado mental del paciente usando la técnica de *Q-sort*. Esta secuencia incluye seis pasos: a) el médico o la enfermera observan el comportamiento del paciente durante el período de tratamiento o durante entrevistas especiales; b) se selecciona el formulario preimpreso para la situación; c) se marcan características para cada dimensión del comportamiento de acuerdo con la cualidad descriptiva del paciente en ese momento; e) se introduce esa información en el computador como entrada al paquete de programas de *Q-sort* para que el computador lo cuantifique; f) el computador calcula e imprime un gráfico del estado mental del paciente, se grafican las cuantificaciones (*T scores*) para cada una de las 20 escalas consideradas. El *T score* representa una comparación normalizada del paciente contra un conjunto de normas que han sido acumuladas sobre la base de la experiencia en una gran población a la que se ha administrado el *Q-sort*. Un *T score* con valor 50 coincide con la media de la población y 10 unidades de *T score* indicarán un desvío standard respecto al valor medio; g) El *T score* del paciente puede ser usado por el médico para evaluar la condición del paciente, o puede ser comparado con el de otros pacientes (o el del mismo paciente en otras oportunidades) para investigaciones sobre los efectos de diversos tratamientos.

Los 20 aspectos considerados en el MHPA son: 1) Deficiencia en capacidad afectiva; 2) Deficiencia en conformidad social; 3) Deficiencia en comunicación social; 4) Inefectividad social; 5) Autoconceptualización desintegrada; 6) Autoconceptualiza-

ción deficiente; 7) Defensas compulsivas; 8) Defensas de ansiedad; 10) Aislamiento social; 11) Comportamiento agresivo hostil; 12) Irritabilidad; 13) Intereses disociativos; 14) Tendencia disociativa; 15) Pensamiento desordenado; 16) Pensamiento paranoico; 17) Tendencias homosexuales; 18) Depresión; 19) Anhedonia, y 20) Tendencia autodestructiva.

También en el Institute of Living se usa el Minnesota Multiphasic Personality Inventory (MMPI) cuya ventaja principal consiste en la posibilidad de que el paciente se lo autoadministre. Este "inventario" ha sido desarrollado por W. M. Swenson, y proporciona en 18 segundos una especie de retrato mental del paciente en una trama de 9 características; a) hipocondría; b) depresión; c) histeria; d) desviaciones psicopáticas; e) grado de masculinidad y femineidad; f) paranoia; g) psicoastenia; h) esquizofrenia; i) hipomanía; mediante la información proveniente de respuestas a un conjunto de 550 proposiciones. Las respuestas son sólo: "cierto", "falso", "no puedo contestar"; en realidad clasifica tarjetas preimpresas en grupos. El test está libre de sesgo del interpretador o entrevistador, ya que las tarjetas ya clasificadas por el paciente se introducen directamente en el computador.

Otra aplicación que se ha implementado en el Institute of Living orientada más hacia la investigación es la referente al tratamiento de grupos de pacientes; para ello se suministran al computador datos sobre diez grupos, con el juicio del terapeuta sobre los cambios ocurridos en cada uno de los pacientes después de nueve o diez meses de tratamiento en grupo. De acuerdo con esta evaluación se coloca a los pacientes en una de estas tres categorías: los que según el criterio del terapeuta respondieron muy bien al tratamiento grupal, aquellos cuyo progreso es dudoso, y los que reaccionaron pobremente.

Luego el investigador revisa la información tomada al momento del ingreso y aplica técnicas estadísticas, tratando de determinar denominadores comunes dentro de los grupos. Si puede aislar ciertos factores estará en condiciones de poder comparar el cuadro de un paciente con los cuadros compuestos de los grupos y predecir las posibilidades de progreso que puede ofrecerle el tratamiento en grupo.

2.3. Autómatas que sienten.

Suponiendo que se llegara a aceptar la posibilidad de autómatas que piensan, permanecería aún dudosa la posibilidad de autómatas que desarrollen sentimientos.

Un científico llamado N. M. Amosov ha trabajado especialmente en esta área partiendo de modelos altamente elaborados de pensamiento y comportamiento humanos, en los que posteriormente incorpora programas de sentimiento, conciencia y creatividad.

En su trabajo, Amosov identifica seis tipos de programas como determinadores del comportamiento humano:

- a) un programa de sensaciones físicas, percepción y procesamiento de información referente al sistema físico;
- b) programas de percepción de influencias del contexto y para el reconocimiento de las cualidades del contexto;
- c) un programa para la acción, es decir, para impartir energía e información sobre el contexto;
- d) un programa de comunicación;
- e) un programa para el sistema de la conciencia, y
- f) un programa para creatividad y descubrimiento.

Así ha elaborado un modelo psicocibernético del comportamiento inteligente.

Las estructuras nerviosas denominadas globalmente "sistema límbico", situadas bajo la corteza cerebral humana constituyen el "cerebro emotivo", es decir, la zona donde físicamente se desarrollan los "sentimientos básicos" (miedo, placer, ira, sensualidad, alegría, etcétera).

Es evidente que ante la falta de los aportes de tipo emotivo provenientes del sistema límbico, la corteza cerebral no es más que una computadora sumamente sofisticada. Las emociones límbicas son *filtradas* por una intensa superestructura cortical de "memorias", de asociaciones y en especial por el mecanismo del lenguaje, que filtra, modera y desvía las emociones primordiales. La psiquiatría está ahora sumamente interesada en los mecanismos que se desarrollan en el sistema límbico y en determinar la forma en que se establecen en el cerebro las conexiones emotivas.

Esta dirección es sumamente promisoría en el área en que estamos interesados, ya que la combinación de este análisis con un estudio en términos de procesamiento de información podrá brindar aportes de inmenso valor para la comprensión del comportamiento emocional humano.

BIBLIOGRAFIA

Amosov, N. M., *Modeling of thinking and the mind*, trans. by L. Fingold and L. J. Fogel. New York, Spartan Books, 1967.

- Borko, H., ed., *Computer Applications in the Behavioral Sciences*, Englewood Cliffs, N. J.: Prentice Hall, 1962.
- Borko, H., *A Factor Analytically Derived Classification System for Psychological Reports, Perceptual and Motor Skills*, 1965, 20 (2), pp. 393-406.
- Bown, O. H., *The Development of a Self-Report Inventory and its Function in a Mental Health Assessment Battery*, *American Psychologist*, 1961, 16 (7), p. 402.
- Cogswell, J. F. & Estavan D. P., *Explorations in Computer Assisted Counseling*, System Development Corporation, TM-2582, 1965.
- Colby, K. M., *Experimental treatment of neurotic computer programs*, *Archives of General Psychiatry*, 10, 220-227 (1964).
- Colby, K. M. & Gilbert, J. P., *Programming a computer model of neurosis*, *Journal of Mathematical Psychology*, 1, 405-417 (1964).
- Colby, K. M., *Computer simulation of neurotic processes*, en *Computers in Biomedical Research*, Vol. I, Stacey & Waxman, Eds., Academic Press, New York (1965).
- Feldman, J., *Simulation of Behavior in the Binary Choice Experiment* en H. Borko, ed. *Computer Applications in the Behavioral Sciences*, Englewood Cliffs, N. J.: Prentice Hall, 1962.
- Gullahorn, J. T. & Gullahorn, J. E., *Computer simulation of human interaction in small groups*, *Proceedings AFIPS*, 1964, Spartan Books, Baltimore, pp. 103-113.
- Helm, C. E., *Simulation Models for Psychometric Theories*, *Proceedings of American Federation of Information Processing Societies*, Washington: Spartan, 1965, Vol. 27, Part. 1.
- Hilf, F., Colby, K. M., and Wittner, W., *Machine-mediated interviewing*, *Stanford Artificial Intelligence Project*, Memo AIM-112, march 1970.
- Holsti, O. R., *An adaptation of the general inquirer for the systematic analysis of political documents*, *Stanford University*, 1964 (mimeo).
- IBM, *Application Brief, Computer Techniques in Patient Care at the Institute of Living*.
- Kleinmuntz, B., *Personality test interpretation by digital computers*, *Science*, 1963, 139 (3553), 416-418.
- Loehlin, J. C., *A computer program that simulates personality*; en S. S. Tomkins & S. Messick, eds., *Computer Simulation of Personality*, New York; Wiley, 1963.
- Moseley, E. C., D. R. Gorham and E. Hill, *Computer scoring of inkblot perceptions*, *Perceptual and Motor Skills*, 1963, 17 (2), p. 498.
- Peck, R. F. & Veldman D. J., *Computer analysis of personality, research plan and supporting data*, *Personality Research Center*, University of Texas, 1962 (mimeo).
- Starkweather, J. A. & Decker, B. J., *Computer analysis of interview content*, *Western Psychological Association*, april 1964, *American Psychologist*, 1964, 19 (9), p. 707.
- Starkweather, J. A., *Computest, a computer language for individual testing, instruction and interviewing*, *Psychological Reports*, 1965, 17 (1), pp. 227-237.
- Swenson, W. M., H. P. Rome, J. S. Pearson and T. L. Brannick, *A totally automated psychological test*, *Journal of American Medical Association*, 1965, 191, pp. 129-131.
- Ozkaptan, H., *Computer simulation of man integrated systems*, *Behavioral Science*, 8.3, july 1963.

TRABAJOS PRESENTADOS EN EL SIMPOSIO DE LA AMERICAN
SOCIETY FOR CYBERNETICS: CYBERNETICS & THE MANAGE-
MENT OF ECOLOGICAL SYSTEMS (OCTUBRE DE 1970)

- Maturana Romecin, H. (Universidad de Chile, Santiago), *Biology of cognition.*
- Heiser, John F., *Psychopathology and Cybernetics.*
- Gose, Earl E., *Remarks on the design of computerized pattern recognition systems and some comparisons to human pattern recognition.*
- Ricci, Fred J., *A cybernetical model of man in his cultural environment.*
- Austin, George M., *Human brain stimulation and alterations in perception.*
- Albus, James S., *The cerebellum: A substrate for list processing of the brain.*
- Sutro, Louis, *First steps toward a model of the human nervous system.*

Capítulo 3. INTELIGENCIA ARTIFICIAL

¿Dominadores de la humanidad, o esclavos obedientes? ¿Entes capaces de razonar y sacar provecho de la experiencia, o un conjunto de piezas electrónicas que sólo pueden interconectarse y disponerse de tal modo de llevar a cabo los procesos que desee quien lo posee? ¿Cerebros gigantescos que provocarán una revolución en los conocimientos única en la historia o simples "calculadores" numéricos? El desarrollo de las técnicas de computación de los últimos años se presta a estos y a muchos otros interrogantes, algunos de los cuales pertenecían hasta hace muy poco al dominio de la ciencia ficción, del mismo modo como los viajes a la Luna antes del primer paso de Armstrong. Sin embargo, contrariamente a lo sucedido con las experiencias espaciales, que sólo tienen efecto directo sobre la mayor parte de la humanidad desde el punto de vista emocional y que pronto pasan a un segundo plano, de las hipotéticas "inteligencias artificiales" resultaría un cambio radical en las condiciones de vida en nuestro planeta. Pero el estudio de estas consecuencias escapa en sí al propósito de este trabajo, ya que nuestro objetivo es más bien el analizar la posibilidad de la existencia de estas "máquinas de pensar".

Es probable que estos "cerebros" se basarán técnicamente en dispositivos similares a las actuales computadoras electrónicas digitales¹, con ciertas posibles modificaciones de diseño ("computación paralela", por ejemplo), que harán posible mejorar gran parte de los procesos (en cuanto a velocidad y organización de los datos), aunque probablemente no lleven a un

¹ En los últimos años se han llevado a cabo estudios para combinar computadoras digitales y analógicas (dispositivos "híbridos"), pero es difícil predecir qué papel podrían llegar a desempeñar éstos en los estudios aquí presentados. Si bien los avances tecnológicos pueden conducir a la obtención de computadoras más veloces (lo cual podría implicar la posibilidad de realizar operaciones por el momento imposibles), es aún más probable que los próximos años se distingan más por una revolución en los conceptos de programación ("software") que en el área técnica propiamente dicha.

enfoque conceptual distinto. Esto se debe a que los actuales sistemas de procesamiento de la información serían suficientemente generales como para permitir modelar y prever dichas situaciones alternativas, y no a su naturaleza electrónica en sí: es un prejuicio bastante común el intentar buscar analogías entre nuestro sistema nervioso y las computadoras basándose en el aspecto eléctrico de ambos; sin embargo, además de ser ésta una analogía falsa (se pasan por alto las reacciones químicas sinápticas y el hecho de que las primeras computadoras usaban medios mecánicos), no es de importancia, ya que nos interesa el aspecto *funcional*, el *comportamiento* de los sistemas en consideración.

En lo sucesivo, cuando nos refiramos al término "computadora", estaremos aludiendo, en la mayoría de los casos, a una computadora con un programa *ya almacenado*, del mismo modo que al referirnos a un "hombre" estamos teniendo en cuenta su aspecto psicológico. Por otra parte, es preferible, de acuerdo con Minsky², no considerar la computadora como un simple conjunto de conmutadores dispuestos de cierto modo sino, de un modo más global, como "un grupo de elementos de asociación de símbolos y control de procesos, y los programas... una red de procesos entrelazados, formuladores de metas y evaluadores de medios y fines". Las razones para adoptar este enfoque se verán más claras a medida que avancemos en nuestro estudio. Una consecuencia inmediata, al menos, es que ahora no será fácil distinguir las implicaciones últimas del concepto, lo cual se nos aparenta posible con una definición simplista.

El papel actual de la computadora en la sociedad es el de una "herramienta intelectual" que permite, mediante su capacidad para el procesamiento veloz de grandes volúmenes de información, ejecutando cientos de miles de instrucciones por segundo, cumplir con las tareas antes consideradas prácticamente imposibles o, al menos, que hubiesen consumido un tiempo y costo inmensamente mayor. Ciertamente ni las grandes proezas espaciales ni las facilidades de las comunicaciones actuales ni gran parte de los dispositivos técnicos que se crean constantemente —¡aun los modelos modernos de computadoras!— hubiesen sido posibles sin haber contado, en su momento, con las computadoras necesarias. En cierto modo las podemos comparar —en su aspecto de "herramienta"— con el teléfono y las comunicaciones en general: si bien podríamos manejarlos con mensajeros y, a corta distancia, con señales, es innegable que la

² Minsky, Marvin, *Semantic information processing*, M.I.T. Press, Cambridge, Massachussets, 1968.

vida moderna tendría otro aspecto completamente diferente de no ser por su presencia.

En síntesis, nos encontramos con computadoras en todas partes del mundo, cumpliendo las tareas más diversas de un modo preciso, seguro y veloz, aparentemente llevando a cabo siempre al pie de la letra las órdenes que se les han impuesto: una cantidad de esclavos sumisos y cumplidores, que sólo se “equivocarán” debido a un error del programador (¡gestos humanos!); o quizá debido a alguna falla técnica, en cuyo caso lo notificarán debidamente mediante su sistema de detección de errores. ¿Entonces las máquinas pueden *pensar*? ¿pueden dar algún signo de “inteligencia”, “sentido común” o “entendimiento”? No, ciertamente no en *esas condiciones*.

En el capítulo anterior, vimos cómo es posible utilizar las computadoras con el fin de simular variados procesos cognoscitivos. Pero aun allí, lo que se deseaba era *imitar* al hombre con fines de estudios psicológicos y sociológicos, mediante un programa previamente elaborado. Sin embargo, se dejó entrever algunas de las características posibles a las máquinas en las que no se había reparado adecuadamente con anterioridad: la capacidad de aprendizaje, de modificar su comportamiento de acuerdo con los resultados de la interacción con el medio ambiente, la posibilidad de cierto razonamiento y formación de conceptos. Pero si le es posible imitar, surge inmediatamente la posibilidad de intentar que actúe por sí misma.

Claro que esto no sería por muchos clasificado como “pensar”, ya que un programador se tomó el trabajo previo de especificar las “reglas del juego”. Sin embargo, esta objeción no puede considerarse válida por variadas razones; luego las analizaremos con detalle, pero baste aquí señalar una de las más lógicas y al mismo tiempo intuitivamente fuertes: nosotros, al nacer, nos encontramos *exactamente* en las mismas condiciones. A través de nuestros ácidos nucleicos, en cada una de nuestras células, traemos el programa que determinará nuestro crecimiento y desarrollo —y por lo tanto nuestro comportamiento inicial ante un medio dado— almacenado en nuestro “código genético”. La información genética se encuentra, de acuerdo con los resultados conocidos de las investigaciones, en el ADN (ácido desoxirribonucleico) de nuestros cromosomas y lo que nos determinaría por completo es la ordenación de los varios “eslabones” diferenciados de que se compone, no su tipo en sí. Sin embargo, esto no difiere del caso de una computadora con un programa ya almacenado, que consiste, si tomamos sucesivamente todas las posiciones de memoria, en una sucesión de ceros y unos, y una diferencia radical del comportamiento y estructura

(otro programa) se podrá obtener sencillamente cambiando la disposición de estos dígitos binarios.

Volviendo a nuestro tema; si *las condiciones* han cambiado, y ahora ya no hablamos de programas estáticos sino de entes capaces de aprender y automodificarse, en definitiva, ¿pueden las máquinas pensar? ¿pueden desarrollar las computadoras electrónicas actividad clasificada como "inteligente"?

Para responder objetivamente a estas preguntas, debemos dejar de lado una cantidad de prejuicios. En primer lugar, quien considere el pensamiento como una actividad misteriosa del "espíritu" y propia del ser humano *por definición*, obviamente no podrá nunca aceptar una discusión. Luego, el considerar a la inteligencia como una cualidad, un atributo, propio exclusivamente al ser humano. Y aun si consideramos abstractamente la inteligencia, es común subestimarla en cuanto entendemos y podemos describir el mecanismo de alguna acción particular que se le relacione. Este cambio constante de valoraciones nos será adverso al intentar juzgar el comportamiento de un programa, en el cual es más fácil observar los pasos particulares que en el cerebro; por lo tanto, sería útil que el lector se detenga un momento a recapacitar sobre lo común de este comportamiento.

Aceptando entonces la inteligencia y el pensamiento como hechos experimentales, correspondería entonces ponernos de acuerdo sobre el significado de estos conceptos y otros relacionados. Notemos ante todo que las definiciones que se puedan dar nunca llegan a ser suficientemente aclarativas, ya que hacen uso de conceptos que a su vez deberán ser precisados o que traen consigo una dosis de ambigüedad tan grande como la palabra definida. Es por eso que muchas veces nos contentaremos con dar sólo *propiedades* de los conceptos, dejándolos en gran parte indefinidos; otras veces proponemos problemas postulando que su resolución requiere la cualidad en cuestión, lo que configuraría un tipo de definición "implícita"; y en gran parte de los casos, postulamos *definiciones operativas* al presentar los diversos problemas.

En psicología encontramos definida la "inteligencia" como un complejo general de rasgos y aptitudes referidas a la flexibilidad de adaptación al medio ambiente, que incluyen principalmente las aptitudes de aprender, pensar en forma abstracta, usar el lenguaje y percibir relaciones complejas. Veremos cómo se pueden conseguir tipos similares de comportamiento en sistemas artificiales, en particular el aprendizaje (3.1), la percepción de relaciones y configuraciones (3.2) y el razonamiento simbólico (3.3) con uso del lenguaje.

El término "inferencia" es otro que aplicaremos en repetidas oportunidades. Es un concepto que aparece en diversos contextos, relacionado en general a los procesos discursivos, aplicándose a la deducción, analogía e inducción. Para obtener una definición operativa en relación a nuestro estudio, sólo diremos que se ha realizado una *inferencia* cuando se ha obtenido una respuesta a partir de más de una proposición inicial mediante el uso de ciertas reglas, exceptuando el caso de transformaciones puramente sintácticas de una entrada. Las "proposiciones iniciales" contendrán, generalmente, conocimientos previamente establecidos (inclusive axiomas lógicos, etcétera).

En cierto modo, lo que se desea es obtener máquinas capaces de resolver problemas que requieran de "ingenio": "procesos de decisión complejos en medios potencialmente infinitos e incontrolables" y "sentido común": "el poder deducir automáticamente y por cuenta propia una cantidad amplia de consecuencias inmediatas de lo que se le dice y los conocimientos que ya posee" (McCarthy) ³.

Un *problema* se presenta cuando existen varias alternativas de acción y se requiere encontrar cuál es la correcta. Para resolverlos contamos los seres humanos con *modelos* internos (ver cap. 1) del mundo exterior; es así como podemos evaluar internamente nuestras decisiones antes de llevarlas a la práctica, recreando la realidad mediante nuestra imaginación (volveremos al tema en 3.3.), y en este contexto es posible definir el aprendizaje como el proceso de completar y mejorar el modelo, modificando, en consecuencia, nuestro comportamiento en un sentido cualitativo a definirse de acuerdo con nuestros fines. Grey Walter considera el "aprender" como el reconocimiento de configuraciones (3.2), tomando este último término en el sentido general de "agrupamiento de acontecimientos distinguibles en el espacio-tiempo".

Un método útil de representar la inteligencia es mediante un continuo n-dimensional, siendo cada una de las dimensiones correspondiente a un "factor" psicológico, de acuerdo con Paul Armer ⁴, quien además propone reemplazar el término "pensar" por su equivalente "habilidad para procesar información". En el continuo así definido, sería lícita la comparación del hombre y la máquina de un modo general, evaluando los méritos de cada uno en su eje respectivo. Es claro que en ese caso será difícil

³ McCarthy, John, *Mechanization of thought processes*, Vol. I; Proceedings Symp. Mat. Phys. Lab., Londres, noviembre 24-27, 1958.

⁴ Armer, P., *Attitudes toward intelligent machines*, Symposium on bionics, Wodd Technical Report 60, reproducido en *Computers & Thought*, McGraw Hill, 1963.

compararlos directamente (no es un orden lineal), pero una tal comparación (como en el caso del cociente intelectual) sería simplemente una simplificación, y difícilmente nos pueda interesar. Un ejemplo en dos dimensiones de lo expuesto se puede obtener considerando solamente la "velocidad" y la "complejidad" de los procesos en ambos. La máquina se encontrará ciertamente más lejos del origen en el eje de "velocidad", pero el hombre tendrá su ventaja sobre el restante. Entonces la meta de los estudios para la obtención de inteligencias artificiales, siempre según Armer, se podría explicar gráficamente como el intento de "alejar del origen" el punto que represente al comportamiento de la máquina.

En 1950, el matemático inglés A. M. Turing⁵ propuso lo que se ha dado en llamar el "test de Turing" para decidir si se puede considerar a una máquina como inteligente. Este test consiste en un hipotético juego que llamó "el juego de las imitaciones", una especie de "Quién es quién" a tener lugar entre una computadora y dos hombres, todos ubicados en lugares diferentes y comunicados por un medio tal como una teletipo (para eliminar consideraciones de tipo físico y poder concentrarse en los aspectos intelectuales de los participantes), siendo el objeto de uno de los participantes humanos el identificar mediante una sucesión adecuada de preguntas y respuestas, cuál es la computadora y cuál el hombre. El papel del otro humano sería el de tratar de ayudar al primero, mientras que la máquina tendría como objetivo tratar de confundirlo. Notemos que mientras más "inteligente" ésta sea, más se le ocurrirán modos de llevar a su interlocutor a una decisión errónea. Aunque este "test" puede ser criticado en varios aspectos, es interesante interpretar la pregunta "¿puede una máquina pensar?" como "¿puede una máquina ganar el juego de las imitaciones?". Quizá uno de los aspectos más interesantes de esta proposición sea el hecho de que Turing ideó todo esto, e inclusive dio pautas para las investigaciones posteriores, cuando todavía no existían gran parte de los adelantos conceptuales en materia de computación —lenguajes de alto nivel, etcétera— que hoy hacen posible prever computadoras con tales aptitudes.

Los intentos de construcción de sistemas con capacidad propia de decisión basándose en los seres vivientes y, recíprocamente, el estudio de los sistemas vivos basándose en resultados obtenidos de la experimentación con modelos, tuvo su punto de partida concreto en los primeros trabajos de cibernética ("la ciencia del control y la comunicación en el hombre y la má-

⁵ Turing, A. M., *Computing machinery and intelligence*, Mind, octubre 1950 (rep. en C. & T.).

quina", Wiener), llevados a cabo durante la Segunda Guerra Mundial, y uno de sus principales objetivos fue el estudio de los sistemas de regulación desde el punto de vista *funcional*, abstraéndose así de sus propiedades físicas, ya fuesen organismos biológicos o técnicos. En 1948, Wiener publicó su obra "Cybernetics", una de las básicas en la materia. Se realizaron algunos experimentos prácticos y se crearon gran cantidad de teorías sobre el comportamiento y el aprendizaje.

Luego, ante la aparición de las máquinas electrónicas y la posibilidad de probar las teorías mediante la simulación, los estudios sobre el tema se dividieron en varios caminos. Por un lado, quienes siguieron elaborando teorías y experimentando con los principios básicos —sistemas adaptativos, reguladores por retroalimentación, etcétera—, con la idea de que a partir de una cantidad de circuitos ricamente interconectados se podría llegar a obtener comportamientos de alta complejidad, evolución o lograrse mediante el aprendizaje y "adaptación". En este grupo es fundamental el estudio de la capacidad de regulación, definida por Ashby⁶ como la capacidad para disminuir la influencia de las variaciones externas sobre el estado del sistema. Es un principio básico a todo organismo; difícil sería la vida sin la posibilidad de protegerse ante las perturbaciones exteriores: los seres humanos tenemos mecanismos de control de nuestra temperatura interior (regulación activa) un barco tiene estabilizadores que evitan el efecto de las olas (regulación pasiva), un termostato controla la temperatura de una habitación, etcétera. Los más dedicados a este aspecto experimental se han basado generalmente en la observación de estos principios en los seres vivos, convencidos en su superioridad actual, y han dado lugar a la ciencia denominada "biónica". En realidad no difieren mucho de los que han enfocado el problema desde un punto de vista más teórico, ya que, por una parte, los modelos matemáticos se basan generalmente en observaciones de la realidad, y por otro lado es poco lo que se conoce acerca de los principios básicos en la biología y la neurofisiología; razones por las cuales los estudiaremos conjuntamente en la sección 3.4 con mayor detalle.

Es así que aunque este primer grupo no ha alcanzado resultados que podemos clasificar quizá como trascendentales, han aclarado, sin embargo, gran cantidad de procesos elementales de aprendizaje y reacciones de respuesta similares a los reflejos condicionados. Dentro del enfoque teórico a que nos referimos, es interesante el paralelismo trazado por Ashby entre "inteligencia", "poder de selección" y "regulación"; comienza por

⁶ Ashby, W. R., *An introduction to cybernetics*, Wiley, 1956.

establecer los conceptos de *amplificación* de la regulación mediante la "suplementación", obteniendo mayor capacidad recurriendo a fuentes económicas o accesibles. Así llega a la conclusión de que el cerebro cumple las funciones de obtener esta "suplementación" de capacidad de decisión a través de la experiencia, obteniendo así una capacidad de regulación mayor que la obtenida por la información genética, siendo ésta quizá la razón por la cual los seres con sistemas nerviosos más desarrollados son los que tienen más posibilidades de adaptarse y sobrevivir. Siendo la capacidad de resolver problemas equivalente en cierto modo a la de elegir o seleccionar una solución entre varias (por ejemplo, elegir una estrategia en el ajedrez), se establece una equivalencia entre "poder intelectual" y "poder de selección", y siendo el último amplificable (ver ejemplos en la obra de Ashby), se deduce la posibilidad de la "amplificación de la inteligencia".

Un segundo grupo de investigadores prefiere tratar el problema de la inteligencia como una "caja negra", tratando de obtener conclusiones, sobre la base de un estudio de "entradas y salidas"; que se traducen en teorías psicológicas específicas, o sea, en un intento de entender el funcionamiento lógico del cerebro para así aplicarlo a la síntesis de sistemas artificiales y a la misma terapia del sujeto. Una gran parte de estos esfuerzos resultaron en teorías "probabilísticas" del aprendizaje, y está relacionada más con los temas de simulación (capítulo 2) que con los demás grupos de inteligencias artificiales.

El último grupo es el que trata de crear una ciencia de "inteligencia artificial" por su cuenta, sin prejuicios hacia lo "simple, biológico o humanoide" (Minsky). Actualmente no es posible distinguir bien las diferencias entre los dos últimos campos (incluso algunos investigadores realizan trabajos difíciles de clasificar en este sentido), debido a que ambos se basan en la capacidad del hombre para resolver sus problemas como el modelo más importante, ambos usan de la computación para la simulación, y existe un estrecho contacto entre ambos grupos. Sin embargo, de acuerdo con Minsky, es de esperar una sustancial divergencia en el futuro, en cuanto al uso por el último grupo de ciertos formalismos lógico-matemáticos no característicos al ser humano.

Otra diferencia importante de resaltar es que al que trabaja en simulación le interesan tanto los aspectos positivos como los negativos del comportamiento y centra su atención en casos particulares, mientras que el investigador en I.A. trata de generalizar las características positivas que encuentra en el hombre.

Para entender la necesidad de investigaciones que no par-

ten de la simulación es quizá útil el siguiente ejemplo de Armer: cuando el hombre quiso volar, estudió primeramente las aves, siendo éstas una evidencia de que "algo más pesado que el aire" podía volar. Sin embargo, grandes fueron los esfuerzos que se hicieron para tratar de obtener máquinas voladoras que moviesen las alas como los pájaros y no se consiguieron resultados prácticos: sólo al descubrirse los principios básicos de la aerodinámica se pudo construir el primer avión.

Como ya vimos, un "problema" existe cuando se debe efectuar una "elección"; un modo trivial de conseguir entonces "máquinas que resuelvan problemas" sería enumerar las posibles soluciones, generándolas por un método apropiado, y probar una a una. Cuando esto es posible, y tenemos la certeza de que la solución será generada en un intervalo de tiempo finito, contamos con un algoritmo. El papel de los algoritmos en las inteligencias artificiales es relativo ya que, como lo hicimos notar anteriormente, es común la tendencia a desdeñar un comportamiento totalmente comprendido como "inteligente".

Como un algoritmo corresponde generalmente a una formalización, nos referiremos a estos métodos en la sección (3.4). Sin embargo, imaginemos el caso en que no conocemos el problema suficientemente; si existe incertidumbre deberemos hacer uso del método básico en las investigaciones de inteligencia artificial: el *método heurístico*, un método que no nos dará garantías de encontrar la solución óptima, pero en el que podremos confiar en buen grado. En las próximas secciones nos referiremos extensamente al método heurístico, pero señalemos aquí un detalle fundamental: el método heurístico es tal mientras no exista un conocimiento completo; en cuanto percibimos totalmente la situación y el procedimiento nos da certidumbre sobre las soluciones, habremos *convertido* el método en uno algorítmico. Un ejemplo claro es el dado por el ajedrez: conocemos en cualquier momento sólo parte de la situación, no el resultado de todas las acciones posibles, y nos basamos en criterios parciales tales como "controlar el centro del tablero", "proteger una pieza bien ubicada", etcétera, que sin embargo *no* son infalibles. Si nos fuese posible analizar todas las movidas posibles hasta el final del partido, podríamos encontrar, si existe, una estrategia ganadora, y tendríamos así un método algorítmico.

El párrafo anterior ya nos sirve para desligarnos de otro preconcepto bastante común: el creer que la computadora puede demostrar un comportamiento "inteligente", ya que resolvería los problemas considerando todas las posibilidades de solución; esto no sucede más que en casos muy simples, ya que inmediatamente se plantea un complicado problema de tipo

combinatorio. Estas resoluciones por "fuerza bruta" casi no tendrán aplicación en la práctica; en el ajedrez, por citar como ejemplo, la cantidad de estrategias posibles se calcula en 10^{120} ; una resolución por este método no será posible nunca. La solución residirá entonces en la *exploración selectiva* de posibilidades.

Para experimentar los complejos métodos heurísticos, se recurre generalmente a áreas bien definidas, dominios regulares donde la "inteligencia" y el "razonamiento simbólico" juegan un papel importante. La razón de esto es que se necesita ignorar las complicaciones de detalle de modo tal de poder concentrarse en los verdaderos problemas. Pero esto *no* significa que los resultados se apliquen solamente a estas áreas, del mismo modo como los físicos estudian los "gases ideales" y luego aplican sus resultados a la realidad. Entre estas áreas se encuentran las matemáticas (diseño de programas para probar teoremas en lógica, álgebra, geometría, etcétera) y los juegos (ajedrez, damas, etcétera), dominios donde con el menor esfuerzo de programación se pueden obtener grandes resultados, son áreas que desarrollan gran complejidad a partir de estructuras simples.

Quizá este propósito no haya sido bien comprendido por los críticos de las investigaciones en inteligencias artificiales, quienes llegaron a la conclusión (falsa, claro está) de que una máquina sólo puede resolver problemas de dominios "formales". De cualquier manera, la presentación en este capítulo de computadoras con capacidad de percepción, de hablar en lenguajes naturales, de resolver problemas de "sentido común" y de obtener altas calificaciones en tests de inteligencia (el programa de "Analogías" de Evans, por ejemplo), creemos alcanzará para desmentir estas afirmaciones.

Y si bien aun estos programas parecieran estar restringidos a problemas particulares, esta especialización es sólo aparente en cuanto a forma, y sólo es de contenido; cumplen su propósito de permitir el estudio de los métodos de manejo de la información, y "las herramientas que eventualmente nos darán generalidad son en sí mismos métodos rigurosamente estructurados y deben ser primero estudiados y probados en las usuales áreas de problemas bien entendidas (y consecuentemente estrechamente definidas)". (Minsky, op. cit.).

Otro de los principios generales en que se basan las investigaciones proviene de la consideración de que la mente de adulto es en realidad resultado de un estado inicial, una educación y una constante interacción con el medio ambiente, de modo que su conducta "inteligente" está dada por sus mecanismos primitivos más el resultado de la experiencia. Suponiendo una pequeña estructura inicial, y teniendo en cuenta que es más fácil

“educar” que “construir”, gran parte del esfuerzo de programación será eliminado diseñando sólo las estructuras fundamentales (agregando, opcionalmente, gran parte de definiciones y métodos de inferencia a priori), y permitiendo a la máquina interactuar con el resto del mundo (aquí son sumamente importantes los métodos de reconocimiento y clasificación de configuraciones, en cuanto partes integrantes de la percepción), obteniendo así su “inteligencia”.

El simple método de enseñanza por “recompensa y castigo” puede ser en general implementado fácilmente, y lo encontraremos repetidas veces en nuestro estudio. Sin embargo, desde el punto de vista teórico adolece de un defecto fundamental, la *máxima* cantidad de conocimiento que se podría adquirir no podría sobrepasar (medida en cantidad de información) al número de “recompensas y castigos” total. Una alternativa explotada grandemente en los últimos años consiste en intentar que aprenda conceptos más generales, “entendiendo” (3.3) las observaciones que se le hagan en un lenguaje simbólico, incluyendo la obediencia a las sentencias imperativas; un aprendizaje conceptualmente distinto al de “ajustar parámetros”, consistente en cambio en la modificación de los mismos métodos de resolución de los problemas y representación de los datos, lo cual estaría más de acuerdo con el aprendizaje de “alto nivel” observable en los seres humanos.

Un sistema que demuestre “inteligencia” necesariamente deberá ser capaz de crear por su cuenta nuevos métodos de resolución de problemas. Estos métodos, sin embargo, no serán eficientes a menos que sean suficientemente generales, ya que es muy remota la posibilidad de que se presente dos veces exactamente la misma situación, que se repitan idénticamente los factores presentes en el momento en que se obtuvo el método. Para conseguir esta generalización, se debe contar tanto con métodos adecuados para distinguir *analogías* entre problemas como para la inferencia *inductiva*.

La cuestión de las analogías entre distintas situaciones depende fundamentalmente del reconocimiento de “configuraciones”, de la capacidad de reconocer similitudes entre grupos de objetos y aun entre métodos. Es de suponer que una “inteligencia”, ya sea humana o no, debe poseer gran cantidad de métodos disponibles, y para su aplicación contar con un modo de estimar el “tipo” de problema y el medio de resolución más adecuado para cada “clase” particular. Muy ligada a esta capacidad, está la de poder “describir” bien las situaciones, de innegable valor para establecer las “similitudes”. Estas descripciones pueden ser efectuadas mediante “listas de propiedades”

cada una de las cuales indica la presencia (y aun el grado) de alguna cualidad particular, o por medio de sentencias descriptivas en algún lenguaje simbólico.

La posibilidad de generalizar a partir de casos particulares, la inducción incompleta, base de las ciencias experimentales, es otro de los comportamientos, además de la posibilidad de distinguir las clases de situaciones, que seguramente deberá formar parte de la "inteligencia artificial". Su implementación depende también en gran medida de la posesión de buenas descripciones y modelos, ya que comparando las propiedades comunes a varios casos, se pueden encontrar las características generales, ya sea de un tipo de problema, de métodos o de estados del medio ambiente.

Esta propiedad de "generalizar" está estrechamente ligada con la capacidad de "predicción" del comportamiento futuro del mundo exterior basándose en su estado actual; es el medio que permitirá a nuestro sistema describir las "leyes" que rigen el comportamiento de la naturaleza o, al menos, ciertos tipos de eventos.

Una interesante formalización de la inferencia inductiva está dada por los trabajos sobre descubrimiento de *gramáticas*. La idea detrás de este método está en la posibilidad de la existencia de un lenguaje formal (ver capítulo 1) de tal modo que todas las posibles sentencias del lenguaje, obtenidas a partir de la aplicación reiterada de las reglas gramaticales, describan (simbólicamente) las situaciones posibles, mientras que los sucesos "imposibles" se corresponderían con las sucesiones de símbolos del alfabeto que no formen parte del "lenguaje". Entonces una vez formalizada el área en estudio, el problema se reduce al de buscar una gramática adecuada. Una vez obtenida, estas reglas constituyen de por sí las "leyes" y relaciones que gobiernan los tipos de sucesos en estudio, pudiendo entonces, inclusive, derivarse el lenguaje completo obteniéndose sentencias que no aparecieron durante el "aprendizaje".

El primer problema, la formalización, podrá tomar diversos caminos. En ciertas áreas (álgebra y lógica, por ejemplo), hay lenguajes naturales (ecuaciones algebraicas y lógica formal, respectivamente), mientras que para otras deberemos quizá basarnos en estructuras sintácticas comunes de sus descripciones en lenguajes simbólicos apropiados. Y en cuanto al segundo problema, la gramática en sí, se han conseguido ya algunos algoritmos para obtener las gramáticas "más probables" a partir de ejemplos tomados al azar de un lenguaje, siendo el trabajo actual en el área bastante intenso.

En caso de ser un problema muy complicado, puede no exis-

tir un método apropiado de solución directa, pero puede ser posible en ocasiones el "analizarlo" convenientemente, decidiendo una *planificación* para su resolución. Esta planificación podría consistir en la formulación de una secuencia de "subproblemas" para ser resueltos con el fin de obtener una solución para el problema original. Cada uno de estos subproblemas podría a su vez ser resuelto directamente o mediante una nueva aplicación de este método. Este procedimiento recursivo puede ser llevado a cabo mediante el estudio de "medios y fines" como en el caso del GPS (3.1.5), o mediante matrices que relacionen las características del problema con los métodos que se poseen.

El último de éstos, debido a Minsky⁷, consiste en el uso de una matriz donde tanto filas como columnas corresponden a "tipos" de problemas (ya determinados, por ejemplo, por un sistema de reconocimiento de configuraciones), y en la que en la posición (i, j) se encuentra una lista (posiblemente *vacía*) de "métodos" que permitan transformar problemas del tipo *i* en otros del tipo *j*. Si no podemos encontrar un método *directo* para transformar el problema que poseemos en otro (la etapa siguiente del plan, por ejemplo, o uno para el que ya conocemos la solución), formamos la "matriz de conexiones" (con *unos* cuando existe al menos un método, y *ceros* si la lista es vacía), y mediante su elevación a potencias se investiga la posibilidad de caminos de más de un paso (para lo cual existen variados algoritmos en programación dinámica, etcétera), lo que se podría interpretar como indicación de la existencia o no de "métodos compuestos" (aplicación sucesiva de los métodos listados) para pasar de un problema de tipo *i* a uno de tipo *j*. El principal inconveniente con este método está, por supuesto, en la obtención de la matriz adecuada, incluyendo la clasificación de los problemas.

Un estudio más detallado de los métodos heurísticos, el reconocimiento de configuraciones y los métodos de representación los llevaremos a cabo en las primeras tres secciones del capítulo, mientras que en la cuarta volveremos sobre los métodos de la biónica, algunos métodos algorítmicos y generalizaciones de algunos de los métodos heurísticos presentados. Son muchas las objeciones que se podrían interponer a los temas tratados, por lo que trataremos de adelantarnos con algunas observaciones.

En primer lugar, volviendo al tema de la existencia del programador y olvidando momentáneamente el argumento bio-

⁷ Minsky, Marvin, *Steps toward artificial intelligence*, Proc. of the Institute of Radio Engineers, enero 1961.

lógico antes presentado, ¿es posible predecir todas las posibles acciones de un programa y por lo tanto de una "inteligencia artificial", conociendo su estado inicial? ¿pueden llegar a "tomarnos por sorpresa", o harán exactamente lo que les hayamos indicado que hagan, y nada más? y ¿cómo poder afirmar que han "creado" algo por su propia cuenta, si alguien les especificó sus acciones en detalle? Este último punto es altamente ambiguo, dependiendo del significado de la palabra "crear"; pero aun suponiendo el significado intuitivo, baste contestar con una pregunta: ¿en qué podemos basarnos para afirmar que *nuestros* actos de "creación" e "invención" no siguen ciertas reglas y principios generales estrictos a partir de nociones adquiridas con la experiencia? La segunda pregunta, por otra parte, no configura en lo absoluto una disyunción exclusiva: si bien es perfectamente cierto que harán exactamente lo que se les haya indicado y "nada más", esto no implica que conozcamos las consecuencias últimas de su accionar, como lo veremos en el párrafo siguiente.

Es así que nos queda sólo el primer interrogante. Para contestarlo se debe, ante todo, recordar que lo que se proponen son programas donde se han especificado acciones tendientes a una interacción con el medio ambiente, una modificación debida a la experiencia. Y si es cierto que en principio es posible prever los estímulos a que serán expuestos no sería entonces menos cierta la afirmación determinista de que podríamos conocer perfectamente el futuro basándonos en el estado, en un momento dado, del universo entero (cosa, por otra parte, teóricamente imposible, si nos atenemos al principio de indeterminación), y similarmente sería posible conocer todo el comportamiento futuro de un ser humano sobre la base de un análisis completo de su estructura inicial y a un similar análisis de las experiencias. Trazando un paralelismo más abstracto, tal capacidad de predicción equivaldría a la posibilidad de conocer, dado un sistema axiomático suficientemente complicado, la forma de todos los teoremas posibles.

Por otra parte, no se puede suponer que el programador tendrá capacidad para manejar las cantidades de información y los complejos procesos que puede llevar a cabo la computadora, por lo que sería falso inferir que ésta no podrá exhibir más inteligencia que su diseñador. Una de las razones quizá más claras de esta complejidad se puede encontrar en el carácter modular de los grandes programas: se trata en realidad de un conjunto de programas más pequeños o "rutinas", que tienen por fin realizar distintas acciones, organizados jerárquicamente, de modo tal que uno encargado del "control" dirigirá sus accio-

nes, pero éstos podrán interactuar e inclusive aplicarse a sí mismos recursivamente, dando lugar a un comportamiento de una complejidad inimaginable, una situación quizá no muy diferente de la que se da en el cerebro humano, mediante varios mecanismos y procesos que intervengan en lo que englobamos como razonamiento, con diversos "centros de memoria" que trabajen con algún tipo de acceso mutuo. Refiriéndose en particular a un procedimiento de eliminación de ambigüedades en un programa que analiza el texto en inglés (programa "SIR", ver 3.3), Minsky da una imagen muy explicativa de lo anterior:

"Es preferible considerarlo no como un procedimiento incluido directamente en una secuencia pre-especificada de resolución, sino más bien como una corte de apelaciones, a la cual el programa principal consulta cuando encuentra inconsistencias o ambigüedades. Cuando escribimos un programa extenso, con gran cantidad de tales cortes, cada una con la posibilidad de llamar a las otras para solicitar ayuda, se vuelve inútil seguir considerando el programa como una "secuencia de instrucciones". Aun cuando el programador ha consignado los principios 'legales' que permiten tales 'apelaciones', no tendrá más que un vago conocimiento respecto de cuándo y dónde, en el curso de un programa, estas rutinas se llamarán entre sí. Resumiendo, debemos imaginar a las partes de los grandes programas heurísticos como especificaciones para individuos de pequeñas sociedades. Por mucho que lo intente, el programador en general será incapaz de antemano todos los detalles de sus interacciones."

Otra importante observación que se hace es respecto de la continuidad de los procesos en los seres humanos: ¿cómo simular un comportamiento continuo con máquinas discretas? Caben aquí varias respuestas. Ante todo, ya aclaramos que no es la imitación el propósito, sino la obtención de sistemas que demuestren un cierto grado de inteligencia, y a priori no hay ninguna indicación de que ésta esté tan estrechamente ligada a la continuidad. Por otra parte, es nuestra creencia de que la continuidad depende en gran medida del grado de resolución con que se observa un sistema, con lo que se tornaría bastante relativa la continuidad de nuestro sistema nervioso central, además de la observación de que nuestra vista, por ejemplo, discretiza la información al proyectarla sobre los receptores colocados en la retina. Por último, señalemos que es posible el agregado de dispositivos "analógicos" de tipo continuo a una computadora si esto fuese necesario, cosa que ciertamente dudamos, excepto para el caso de unidades periféricas.

En cuanto a la variedad de conductas posibles que puede

demostrar el ser humano y que aparentemente no se encontraría en los sistemas artificiales, los programas que describiremos cubren en conjunto gran parte de los comportamientos considerados privativos de una "inteligencia". Si bien cada uno en sí demuestra una marcada especialización eso se debe en gran parte, como ya lo hicimos notar, a que fueron diseñados con el fin primordial de estudiar los principios en sí, pero es de esperar que, con mayores disponibilidades de memoria y otras facilidades técnicas, se puedan complementar y lleguen a formar parte —al menos sus principios conceptuales— de sistemas más "generales".

Gran cantidad de argumentos que se levantan en contra de la posibilidad de "máquinas inteligentes" provienen tanto de la ignorancia de los progresos alcanzados como de la costumbre al uso de sistemas de computación pequeños y limitados por parte de muchos programadores, y aun de razonamientos basados en premisas completamente falsas, muchas de las cuales parten de atributos supuestos de la inteligencia. Uno de los más usuales es el de "una computadora no puede ser inteligente porque no puede sentir, no puede escribir poemas ni crear música".

Pero acá caben varias acotaciones: Ante todo, ¿pueden *todos* los seres humanos escribir poemas y crear música? Si se trata de "crear música" desde un punto de vista abstracto, también hay computadoras que lo hacen, y en cuanto a la creación artística en sí, son realmente contados los seres humanos que poseen tales aptitudes. Sin embargo, no llamamos al resto de los seres humanos "idiotas" sólo porque no son capaces de tener ciertas aptitudes especiales. Si bien es difícil estimar la relación entre los valores estéticos y morales y el área intelectual, debemos aceptar la influencia de la educación sobre los primeros, por lo que desde un punto de vista abstracto, una máquina "educada" también desarrollará una cierta escala de valores. Sin embargo, y para no caer en una discusión de la que seguramente no obtendremos ninguna conclusión, aceptemos la presencia de ciertos factores "inescrutables" en el ser humano. Aun así, el dominio de lo racionalmente alcanzable es amplio, y es muy probable que mucho más de lo que se pueda apreciar a primera vista. En definitiva, dejemos el tema de las "máquinas que sienten" a la ciencia ficción y al futuro; aun así podremos llegar a grandes resultados en "máquinas intelectuales", concentrándonos en los aspectos racionales del comportamiento, en aquellos que se puedan "entender", dejando con este último término abiertas las posibilidades para toda área que se pueda llegar a analizar, ya sea ahora o en el futuro. Los límites de

las investigaciones están aún muy lejos de nosotros y ninguna prueba científicamente válida ni ningún argumento lógico se ha presentado aún para indicar que los hayamos alcanzado.

Por otra parte, aun cuando los estudios recién están en sus comienzos y debamos todavía esperar diez o veinte años para llegar a los resultados finales (si tiene algún sentido hablar de *final*), ya hoy podemos ver las aplicaciones de los “sistemas adaptativos”, los “autómatas que aprenden” en áreas tan variadas y útiles como la traducción automática (aunque por el momento con grandes limitaciones), los grandes bancos de datos, estudios psicológicos y sociales, control automático de procesos, etcétera.

Los que en un principio fueron denominados “autómatas” y “servomecanismos” (el término “autómata” proviene de autonomía, “la realidad regida por una ley propia”, y se aplicó a mecanismos capaces de llevar a cabo sus fines —los de quien lo construyó— independientemente del mundo exterior), hoy amenazan invadir, gracias a sus antes subestimadas posibilidades de autoorganización y aprendizaje, todos los terrenos en donde los hechos sean descriptibles por relaciones causales y lógicas, temas hasta ahora sólo reservados a la inteligencia humana.

En definitiva, en un momento en que el hombre y la civilización hacen de la máquina una necesidad, en que la guerra y la paz, las actividades productivas y la investigación científica dependen en gran parte de la posibilidad de pensar “con” la máquina, se comienza a vislumbrar la posibilidad de las computadoras inteligentes a través de las investigaciones que son llevadas a cabo en los más importantes centros de estudio del mundo. Es una tarea difícil la que hay aún por delante, pero los primeros pasos ya se han dado con resultados aparentemente muy positivos. Es una tarea en la que sólo el trabajo en equipo, el estudio interdisciplinario por parte de matemáticos, psicólogos, biólogos, ingenieros y expertos en lingüística, entre otros, podrá dar frutos finales concretos. Una tarea, por otra parte, que debe ser seguida con atención —y si es posible tomar parte en ella— por los estudiosos de países que no quieran quedar marginados del futuro de la humanidad. Cerramos esta introducción con un fragmento de un reciente informe del grupo de Inteligencia Artificial de la Universidad de Stanford, que realiza sus investigaciones bajo contrato de la Agencia para Proyectos de Investigación Avanzados de la Secretaría de Defensa del Gobierno de los Estados Unidos ⁸.

“La Inteligencia Artificial es el estudio teórico y experi-

⁸ Stanford Artificial Intelligence Project, Memo AIM-117, abril de 1970.

mental de los procesos perceptivos e intelectuales usando computadoras. Su objetivo último es el entendimiento de estos procesos tanto como sea necesario para conseguir que una computadora perciba, entienda y actúe en formas ahora sólo posibles a los seres humanos. La información para este estudio proviene en parte de la observación del comportamiento humano, incluyendo la autoobservación, pero principalmente de las experiencias con programas diseñados para la resolución de problemas elegidos de tal modo de requerir los procesos intelectuales en estudio. En muchos callejones sin salida nos hemos y estamos introduciendo, y muchos errores se están cometiendo. Sin embargo, se está acumulando una colección de conocimientos fundamentales.”

3.1. El método heurístico.

Como ya apuntamos en la introducción, el uso de los llamados “métodos heurísticos” es el que puede, en definitiva, darnos la solución al problema de la inteligencia artificial, en contraposición con los mucho menos efectivos sistemas que se “autoorganizan” y con los métodos propiamente algorítmicos, con menor rango de aplicabilidad. También se hizo notar la relatividad del concepto de “heurístico” dada por la cantidad de información que se posee sobre el asunto en cuestión, un ejemplo de lo cual estaría dado por la resolución de un problema como el siguiente: se tiene un gran número de cajas de similar tamaño y se sabe que cada una contiene cierta cantidad (desconocida) de un cierto artículo; un rápido método heurístico para conocer la cantidad total del artículo será abrir una de las cajas, contar la cantidad contenida, y multiplicar por el total de cajas; será “heurístico” porque no sabemos *con seguridad* que así se obtenga la respuesta correcta, sino que nos basamos en la suposición (debida a nuestra experiencia con casos similares) de que los artículos se empaquetan en cantidades fijas. Sin embargo, para quien llenó las cajas, si lo hizo de tal modo de equidistribuir el artículo, el mismo método será un algoritmo con el cual sabe que obtendrá una respuesta correcta.

Otra precaución que se debe tomar al manejar el término “método heurístico” (equivalentemente, una “heurística”) se debe al uso diferente de esta palabra según el interlocutor. Mientras que es común denominar con estos términos a todo procedimiento “informal”, y aun asociado a la experimentación y la inducción incompleta (“heurística es el arte de la investigación según métodos científicos”, Enciclopedia Herder, Barcelo-

na), Minsky (1961) nos previene de usarlo como opuesto a lo totalmente efectivo, definiéndolo simplemente como cualquier método "relacionado con el mejoramiento del desempeño en la resolución de problemas" y concretándolo más posteriormente (1968) como "característica que permite perfeccionar la eficacia de resolución o el campo de aptitud del sistema. Estas (características) se extienden desde artificios diseñados ad-hoc para tipos especiales de problemas hasta los principios más generales de administración eficiente y asignación de recursos", lo cual se contradiría, en cuanto a la generalidad, con considerarlo como "un procedimiento provisional y razonable cuyo propósito es descubrir la solución del problema particular en consideración" (Gelernter, 1959, quien además da una interesante imagen de una "heurística" como un "filtro" situado entre el "generador" y el "evaluador" de soluciones para una clase dada de problemas).

Fuera del campo específico de la inteligencia artificial, los métodos heurísticos juegan también un papel importante. Ejemplo de esto son algunos métodos de optimización en grafos, como la búsqueda de permutaciones óptimas mediante la subdivisión previa en subgrafos fuertemente conexos, optimizando luego cada uno algorítmicamente, u otros métodos "semianalíticos" o "aproximados" descritos por Kaufmann, quien, por su parte, define heurístico a un "método que no puede aceptarse del todo rigurosamente, pero que proporciona resultados suficientes para la práctica".

En cuanto al significado de "programación heurística", que denotaría en principio a la técnica de programar usando métodos heurísticos, se plantean también ambigüedades, dándose incluso el caso (Garvin et alii, 1963) de reducirlo a "un sistema de programación de computadoras de modo tal que se pueda decir que la máquina 'aprende' mediante un procedimiento de ensayos sucesivos".

Por nuestra parte, preferimos la definición de Feigenbaum y Feldman¹ (1963): "una heurística (regla heurística, método heurístico) es una regla empírica, estrategia, artificio, simplificación o cualquier otro recurso que limite drásticamente la búsqueda de soluciones en problemas con un amplio espacio de soluciones. Estos métodos no garantizan soluciones; todo lo que puede afirmarse sobre un método heurístico útil es que ofrece soluciones adecuadas la mayoría de las veces", aunque sería preferible que el lector capte la esencia del concepto por sí mismo mediante el análisis de los ejemplos que presentamos.

¹ Feigenbaum, E., Feldman, J. (edits.), *Computers and thought*, McGraw Hill, 1963.

La aplicabilidad de un método heurístico existiendo un algoritmo adecuado depende fundamentalmente del costo y el tiempo asociado al último, ya que si bien un algoritmo nos asegura una solución en un intervalo "finito" de tiempo, éste puede ser suficientemente extenso como para impedir su realización práctica, como en el caso expuesto del ajedrez; se debe entonces decidir sobre la base del esfuerzo requerido por cada método y a la proporción calculada de soluciones capaz de encontrar, asignando por ejemplo un costo por cada intento frustrado.

Analizaremos a continuación la resolución de problemas de algunas áreas particulares (juegos, teoremas, etcétera) como así también algunos enfoques de tipo más general, basados en estos primeros intentos. Posteriormente (3.4) presentaremos otras generalizaciones y formalizaciones de métodos ya considerados, que tengan quizá más posibilidades de poseer una gran capacidad y campo de aplicación.

3.1.1. Máquinas que juegan.

El campo de los juegos es uno de los más interesantes y a la vez instructivos en los que se pueden experimentar las ideas y obtener otras nuevas en la aplicación de métodos heurísticos. Fundamentalmente, es un dominio bien definido donde los fines son claros, dados por las reglas del juego, y donde los procesos de decisión pueden alcanzar una extrema complejidad. El razonamiento que lleva a cabo un jugador, tiene en cuenta una cantidad inmensa de factores, muchos de ellos quizá a nivel subconciente, que lo llevan a ser altamente selectivo en las posibilidades que explora concientemente. Es una área en la que, por otra parte, a la capacidad de aprendizaje se le puede asignar un papel fundamental. Y, finalmente, uno de los principales atractivos en el diseño de máquinas que jueguen está indudablemente en la posibilidad de competir que se plantea de un modo natural entre el hombre y la máquina, oportunidad para comparar las respectivas capacidades, especialmente con los actuales equipos, que permiten una confrontación directa, "On-line", económica gracias al sistema de tiempo compartido. Entre las posibilidades más interesantes que se plantean para el futuro, está el estudio de la importancia del comportamiento adaptativo en varios juegos (el poker y el truco, por ejemplo), donde es importante un cierto conocimiento de la "psicología" de nuestro oponente. Ya se han dado algunos pasos con implementaciones de "jugadores

artificiales de poker"¹, pero falta mucho por investigar para llegar a resultados positivos, cuya aplicación a otras áreas tiene posibilidades insospechadas (considerar, por ejemplo, la aplicación comercial en publicidad o en presencia de competencia).

No es de nuestro interés el análisis de los programas en cuanto jugadores en sí, sino desde el punto de vista del aprendizaje y los métodos. Es así que no resultan comprendidos en el área de "inteligencias artificiales" los numerosos programas existentes para juegos tales como el ta-te-ti, etcétera, donde un programador conocedor de las estrategias óptimas haya diseñado un algoritmo que permita a la computadora ganar todos sus juegos (o al menos empatar, si la estrategia óptima sólo nos asegura eso) siendo en cambio de consideración los casos, por ejemplo, de programas "probabilísticos", donde la experiencia indica el mejor modo (¡aunque no necesariamente infalible!) de jugar. No interesan mayormente, por otra parte, los resultados de la clásica teoría matemática de los juegos (Von Neumann, etcétera), ya que si bien la mayoría de los juegos considerados son de suma cero e inclusive con finitas estrategias, es prácticamente imposible, debido a su cantidad, establecer las matrices correspondientes. Sin embargo, como se verá, se aplica el principio del minimax repetidamente a situaciones locales, a "submetas" que eventualmente permitan ganar el juego en cuestión.

En ciertos programas, cuyo principal exponente es el que Samuel desarrollara a mediados de la década del 50 y que aún hoy continúa experimentándose y mejorándose (especialmente después de la implementación del procedimiento de agrupación en factores, publicado en noviembre de 1967), el énfasis está puesto en las técnicas de aprendizaje, fundamentalmente en el ajuste de parámetros, mientras que en otros se tiende quizá con preferencia hacia los principios de la resolución de problemas en sí. Es común a casi todos los programas recientes, además, el uso de diversas técnicas de búsqueda en árboles de soluciones ("alfa-beta", "gama", etcétera), derivadas de las análogas de la investigación operativa (programación dinámica, "branch and bound", etcétera).

La mayoría de los juegos, tales como el ajedrez y las damas, ofrecen un número finito de alternativas de juego para cada uno de un número finito de posiciones posibles de fichas en el tablero, y pueden ser descriptos totalmente por un árbol (tipo especial de grafo donde no existen ciclos), existiendo en principio un procedimiento algorítmico para evaluar las alternativas

¹ Waterman, D. A., *Generalization learning techniques for automating the learning of heuristics*, en *Artificial Intelligence*, Vol. 1, 1970.

que se presentan a partir de las posiciones terminales (ya sea victoria para alguno de los oponentes o un empate), que da como resultado la estrategia óptima (aquella que da la mejor “recompensa”) a seguir: el procedimiento, el minimax (*minimizar-maximizar*) de la teoría de juegos, consiste en considerar alternativamente las movidas propias y del adversario.

Siendo el método fundamental, conviene estudiarlo mediante un ejemplo. En la Fig. 1, se supone una partida de ajedrez

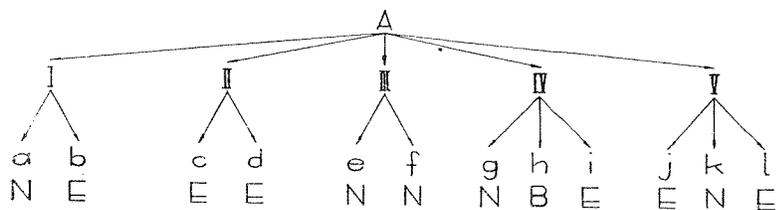


Figura 1.

donde Blancas debe mover a partir de A contando con 5 alternativas: I a V, a las cuales quien juega Negras puede replicar con ‘a’ o ‘b’, ‘c’ o ‘d’, ‘e’ o ‘f’, ‘g’, ‘h’ o ‘i’ y ‘j’, ‘k’ o ‘l’ respectivamente. Supongamos además que, en la figura, B significa que Blancas gana, E un empate y N que gana Negras. El análisis del minimax consiste en lo siguiente: si Blancas juega I, de acuerdo con la movida de Negras se dará que empate o pérdida, si juega II, cualquiera que sea la respuesta de su contrincante se llegará a un empate; si juega III, perderá seguramente, y en caso de jugar IV o V corre el riesgo de que su enemigo juegue ‘g’ o ‘k’, con lo que habrá perdido el partido. Suponiendo que quien posee las Blancas no subestime demasiado a su oponente, difícilmente corra el riesgo de jugar IV (lo que equivaldría a “jugarse” por entero, esperando que el contrario no repare en su jugada ganadora y sólo vea la posibilidad de ‘h’, que lo lleva a perder), y por lo tanto jugará seguramente II, con lo que se habrá al menos asegurado un empate. Este procedimiento se puede hacer ahora recursivo, otorgando ahora a “A” el valor óptimo encontrado, (E), y repitiendo el razonamiento a un nivel superior (una vez calculados todos los valores al mismo nivel que A), ascendiendo así hasta llegar a la posición en que nos encontramos (almacenando apropiadamente en cada paso la mejor estrategia hallada), permitiendo así la decisión óptima.

Este algoritmo, sin embargo, tiene pocas posibilidades de éxito, ya que la cantidad de alternativas a considerar es demasiado grande (se calculan en 10^{120} las combinaciones posibles

sobre la base de una estimación promedio de unas 30 jugadas factibles por posición y unas 40 jugadas como duración de una partida, en el ajedrez, y de 10^{40} en las damas), aunque pueda ser implementado para el análisis de pocas movidas (el análisis de dos jugadas completas en ajedrez implica alrededor de un millón de posibilidades, lo cual sería ya factible) y para resolver problemas simples ("encontrar el mate en dos jugadas", etcétera). Pero para jugar efectivamente, se pueden encontrar métodos heurísticos que hacen sólo uso parcial de esta técnica.

Supongamos ahora que tenemos algún modo (no importa cómo, por el momento) de "calificar" numéricamente cada posición de las fichas en el tablero. Repitiendo entonces el esquema anterior, pero teniendo ahora como nodos terminales del árbol no necesariamente las posiciones ganadoras, perdedoras o de empate, sino simplemente aquellas a las que se pueda llegar luego de una cantidad fija de movidas, se tiene una situación como la esquematizada en la Fig. 2, para una cantidad de dos "jugadas" (cuatro movidas alternativas entre ambos jugadores).

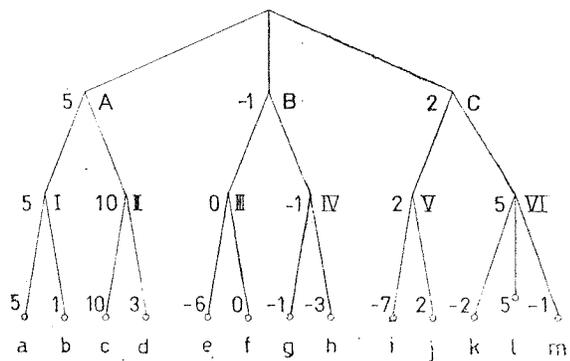


Figura 2.

Supongamos que es nuestro turno jugar y que podemos optar entre A, B o C, teniendo nuestro oponente las alternativas diagramadas, ante lo cual nosotros podríamos, a nuestra vez, responder con las movidas "a" a "m", lo que nos reportaría los "beneficios" indicados en el nivel 4 (supondremos que no hemos aún calculado los números que aparecen en los restantes niveles 2 y 3). Suponiendo que nuestro objetivo sea la maximización de dichos beneficios (ya dijimos que luego veremos cómo se "califican" adecuadamente las posiciones), en principio deberíamos jugar la secuencia A-II-c, con una ganancia máxima. ¡Con el pequeño e inocente detalle de que para ello de-

beríamos requerir la colaboración de nuestro adversario! Pero como en tal espíritu de camaradería no debemos hacer restar nuestras esperanzas, supondremos que nuestro adversario hará todo lo posible por disminuir nuestra posible ganancia cada vez que le sea posible (en este caso, al tomar él su decisión en el nivel 2). Razonemos en consecuencia del siguiente modo: al jugar nosotros en el nivel 3, lo haremos de forma tal de maximizar nuestra ganancia al paso siguiente, por lo que elegiremos *a, c, f, g, j y l* en el supuesto caso de encontrarnos en las posiciones I a VI respectivamente. Ahora consideramos como valor de las posiciones del nivel 3, los correspondientes a las jugadas que consideramos mejores. El próximo paso consiste en suponer que nuestro contrario intentará, estando en una de las posiciones del nivel 2, *minimizar* nuestra ganancia (lo cual supondremos equivale a maximizar la suya), con lo cual elegirá I, IV y V, respectivamente y aplicando una nueva maximización por nuestra parte a los valores así obtenidos para el nivel 2, tendremos que nuestra jugada óptima es "A".

Este razonamiento consiste en la consideración de "qué es lo mejor que podemos hacer si se nos da el peor de los casos", apta para el juego contra un enemigo considerado excelente jugador. Notar muy especialmente que si bien en este caso tendremos un valor *asegurado* de "5", éste es sólo una cota inferior, un mínimo ante cualquier eventualidad, pero si el contrincante no juega a su vez óptimamente, podremos obtener una ganancia o puntaje mayor (considerar, por ejemplo, que ante nuestra "A", éste juegue "II", en este caso jugaremos "C", y habremos obtenido 10 puntos en lugar de 5).

Cabe acotar aquí que, si bien el explicado es el método clásico de obtener los valores de una posición a partir de las posibles continuaciones, es posible realizar importantes variaciones de este esquema. En particular, el programa "M y N" de Slagle y Dixon¹ tiene en cuenta las *M* mejores continuaciones al maximizar y las *N* mejores al minimizar, utilizando la heurística de que es preferible tener varias continuaciones buenas posibles antes que una sola.

El problema se reduciría entonces, en principio, a encontrar un buen modo de asignar un "puntaje" a posiciones intermedias (incluyendo, por ejemplo, un número suficientemente grande para una posición ganadora y uno pequeño para una perdedora). En este momento el lector puede plantearse porqué, si se posee un tal medio de evaluación, no se "calificó" directamente las posiciones del nivel 2; esto se debe justamente a la imperfección

¹ Slagle, J. L., Dixon, J. K., *Experiments with the M & N tree searching program*, Comm. A.C.M., 13, 3, marzo 1970.

de los métodos de evaluación "estáticos", que hacen necesaria una simulación de varias movidas para disminuir así, en cierto modo, la probabilidad de error, como ya veremos. Otra acotación muy importante en este momento es que como se puede apreciar del último ejemplo, la cantidad de nodos en la parte más baja del árbol depende (de un modo aproximado) geoméricamente del nivel; si el inicial es el "1", se tienen k nodos en el nivel siguiente, k^2 al otro, y así sucesivamente si se supone un promedio de k ramificaciones por etapa. Este crecimiento exponencial es el causante de la imposibilidad de aplicación de los métodos algorítmicos ya que la cantidad de alternativas crece hasta sobrepasar los límites del almacenamiento disponible. Es así como somos conducidos a la necesidad de poseer métodos efectivos de "búsqueda" en árboles, genéricamente llamados en inglés de "poda", ya que se encargan de seleccionar las ramas que son útiles.

Esta necesidad de "podar" los árboles de decisión se encuentra constantemente en los programas de inteligencia artificial, asociado a casi todo proceso de solución, y no es muy diferente de lo que realizamos nosotros mismos (quizá a un nivel distinto, ya no conscientemente) cuando en una partida de ajedrez (o cualquier otra situación similar) consideramos sólo las posibilidades "intuitivamente obvias" (¡cuánto misterio encierran estas dos palabras!) que parecieran emerger en cuanto observamos el tablero (o nos ponemos al tanto de la situación). Por otra parte, es interesante, como lo plantea Samuel, especular sobre la posibilidad de que el aprendizaje en los seres humanos incluya el establecimiento de mejoras en sus métodos de búsqueda.

Uno de los primeros en enfocar el problema del ajedrez en relación con las posibilidades de la computación digital fue C. Shannon, quien en 1949 describió, aún sin considerar un programa concreto, varios de los problemas que se presentaban. Propuso la utilización del método del 'minimax' basado en criterios de 'evaluación estática' naturales al ajedrez: una suma ponderada de factores importantes tales como cantidad de piezas de cada jugador, control del centro, etc., otorgando así un "puntaje" a cada posición terminal y luego evaluando la posición actual por el medio indicado, siendo la idea de este método que siendo mayor la cantidad ("profundidad") de jugadas exploradas, menor es la influencia de la regla de evaluación en sí, que es por supuesto muy aproximada. Claro está que si fuese posible llevarlo a su extremo, este método daría el algoritmo anterior, permitiendo un juego perfecto independientemente de la bondad de la regla (con la sola condición que ésta

otorgue un valor suficientemente grande a la posición ganadora, claro está).

Cabe aquí, antes de continuar, trazar un paralelismo entre la discusión precedente y lo explicado anteriormente sobre "planificación" y subproblemas. Es posible interpretar los factores ponderados en menor forma como submetas, que una vez alcanzadas, permitirán alcanzar las metas indicadas por los factores más ponderados (comparar en el ajedrez el tener una torre bien ubicada con la posibilidad de dar jaque, etc.). Otra observación se refiere a la linealidad que se está suponiendo: al considerar los factores numéricos de cada 'propiedad' (movilidad, cantidad de piezas y otras en el ajedrez) multiplicados por un coeficiente o peso, estamos en la práctica procediendo a la evaluación de un polinomio lineal en varias variables (una por cada propiedad en consideración), lo que implica una simplificación muy cómoda, pero también muy perjudicial como lo reconoce Samuel en su segundo artículo.

Del estudio de Shannon se obtiene un método general de trabajo en la creación de máquinas que jueguen, aplicándose su método (con las funciones de evaluación estática correspondientes) a diversos juegos; los detalles que se deben tener en cuenta al analizar los intentos posteriores en este contexto son: cuáles alternativas y hasta qué profundidad se deben considerar y cómo proceder a la evaluación estática en base a la configuración del tablero en un instante dado y a su integración en un valor único para cada movida posible. Una última decisión se debe tomar respecto al modo de elegir la jugada a efectuar, ya sea basándose únicamente en la alternativa que ofrece mayor puntaje o usando algún criterio adicional.

Cuando se considera una cantidad fija de niveles de exploración en el árbol de decisión, se está siguiendo un proceso radicalmente distinto del usado por los seres humanos en situaciones similares; en efecto, cuando en un análisis encontramos una situación de peligro o, por el contrario, muy positiva, continuamos desde esa situación un análisis más profundo. Si se nos hace un examen médico general, por ejemplo, el descubrimiento de alguna aparente irregularidad superficial en un cierto aspecto motivará un examen a fondo de este factor con el fin de descubrir las verdaderas razones de esta irregularidad y determinar su importancia. Del mismo modo, en otro campo, si en el curso de un análisis de una partida de ajedrez descubrimos que con una cierta combinación de jugadas llegamos a una posición en la que estamos dando jaque o se está produciendo un intercambio de piezas, poca será la significación que pueda adquirir una evaluación estática (cantidad de piezas por

bando, por ejemplo) en ese instante, y en cambio tendrá más sentido continuar la exploración hasta una posición más estable. El primero en reconocer la importancia de este razonamiento y de implementarlo explícitamente fue Turing, quien, en base a este concepto, que denominó "posición muerta", diseñó el primer programa detallado (aunque no en un lenguaje de programación, ya que sólo simuló su funcionamiento manualmente) para jugar al ajedrez en base a los principios de Shannon. Sin embargo, el programa de Turing no podía llegar a jugar bien, cometiendo grandes errores; de cualquier modo, no pudo ser experimentado suficientemente debido a la complejidad de la simulación manual.

Continuando con los intentos en el ajedrez, en 1956 se programó la computadora MANIAC I, que si bien no tuvo en cuenta el concepto de "posición muerta" de Turing (consideraba siempre 2 jugadas en su análisis), llegó a jugar razonablemente, pudiendo derrotar a un principiante, pero sin llegar a ser más que regular. Para reducir las alternativas a considerar (le tomaba una hora el análisis de cada movida), se decidió modificar el juego, eliminando los alfiles y dos peones y reduciendo el tablero de ocho por ocho a seis por seis; de este modo, se disminuyó a alrededor de 12 minutos cada análisis (se calculan unas 160.000 alternativas desde cada posición para un análisis de dos jugadas de antelación). La evaluación estática se hacía en base a la cantidad de piezas ganadas y perdidas y a la 'movilidad' que se obtuviese. Años después, parte de los integrantes del grupo de Los Alamos, que habían programado la MANIAC (M. B. Wells y otros), generalizaron el anterior para adecuarlo al tablero normal de 8 x 8, usando ahora la computadora MANIAC II; implementaron además el principio de exploración hasta dos movidas consecutivas que no resultasen en capturas, con un mínimo de 3 y un máximo de 5.

Ya en el programa inicial de Samuel¹ se implementó el principio de profundidad de análisis dependiente de la situación, en este caso aplicado al juego de damas, que presenta una situación mucho más simple que la del ajedrez, aunque suficientemente compleja como para atraer la atención de los investigadores y experimentar las ideas. En principio el análisis se ejecuta hasta una cierta cantidad mínima de movidas de antelación, por ejemplo 3 en las aperturas, y se produce la evalua-

¹ Samuel, A. L., *Some studies in machine learning using the game of checkers*, IBM Journal of Research and Development, julio 1959. No fue éste, en realidad, el primer programa para jugar a las damas, ya que con anterioridad había sido diseñado uno en Inglaterra por Strachey.

ción estática, basada en un polinomio cuyos términos representan diversas características (imposibilidad de movimiento, que equivale al fin del partido, avance, control del centro, movilidad, etc.), a menos que se produzca en alguna posición una de las siguientes alternativas:

- a. la próxima movida es un salto,
- b. la última movida fue un salto,
- c. existe la posibilidad de un intercambio,

en cuyo caso se continúa la exploración desde la posición en cuestión. A las cuatro movidas se evalúa a menos que se produzcan (a) o (c); a las 5 ó más se detiene siempre que la próxima movida no sea un salto (a); a partir de los 11 niveles se detiene si ya existe una ventaja evidente para alguno de los dos bandos, y a los 20 se detiene siempre, debido a limitaciones de almacenamiento, ajustando debidamente las evaluaciones en caso de quedar un salto pendiente. Notar que con este método no se produce una proliferación excesiva de ramas, ya que cada vez que se cumple la condición (a) (lo cual es necesario para pasar de los cinco niveles), las alternativas de continuación que se presentan son pocas. Por otra parte esta es una excelente aplicación del principio de Turing, siendo las evaluaciones realmente "estáticas", y permitiendo la investigación de las continuaciones más interesantes. La jugada ejecutada es la que obtiene una mayor evaluación. Agreguemos que en algunos casos la profundidad de exploración se puede incrementar mediante un artificio que veremos posteriormente.

Aún hay un elemento en este esquema, sin embargo, que resulta poco natural. En efecto, mientras que en algunos juegos (damas, por ejemplo), la cantidad de alternativas a considerar no es —generalmente— muy amplia, en el ajedrez aún un buen jugador no considera más que unas pocas alternativas desde un principio, desechando la gran mayoría. Y si bien el propósito no es imitar al jugador humano, cualquier heurística que éste use será bien venida en el diseño de un programa. Un jugador se atiene normalmente a una cierta cantidad de objetivos, tales como defender sus piezas y atacar, y considera primeramente las acciones que le permitirían concretarlas. Este principio fue reconocido por Alex Bernstein, quien a fines de la década del cincuenta diseñó un programa basado en criterios de selección de alternativas materializados en una colección de subrutinas llamadas "generadores de movidas plausibles", estando cada una de ellas relacionada con alguna característica particular (defender al rey si está en jaque, obtener mayor movilidad, posibilidad de dar jaque, etc.), y operando estos generadores de movidas "razonables" de un modo priori-

tario hasta totalizar siete alternativas posibles ante cada posición. Es de notar que ahora un análisis de dos jugadas lleva la cantidad de continuaciones a explorar a una cifra mucho inferior: 2400, aunque no se obtiene necesariamente una mayor velocidad, debido a lo complicado del programa (unos 8 minutos por movida, con un programa 10 veces más largo que el de la MANIAC I).

El programa de Bernstein² usaba como función de evaluación estática una suma ponderada del total de piezas, defensa del rey, control y movilidad para las Blancas dividida por una similar para las Negras, con lo que si bien se consiguió limitar sustancialmente la cantidad de movidas, se perdió capacidad en cuanto a la profundidad del análisis, en particular en lo que atañe a las respuestas esperadas ante las movidas propias. Estos inconvenientes pueden ser, sin embargo, subsanados en su mayor parte mediante el agregado de otros "generadores" respondiendo a distintos objetivos.

Un enfoque distinto a los anteriores en muchos aspectos es el adoptado por el grupo formado por los doctores Newell, Shaw y Simon, pioneros en las investigaciones en simulación de procesos cognoscitivos y en inteligencias artificiales. En el contexto de su método basado en la existencia de submetas y subproblemas usado en sus diversos programas, crearon en 1958 un programa de ajedrez para la computadora JOHNNIAC³, pudiendo manejar una gran complejidad mediante el uso del IPL-IV, un lenguaje procesador de listas creado por ellos mismos. El programa se escribió usando un vocabulario orientado hacia el ajedrez, lenguaje a su vez implementado en base al IPL-IV mencionado. Las características peculiares de este programa, en especial su estructura modular, lo hacen apto para la experimentación en cuanto a la introducción de modificaciones y agregados.

A diferencia de los programas anteriores, la evaluación no se hace en términos de una linearización con coeficientes constantes, sino que se varía la cantidad de términos significativos de acuerdo a la meta que se desea alcanzar; la decisión final, además, no se efectúa directamente mediante un minimax, sino que atiende a otros criterios, ya sea de eficiencia eligiendo la primera jugada "buena" a considerarse, lo cual implica un considerable ahorro de tiempo, ya sea de sofisticación mediante

² Bernstein, et alii, *A chess playing program for the IBM-704 Computer*, Proc. Western Joint Computer Conference, 1958.

³ Newell A., Simon H., Shaw J. C., *Chess playing programs and the problem of complexity*, IBM Journal of Research and Development, octubre 1958.

la consideración de situaciones especiales (sacrificios de piezas para conseguir un propósito a más largo plazo, etc.), o bien de aplicación de principios especiales tales como jugadas con "doble función", que permiten satisfacer más de una submeta simultáneamente.

El programa consta básicamente de un conjunto de rutinas, asociadas con las diversas metas intermedias que se puedan formular durante un partido (seguridad del rey, balance de material, etc.), cada una de las cuales lleva asociada a su vez una serie de secciones: un "especificador" de la meta, un "generador de movidas", un "evaluador estático" y un "generador de movidas para análisis"; además de rutinas comunes de todo el programa, como las de obtención del valor efectivo para cada movida propuesta y la encargada de la elección final. Cuando le corresponde jugar, el programa realiza un "análisis preliminar", decidiendo en base al estado del juego, las metas que se propone alcanzar en lo inmediato (defensa si hay peligro, etc.), lo cuál se efectúa de acuerdo a la información suministrada por las especificaciones respectivas (no intentará controlar el centro, por ejemplo, si no quedan ya peones disponibles para esta tarea).

Una vez decididas estas submetas, y colocadas en un orden *prioritario*, actúan por única vez los generadores de movidas, que presentan para cada una de éstas una cierta cantidad de alternativas para llevarlas a cabo. Al analizar estas alternativas, estudiando si es necesario varios niveles, no se utilizan ya los generadores iniciales, sino que se establece un proceso recursivo mediante el uso de los "generadores de movidas para análisis". El valor que se asignará a cada movida consistirá de un vector, correspondiendo cada componente a uno de los "propósitos" actuales ordenados de acuerdo a su importancia, e indicando la evaluación de esa posición de acuerdo a los respectivos propósitos. Esta evaluación la realiza a su vez cada generador de movidas para análisis, teniendo en cuenta el concepto de "posición muerta" ya enunciado. Si la posición es de algún modo "inestable" se encargan de generar continuaciones a partir de la alternativa considerada, repitiéndose el proceso hasta encontrarse una posición suficientemente estática (para algunas metas, como la del control del centro, el concepto de estabilidad de la posición carece de sentido, y la evaluación se efectúa directamente). El minimax se efectúa suponiendo a estos vectores ordenados del siguiente modo: si en la primera coordenada (correspondiente como se recordará al propósito más importante en el momento) las componentes difieren, se elige el vector que posee la "preferible", si son iguales se analiza la si-

guiente coordenada, y así sucesivamente. Notar que no se hizo la suposición de que estas componentes sean numéricas, sino solo que se puedan comparar en cuanto a su bondad para la obtención de la meta, pudiendo ser simplemente un indicador de si una propiedad se cumple o no.

Con un ejemplo (Newell, Shaw, Simon), se verá más claro el anterior razonamiento. Supongamos, como lo señala la Fig. 3,

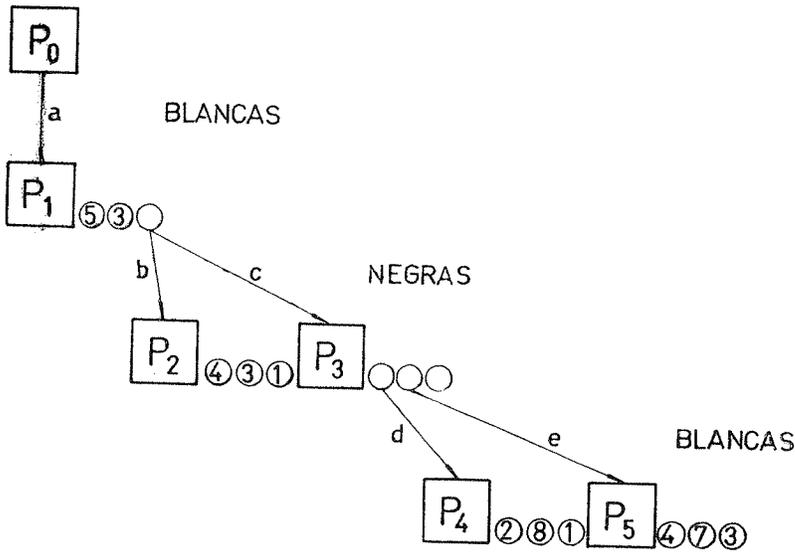


Figura 3.

que el programa, con Blancas, debe mover a partir de la posición P_0 , y se desea analizar una de las alternativas generadas: P_1 , siendo, para simplificar, todos los valores numéricos. Tomando cada meta se intenta evaluar estáticamente P_1 , lo cual se logra excepto para la última componente, obteniéndose "5" y "3" para las dos primeras. Pero no siendo la última estática, se generan dos movidas "b" y "c", por medio de la rutina respectiva. Para la primera, que lleva a P_2 , la evaluación tiene éxito. Para la segunda, la evaluación no puede efectuarse para la primera coordenada, lo cual da lugar a la generación de un nivel más, llegando a P_4 , la cual es estática, y por lo tanto se evalúa. Con la segunda coordenada se efectúa un proceso similar obteniéndose P_5 , también estática, y con la tercera no hay inconvenientes en una evaluación directa. Ahora Blancas maximiza en el último nivel, eligiéndose el valor de P_5 y Negras mini-

miza entre (4, 3, 1,) y (4, 7, 3), tomando (4, 3, 1), que se convierte así en el valor efectivo —reemplazando las dos primeras coordenadas, que ya han perdido significación— de la movida “a” que lleva a P_1 . En realidad el minimax se produce simultáneamente con la generación de alternativas, lo cual en ciertos casos se traduce en un beneficio; si P_5 en el ejemplo se hubiese generado en primer término a partir de P_3 , ya no habría necesidad de generar P_4 , ya que el valor de P_3 será *siempre* mayor o igual que P_5 (Blancas maximizan), y entonces P_2 es elegido sin más, ya que Negras minimizan. Esta es aparentemente, la primera vez en que se aplicó explícitamente un procedimiento de “poda” de árboles.

Dos observaciones finales: la elección definitiva de la alternativa puede basarse, como ya indicamos, en diversos criterios, siendo aparentemente el más útil (especialmente en cuanto a tiempo, uno de los factores más importantes en este programa, que podía tardar, según la complicación de la situación, desde algunos minutos hasta varias horas) el de fijar un “mínimo aceptable” como definición de “jugada buena” y tomar la primera alternativa cuya evaluación sobrepase dicho límite; si no se genera ninguna así, se elige entonces la mejor de acuerdo al orden indicado. Por último, la evaluación con vectores de cantidad variable de componentes, podría considerarse (si se toman todas las coordenadas numéricas) como una lineal con coeficientes que se ajustan según la situación (y de modo tal que la mayor importancia de unas metas sobre otras equivalga a mayores coeficientes). En la misma línea, pero solo con el fin de resolver combinaciones de mate en pocas movidas, se encuadra el programa “MATER” de Simon y Baylor ⁴.

A partir de 1960 A. Kotok y un grupo del M.I.T. desarrollaron una serie de programas de ajedrez, primero para una IBM-704 y posteriormente para la IBM-709 y la 7090 ⁵; analizaba las jugadas hasta 8 movidas o hasta la primera “posición muerta” (lo que sucediese primero), tardando de 5 a 20 minutos para el análisis de cada posición. Ante la comprobación experimental de que dedicaba demasiado tiempo al análisis de jugadas irrelevantes, se decidió incorporar métodos de exploración, utilizándose el “alfa-beta” de McCarthy, el cual detallaremos posteriormente. La evaluación era de tipo lineal, teniendo en cuenta el balance de material, control del centro y desarrollo de las piezas. Su nivel de juego se calculó similar

⁴ Simon, H. A. y Baylor G. W., *A chess mating combinations program*, Proc. Spring Joint Comp. Conference, abril de 1966.

⁵ Kotok, Alan, *A chess playing program*, Artificial Intelligence Project Memo 41, M.I.T., 1962.

al de un aficionado con unas 100 partidas de experiencia, pero cometía de a tanto, gruesos errores. Una variante de este programa fue usada por John McCarthy en un juego contra un similar soviético.

Los programas para jugar al ajedrez se han continuado desarrollando en base a estos principios; en 1965 Burton H. Bloom, del M.I.T., describió un programa basado en evaluadores estáticos y generadores de movidas plausibles, agregando estrategias heurísticas para aprovechar los errores del adversario, estando el programa ya al nivel de un jugador amateur aunque aún bajo el promedio. Con posterioridad, el mismo investigador estudió el problema de la eficiencia de los métodos de exploración en árboles, aplicando los algoritmos obtenidos al problema de ajedrez, que presenta, como ya vimos, un campo altamente propicio para tales experiencias.

Otro esfuerzo en ese sentido lo realizó Arnold K. Griffith, quien en 1967⁶ describió un programa que juega a las damas, y que no realiza una búsqueda extensiva sino que propone un número limitado de alternativas en cada situación; la búsqueda de la jugada óptima se realiza, por otra parte, mediante una exploración aleatoria usando el método "alfa-beta", y se introduce la forma de la función de evaluación mediante un análisis estadístico de juegos publicados.

Es importante hacer notar que, a excepción del programa de Samuel, los estudios de programas complejos para estos juegos se han centrado casi exclusivamente en la calidad de la operación en sí, dejando de lado el aprendizaje; éste es realizado más bien por el sistema programador-máquina introduciendo el primero los cambios necesarios de acuerdo al rendimiento observado. Esto permite concentrarse en el difícil problema que representa la elección de heurísticas de juego apropiadas, además de estar motivado por razones prácticas, ya que un aprendizaje automático requiere generalmente (con las técnicas básicas) un estudio estadístico, y por lo tanto una cantidad demasiado grande de juegos para dar frutos concretos.

En los últimos párrafos nos referimos a los métodos de búsqueda, y a ellos nos dedicaremos ahora. Para comprender su origen, esquematicemos brevemente al método de "Branch and Bound" (literalmente, "Ramificación y Acotación") o "exploración dirigida" usado en la Investigación Operativa. Los problemas a resolver tienen la forma general: "hallar en un cierto conjunto "A" de alternativas la solución óptima entre las que cumplen ciertas restricciones", dado este óptimo, por ejemplo,

⁶ Project MAC Progress Report IV, July 1966/July 1967, M.I.T., 1967.

por un valor numérico asociado a cada solución que se desea hacer mínimo (un costo, por ejemplo) o máximo (una ganancia). Suponemos además que poseemos algún "criterio de separación" en base al cual podemos dividir al conjunto en clases disjuntas que darán las restricciones, (lo cual corresponde en los juegos, por ejemplo, a separar según los posibles objetivos) y —esto es fundamental— que podemos calcular (si el problema es de minimización) un mínimo* para los valores de las soluciones en cada uno de los conjuntos considerados. Se elige entonces la alternativa con menor cota, sea ésta A_1 . Se procede ahora a bifurcar a partir de A_1 , obteniéndose por ejemplo A_{11} y A_{12} (ver Fig. 4), para los cuales se realiza la evaluación del

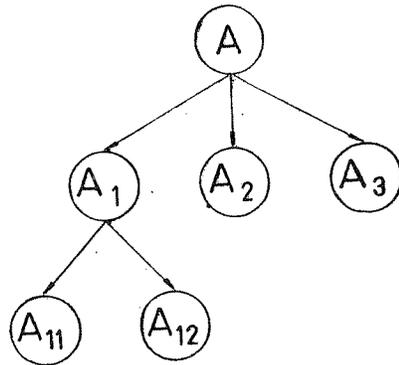


Figura 4.

mínimo. Notemos que cada uno de estos mínimos es, necesariamente, mayor o igual que el de A_1 , ya que A_{11} y A_{12} formaban parte del A_1 . Si ahora se elige el mínimo de éstos, sea A_{11} , esta cota se compara con las correspondientes a A_2 y A_3 ; si sigue siendo menor (el mínimo de A_1 , lo era, pero esta puede ser mayor), se continúa la bifurcación desde A_{12} , caso contrario se sigue desde el menor hallado. Si lo que se desea es maximizar, el método es similar. Luego de varios niveles, correspondientes a otras tantas restricciones impuestas, se llegará al óptimo, sin haber tenido la necesidad de explorar cada una de las continuaciones hasta el final, obviando así un complicado problema combinatorio.

El método explicado es un *algoritmo*, siempre nos da la solución óptima. Desgraciadamente no es posible su aplicación

* En general, nos contentaremos con una cota inferior, no necesariamente la mejor posible, y es ése el sentido con que se debe interpretar "mínimo" aquí.

directa a los problemas que nos interesan. El principal problema con que nos encontramos es la dificultad en calcular los mínimos necesarios; por otra parte se debe modificar levemente —en los juegos, por ejemplo, nos interesa maximizar y minimizar alternativamente—, y en cualquier caso, las evaluaciones finales no se harán casi nunca con “información completa”, sino que serán producto de valoraciones aproximadas. Sin embargo, usando dichas aproximaciones, y mediante el uso de técnicas heurísticas suplementarias para reducir la exploración, el método da origen a varios procedimientos útiles, investigados en el M.I.T., en cuanto a su aplicación a los estudios en inteligencias artificiales (McCarthy y otros). El método “alfa-beta” consiste en no continuar explorando aquellas ramas que en un cierto momento ya se muestran inútiles (que no pasan del valor “alfa”) o que son demasiado malas para nuestro oponente (valor *mayor* que “beta”, recordar que su objetivo es *minimizar*). Nos basaremos en la exposición de Samuel⁷.

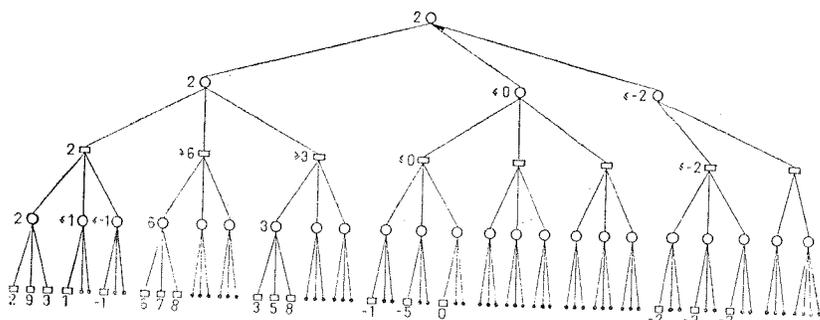


Figura 5 (Samuel).

Antes de explicar el método, recordemos que al razonar nuestra mejor movida con ayuda del minimax siempre suponemos que nuestro contrincante intentará minimizar nuestro beneficio, mientras que nuestro fin es aumentarlo, y que suponemos que podemos evaluar numéricamente este “puntaje” para poder así comparar las alternativas. El árbol de la Fig. 5 da un ejemplo de un caso en que el método resulta altamente eficaz: solo se calculan algunos de los valores (los escritos), mientras que no hay necesidad de investigar la mayoría de las bifurcaciones (líneas de puntos). El lector podrá verificar que se ob-

⁷ Samuel, A. L., *Some studies on machine learning using the game of checkers*, IBM Journal of Research and Development, noviembre de 1967.

tiene la solución correcta cualesquiera sean los valores omitidos, asignando puntajes arbitrarios y efectuando el "minimax" correspondiente. Con este ejemplo estudiaremos el método.

Para seguir la dinámica del proceso con claridad observemos la Fig. 6, obtenida de la anterior eliminando las ramas no evaluadas. Mediante el cómodo procedimiento de invertir alternativamente los signos de las evaluaciones, el proceso se

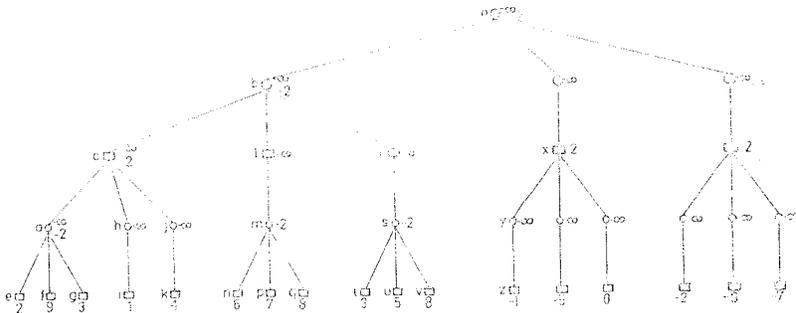


Figura 6 (Samuel).

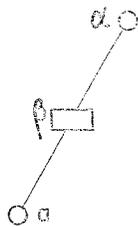


Figura 7.

puede considerar como enteramente de maximización, tanto para nosotros como para nuestro oponente (ya que tomar el mínimo de un conjunto de números equivale a cambiarles el signo y maximizar), y ahora no hay necesidad de distinguir los valores "alfa" de los "beta", ya que juegan papeles simétricos. Los valores "alfa" se calculan a medida que se genera la posición correspondiente, y representan el valor que deberá ser excedido para que el jugador a quien corresponde el turno de mover considere interesante investigar esa rama.

Al comenzar la evaluación, se colocan los "alfas" aún desconocidos en un valor negativo suficientemente pequeño, representado en la figura por " $-\infty$ ". El orden en que se realiza la

búsqueda está indicado alfabéticamente. Al llegar al último nivel (suponiendo siempre una cantidad fija de jugadas anticipadas), evaluamos la posición mediante la función respectiva, obteniendo un 2 para e . El valor de la posición dos niveles más arriba (c , aquí) es usado como "alfa" a los fines de decidir la continuación o no; si el valor actual no lo supera, no tendría sentido continuar por este camino ya que el oponente podría, en d , jugar e y de este modo disminuir el puntaje que ya teníamos asegurado con llegar a c . Del mismo modo, si la inversa del valor actual no supera al de d , estamos suponiendo que nuestro oponente trata de maximizar dicho valor, y en ese caso no jugará e , que le disminuiría la ganancia. Las situaciones del primer tipo, que determinan la no continuación de la exploración, se denominan "puntos de corte". En este caso, sin embargo, el valor de e (2) supera al de c ($-\infty$) y su inversa (-2) supera también a d ($-\infty$); el jugador en d elegirá entonces —por lo menos ese es nuestro razonamiento— la movida que lleve a e , ya que así ganará más (-2), por lo que reemplazamos el valor de d por éste. Considerando ahora f , si bien nos permitiría aumentar nuestra ganancia, no sería aceptado en d , ya que (-9) es menor que (-2). Lo mismo ocurre con g , con lo que ya obtuvimos el valor final para d (-2), el cuál comparamos —repetiendo el proceso anterior a otro nivel— con b ($-\infty$), y siendo mayor se compara su inversa con c ($-\infty$), al cual también supera y por lo tanto lo reemplaza.

Se genera ahora la posición h , a la cual asignamos el valor mínimo que el jugador correspondiente aceptaría (su ganancia "asegurada" es, hasta el momento, el " $-\infty$ " de b). Se evalúa ahora i , obteniéndose un valor "1", lo que indica que si a partir de c jugamos h , el oponente podrá jugar i y así reducir nuestra ganancia, la cual ya habíamos calculado en 2. Existiendo esta posibilidad, no tiene más sentido continuar explorando a partir de h , y repetimos el proceso con j , donde acontece algo similar, y así sucesivamente hasta eliminar todas las posibilidades. Resumiendo, se tiene la situación de la Fig. 8: si se verifica $a > \alpha$ y $\beta > a$ (no trabajamos ahora con los signos invertidos), se reemplaza " β " por el valor de " a ". Una heurística adicional que nos proporciona este diagrama es la siguiente: si dos valores consecutivos de " α " y " β " difieren poco, podemos detener la búsqueda en esa dirección, ya que cualquier valor de " a " que podría hacerlos variar, deberá también diferir poco de éstos (ya que sería $\alpha > a > \beta$).

El ejemplo dado era extremo en cuanto al número mínimo de continuaciones que se debió explorar, y si bien en los casos prácticos las situaciones no son tan adecuadas, se prestan igualmente a un gran ahorro con el uso de este método (reducción

del factor de bifurcación de un 30/40 % en el juego de damas), y puede ser mejorado mediante la aplicación de diversas técnicas heurísticas para decidir el *orden* en que se realizan las evaluaciones, de modo de permitir encontrar los puntos de corte más rápidamente. Samuel calcula que con una exploración ideal el factor "aparente" de bifurcación se puede reducir a su raíz cuadrada; esto significa, para 8 alternativas por movida, la posibilidad del análisis de 5 jugadas de antelación (alrededor de 66.000 casos) en unos 5 minutos, contrariamente a los miles de horas que llevaría analizar los mil millones de posibilidades en caso de no utilizar el método.

Notemos que el método considerado consiste en una exploración "en profundidad" más bien que "a lo ancho", ya que investigamos exhaustivamente en una dirección y recién después atendemos a otros niveles superiores (ver el orden de las letras en la Fig. 6), mientras el clásico minimax se puede aplicar más "paralelamente". Una técnica heurística eficiente es el realizar un "preanálisis" de poca profundidad pero total, que permite así decidir —sobre la base de los valores calculados con una evaluación realizada a los pocos niveles— cuales son las ramas más "prometedoras", y aplicar a éstas la exploración "profunda" en primer término. Esta variante tiene sus ventajas —ordenación de las alternativas que puede llevar a una gran reducción del esfuerzo— y desventajas —algunas estrategias eventualmente buenas pueden comenzar con jugadas aparentemente malas, tales como perder una pieza, etcétera—, y es usada por Samuel en su programa modificado (1967). En este programa, la exploración se finaliza cuando así lo indica un cierto "contador de bifurcaciones" que decide prolongar a mayor profundidad las ramas "preferidas" (las que el análisis previo definió como tales), y continuando aún más cuando se encuentran intercambios de piezas, con el fin de aplicar el concepto de "posición estática".

Siendo los métodos de exploración tan fundamentales al tema de las inteligencias artificiales, conviene analizar algunas variantes más, en particular las usadas por Samuel. Una interesante combinación de análisis y alfa-beta está dada por el siguiente procedimiento: primeramente se consideran todas las movidas y sus posibles respuestas, ordenándose estas ramas de acuerdo con su puntaje, las mayores primero. Luego se continúa por tres niveles más, y se obtienen los nuevos puntajes; la rama con el máximo es elegida para continuar la exploración por dos niveles, y si el último valor así obtenido no supera al almacenado de la rama que seguía a la actual en orden de

prioridad, se elige esta última para continuar el análisis *. Este proceso se repite hasta una cierta profundidad máxima, determinada por diversos factores (ver párrafo anterior), y se realiza luego el minimax basado en el algoritmo alfa-beta. Otra heurística que puede resultar efectiva es la poda "hacia adelante", consistente simplemente en descartar las posiciones que parezcan poco fructíferas, pero su uso excesivo puede llevar a pasar por alto muchas jugadas ganadoras a largo plazo **.

La posibilidad de otros métodos de exploración en árboles es muy amplia, y posteriormente veremos un programa ("Múltiple") que rivaliza en eficiencia con el método alfa-beta, al menos en el juego de "kalah", mucho más simple que los aquí discutidos; en particular veremos cómo un juego puede ser considerado (para su solución) como un *teorema* de lógica formal, dada una cierta formalización para éste, siéndole entonces aplicable gran parte de las técnicas heurísticas usadas en estas áreas.

Volviendo al ajedrez, digamos que en noviembre de 1966 Richard Greenblatt, del proyecto MAC (M.I.T.), puso a jugar a su programa MACHACK VI; siendo éste admitido como miembro honorario de la *Federación Norteamericana de Ajedrez* y, con un nivel de juego similar al del aficionado promedio, fue presentado en la conferencia de la AFIPS de noviembre de 1967, habiendo participado hasta mediados de 1969 en cuatro torneos, ganando alrededor del 86 % de sus partidos en dichos encuentros, y con una experiencia total de unos 2.000 partidos.

El programa⁸ investiga las jugadas juzgadas más interesantes por el generador de movidas plausibles, utilizando un "evaluador de posiciones" que verifica si se cumple o no una condición de inestabilidad (una pieza amenazada y jaque, o

* Este proceso, la "ordenación dinámica" de alternativas, es un refinamiento del "análisis superficial" o "preanálisis", en cuanto permite modificar estas estimaciones previas sobre la base de la información que se va obteniendo sobre la bondad de la rama, causando, si es necesario, un reordenamiento de las prioridades, lo que se traduce en la optimización en la utilización del "alfa-beta", ya que es así mayor la posibilidad de obtener un punto de corte. Un caso particular es el método "A-B" de Slagle y Dixon (1969), en el cual se posee un parámetro que determina la continuación de la exploración de una rama o el reordenamiento. El ordenamiento dinámico es también utilizado por Slagle y Lee en su método "gamma" (ver 3.2.3.).

** Se pueden considerar sólo las "n" mejores posiciones siguientes en cada caso, de acuerdo con diversos criterios, siendo este factor n constante o variable, de acuerdo con la profundidad (atemperación del factor con el fin de evitar una proliferación excesiva de ramas).

⁸ Greenblatt, R. D., et alii, *The Greenblatt chess program*, A. I. Memo 174, abril 1969.

dos piezas en peligro simultáneamente), en cuyo caso se repite el proceso con la posición alcanzada, y si no es así, se investigan los posibles intercambios de piezas, hasta una posición estática; el evaluador estático combina balance, proporción de piezas, posición de peones, seguridad del rey y control del centro, cada una de éstas dependientes a su vez de varias características.

Si bien hace uso de los conceptos fundamentales enunciados, el programa fue posteriormente evolucionando: las evaluaciones, por ejemplo, no se hacen de acuerdo con los métodos exclusivamente numéricos anteriores, sino mediante un razonamiento de tipo más bien "simbólico" consistente en la identificación de las amenazas presentes de modo individual y la generación de respuestas adecuadas ante cada una; el generador de "movidas plausibles", por otra parte, actúa de acuerdo con la situación del juego, si la situación es favorable pueden convenir los cambios de piezas y las simplificaciones, mientras que en caso contrario el juego es de tipo más bien agresivo. En 1971 el programa, considerado ya el mejor "jugador artificial" del mundo, continúa mejorando, pero con un juego aún inferior al de muchos seres humanos.

El razonamiento de tipo simbólico antes que numérico representa un esfuerzo muy actual en el campo de los juegos (en realidad, en todo I.A., ver 3.3.), y se obtienen interesantes resultados mediante la formalización y el uso de lenguajes adecuados. Un ejemplo interesante es el programa de King⁹ (1971) para resolución de juegos "posicionales", o sea aquellos que consisten en la "ocupación" alternativa de casilleros por ambos jugadores, no pudiendo éstos desocuparse y finalizando el juego cuando alguno de los jugadores gana habiendo ocupado alguno de los "conjuntos distinguidos" (estados ganadores), o con un empate si no quedan casillas por ocupar y ninguno ganó. La estrategia básica consiste en identificar (y tratar de alcanzar) aquellos estados que son "obligados", en el sentido de que una vez en ellos el oponente sólo podrá efectuar movidas con las que perderá; para esto se cuenta con descripciones de estos estados, y se comparan los encontrados para verificar si son tales. Estas descripciones se obtienen fácilmente gracias al formalismo utilizado, basado en el uso de grafos, y con la ayuda de un aprendizaje que permite "recordar" los más frecuentes.

En varias ocasiones nos referimos al aprendizaje en los juegos, pero hemos postergado hasta el final su discusión para que se tuviese clara la idea de las estrategias heurísticas de

⁹ King, Paul F., *A computer program for positional games*, Jennings Computing Centre, 1971.

juego en sí. Ahora consideraremos algunos esquemas de aprendizaje; el programa de Samuel es, como ya lo señalamos, uno de los mejores ejemplos —en cuanto a complejidad y eficiencia— de programas que “aprenden” mediante el ajuste de parámetros y otras evaluaciones de tipo más bien estadístico. Si bien en la actualidad se tiende a un aprendizaje de tipo más superior, éste es básico para comprender la importancia que puede tener la automodificación debida a la interacción con el medio externo para los programas “inteligentes”.

3.1.2. Aprendizaje elemental de los juegos.

Ya en la clásica teoría de juegos (de dos personas y suma cero) se tenían algunos métodos que se podrían denominar de “aprendizaje”, llamados también “métodos iterativos”¹, que, aunque de convergencia lenta, tienen un interés teórico en cuanto permiten decidir una estrategia óptima sobre la base de la experiencia. Los juegos que nos interesan, sin embargo, son juegos secuenciales para los cuales, aun cuando es teóricamente posible la formulación de todas sus estrategias y la “matriz de pagos”, en la práctica esto no resulta aplicable, debido al problema de complejidad que se plantea. Se deben encontrar entonces métodos particulares para resolver las dificultades que se suscitan.

El tipo más elemental de aprendizaje consiste en el refuerzo de aquellos aspectos positivos, y la supresión de los negativos que se presenten en el comportamiento. Si bien es un método en apariencia extremadamente simple, en la práctica resulta complicada su implementación debido a la presencia en casi todos los casos de una cantidad relativamente grande de factores que influyen sobre el comportamiento de un modo difícil de determinar con precisión, ya sea debido a lo leve de sus efectos o al hecho de no ser independientes, traduciéndose esto en uno de los problemas básicos de la teoría de aprendizaje: el problema de la “asignación de responsabilidades” individuales. Si bien para algunos juegos simples (ta-te-ti, por ejemplo), puede ser útil considerar como causante de una victoria a las últimas movidas, esto desaparece totalmente al considerar juegos como las damas o el ajedrez, donde una buena apertura puede decidir la suerte de los contrincantes (en realidad, inclusive

¹ Ver. por ejemplo, Gale, D., *The theory of linear economic models*, McGraw Hill, 1960, Cap. 7.

el ta-te-ti, el control inicial del centro puede ser determinante de una victoria).

Un ejemplo sumamente simple de dicho método está dado por un programa implementado en la Universidad de Buenos Aires², el cual, dado un juego cualquiera, mediante su "tabla de transiciones" (una enumeración de las posiciones de tablero o "estados" posibles y una indicación de cómo generar las movidas válidas a partir de éstos) jugaba al principio de modo completamente al azar, pero al finalizar cada partido ajustaba las distribuciones de probabilidades de modo de dar mayor importancia a las jugadas que lo llevaron a la última (si ganó), o de restar valor a éstos (si perdió el partido). Se tuvo en cuenta el caso en que el adversario pudiera jugar mal los primeros partidos (con el fin, por ejemplo, de obligar al programa a producir una estrategia que al repetirse posteriormente, lo llevara a perder), y el crédito no se asignaba solamente a las últimas movidas sino que diversos criterios permitían un progresivo aprendizaje hasta llegar a las movidas iniciales. Pero aunque el programa dio resultados satisfactorios —puesto a jugar consigo mismo mediante la simulación alternativa de los dos oponentes obtenía la estrategia óptima en pocos minutos— no es posible esperar de tales intentos, por supuesto, la solución a los problemas realmente importantes, en los cuales no es posible almacenar las distribuciones de probabilidades para cada posición y, aunque fuese factible, resultaría extremadamente lento. Similares métodos "probabilísticos" tienen interés, sin embargo, desde otros puntos de vista (ver en general sección 3.4).

También en la línea del "aprendizaje por memorización" descrito se encuentra uno de los métodos aplicados por Samuel en la versión original de su programa que juega a las damas. Se trata aquí de almacenar los valores calculados para posiciones de tablero ya encontradas (no de las transiciones en sí, como el anterior, lo que permite varias ventajas, incluyendo la reducción del espacio necesario). La utilización de estos valores se traduce en una mayor profundidad efectiva de exploración, como veremos con un ejemplo. Supongamos (Fig. 8) que en un partido anterior se tenían varias alternativas desde una posición "A", las cuales se investigaron hasta una profundidad fija de 3 movidas, dando la rama que conduce a "B" el valor óptimo (digamos "2"). Este valor "2" se almacena en un sitio especial asociado a la posición A.

² Memorandum interno de GEDECC, Facultad de Ingeniería, febrero de 1971.

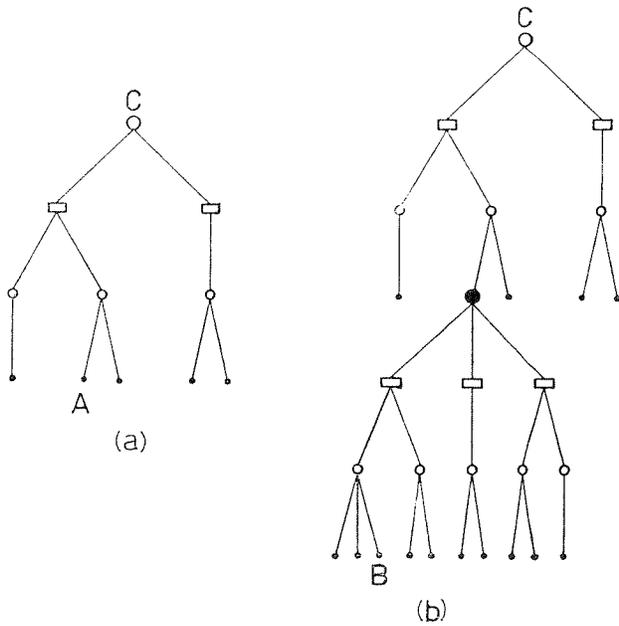
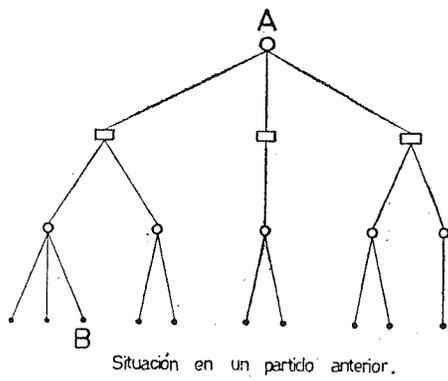


Figura 8. Situación en un partido anterior, y el "aprendizaje por memorización".

En el partido siguiente, figuremos que se efectúa un análisis desde "C" (Fig. 8a), también de profundidad fija en 3, y que "A" se encuentra entre los nodos del último nivel. Cuando se deba calcular ahora el valor de "A", no se hará mediante el evaluador estático sino refiriéndose al valor almacenado (en general, usando la función de evaluación para una posición

cualquiera sólo en el caso de no tener ésta un valor anteriormente "memorizado"), lo que equivale a una profundidad para algunas ramas de 6 movidas, es decir, el resultado es el mismo que si se hubiese realizado la exploración diagramada en la Fig. 8b, incluyendo todas las evaluaciones respectivas (en ambos casos se efectúa luego el "minimax" habitual para evaluar C).

La efectividad de este método se puede aumentar de varios modos. Por una parte, es útil almacenar para las posiciones calculadas no solo su valor sino también la cantidad de movidas o niveles con que éste fue calculado (en el ejemplo anterior, asignaríamos a "A" un valor "2" y una cantidad de niveles "3"); esto permite ante varias alternativas que ofrecen iguales resultados, elegir la que permita obtenerlos en el lapso menor e inclusive mayor, esto último en la situación en que notamos que estamos perdidos cualquiera sea nuestra respuesta, pero intentamos "hacer tiempo" en la esperanza de que nuestro adversario no note su posibilidad de ganar. En segundo término, puede no tener sentido memorizar jugadas que no son usadas frecuentemente, ya que en tales casos la ventaja en cuanto a profundidad se verá anulada por el gasto en memoria y la velocidad; por otra parte, las jugadas que son a menudo efectivas deben ser "recordadas" ya que su uso es habitual. Este principio, típico en los seres humanos, se aplica en el programa mediante un procedimiento consistente en la asignación de un "índice de edad" a cada valor calculado, el cual se disminuye en cada ocasión de ser usado, y que aumenta automáticamente luego de cada cantidad fija de movidas. La posición para la cual su "edad" haya pasado de un valor máximo es "olvidada", mientras que aquella que haya sido usada repetidamente permanece siempre "joven".

Este método de aprendizaje es especialmente útil para ciertas aperturas y situaciones de fin de partido que aparecen habitualmente, pero no resulta tan efectiva para el juego medio, donde se presenta una gran cantidad de posibilidades en cuanto a posiciones de las fichas. En estos casos, sin embargo, hay algunas propiedades en común de las situaciones de las cuales el buen jugador sabe extraer analogías y plantearse así buenas estrategias globales. El lugar apropiado para este tipo de conocimientos es la misma función de evaluación estática, que representa qué factores (y con qué peso respectivo) deben ser consideradas como útiles o perjudiciales, y un aprendizaje de un nivel superior al anterior (y quizá aún más parecido al de los seres humanos) consiste en modificar esta función de acuerdo a la experiencia.

En la modificación del "polinomio de evaluación" consiste el "aprendizaje por generalización", que actúa alterando los coeficientes al final de cada partido de acuerdo a un complicado sistema de correlaciones entre los distintos valores obtenidos para el polinomio y los coeficientes respectivos durante el juego; con una política consistente (aquí el problema de "asignación de responsabilidades" ya mencionado) en tener en mayor consideración los factores que gravitaron positivamente y restar importancia a los negativos (si se llegó a buenos resultados mediante su utilización) y en el proceso inverso en caso de no haber dado los resultados esperados. Mediante otro proceso, se hace "elegir" al programa los términos que desee usar (por ejemplo, movilidad de las piezas, avance de éstas, control del centro, etc.) de una lista que le fue suministrada (elige unos 16 de un total de 32 factores posibles).

Este método es útil para el juego medio, por lo que en conjunción con el "aprendizaje de memoria" para las aperturas y los finales, puede dar lugar a una gran capacidad de entrenamiento. Esto tiene lugar jugando la máquina "contra sí misma", contra seres humanos o inclusive mediante estudio de partidos publicados, tal cual lo hacen usualmente los jugadores de damas o ajedrez. En un partido jugado en 1965 empató con el campeón mundial W. F. Hellman, pero perdió cuatro partidos jugados con el mismo por correspondencia.

Una limitación fundamental de estos métodos de aprendizaje, y aún del que luego se agregó, consiste en la no creación del programa de los factores a considerar sino que éstos son solamente elegidos y ponderados a partir de una lista ya suministrada. Un modo de conseguir esta mayor "generalización" sólo podría basarse en una especie de "entendimiento" de la situación por parte de la máquina; enfocado desde otro punto de vista, lo que se necesitan son algo así como "heurísticas que permitan aprender otras heurísticas"*. La situación no es tan mala, sin embargo; en el ajedrez por ejemplo, gran parte de los principios en que se basan los campeones se deben al estudio realizado por sus predecesores y los factores señalados ya por éstos.

El método de "agrupación de factores" fue usado primeramente por el ya nombrado Griffit³, y representa un intento de adoptar un medio de evaluación más flexible que el del polinomio lineal pero aún suficientemente rápido y práctico como

* En ese sentido es interesante el programa de Waterman (3.4.3.).

³ Se puede encontrar un resumen en Griffit, A. K., *A new machine learning technique applied to the game of checkers*, A. I. Memo 94, M.I.T. 9-3-66.

para ser implementado en una computadora. Consiste, en síntesis, en el agrupar los términos que pueden tener importancia para un cierto aspecto del juego (en ajedrez, por ejemplo, para "avance de las piezas propias" se podrían agrupar los factores "posición de peones", "posición de los alfiles", etc.), y definir una tabla mediante la cual a cada configuración posible de estos factores se le asigna un valor, de acuerdo a su utilidad. La evaluación final se haría, entonces, de acuerdo a los resultados indicados por cada grupo. Lo fundamental en este enfoque es que permite la interacción de factores asociados de un modo difícil de conseguir en el caso en que todos se ponderaban simultáneamente, sin atender a sus relaciones mutuas (en rigor de verdad, esto era posible hasta cierto punto en el programa original de Samuel mediante ciertos "términos binarios" del polinomio, pero no es claro que se pudiesen obtener resultados importantes de este modo); así es posible comparar las ventajas y desventajas que nos presenta la solución a un problema mediante las comparaciones directas entre dichos factores, en lugar de su estudio aislado y posterior balance conjunto.

En la nueva versión de su programa de damas, Samuel (1967) incorpora el citado método a sus procedimientos de evaluación y aprendizaje. Para la evaluación a partir de los resultados de cada grupo, cuenta con una "jerarquía de niveles", de modo que si dos subconjuntos de factores están a su vez relacionados, se encuentran agrupados en un grupo que a su vez los contiene. El resultado final se obtiene mediante una última tabla, cuyos parámetros son los resultados obtenidos en las jerarquías inferiores. El aprendizaje se lleva a cabo mediante un proceso consistente en el análisis de jugadas publicadas ("book learning"); el programa simula jugar ambas partes, y compara sus movidas con las indicadas en el libro. Si coinciden, asigna a cada configuración de factores que estuvieron presentes en cada uno de los grupos un "puntaje" positivo, mientras que si las jugadas difieren, les disminuye dicho peso (en realidad el proceso es mucho más complicado, pero lo antedicho es suficiente para captar su esencia). Los resultados obtenidos con este método superan ampliamente al anterior, sucediendo algo similar a lo que acontece en casi toda otra ciencia, cuando se dejan de lado las simples y prácticas, pero distorsionantes, aproximaciones lineales.

Radicalmente distinto es el enfoque del Dr. Ranan Banerji (Ohio, U.S.A.) quien ⁴ reconoce lo específico de estos programas,

⁴ Banerji, Ranan B., *Machine learning of games*, Computers and Automation, noviembre-diciembre de 1970.

en el sentido de su eficiencia restringida a un juego particular; el ideal en cuanto a programas que aprendan es por supuesto el consistente en aquellos a los cuales se les deba dar sólo las reglas del juego, y éste encuentra los "trucos" necesarios por sí mismo. Propone para ello la obtención de modelos generales eficientes para los juegos y para las estrategias necesarias; a su vez esto plantea la necesidad de un adecuado lenguaje de comunicación entre el hombre y el programa. La generalidad de dicho lenguaje hace por el momento difícil su desarrollo, pero ya se han llevado a cabo experiencias con otros más restringidos pero aún cubriendo una amplia clase de juegos (los "posicionales" del programa mencionado de P. King).

No acaba aquí todo lo referente a juegos, ya que nos encontraremos a menudo con temas estrechamente relacionados, cuyos resultados pueden ser fácilmente aplicados en dicho dominio.

3.1.3. Métodos heurísticos para temas "formales".

No hay duda de que las matemáticas, el dominio formal por excelencia, son consideradas como una actividad que requiere un esfuerzo intelectual superior, por lo que la invasión de este terreno por parte de los "autómatas" reviste una importancia fundamental. Por una parte, se dispone de dominios axiomáticos, donde todo el esfuerzo del investigador puede dedicarse al estudio de los métodos, antes que a la preparación de datos y otros detalles circunstanciales; por otro lado, dada la creciente aplicación de las matemáticas en todas las ciencias, la obtención de nuevos resultados, quizá fuera del alcance de la limitada mente humana, puede tener una importancia insospechada. Y si se identifican a las matemáticas con la inferencia deductiva, no es necesario seguir remarcando su importancia; los métodos que aquí se obtengan serán utilizados en cualquier tipo de razonamiento.

Dentro de las "áreas formales" a considerar, se encuentra en primer lugar la lógica simbólica (proposicional y de predicados), y luego distintas áreas matemáticas (geometría, análisis, álgebra, etcétera). Como en el punto anterior, además de los ejemplos de esta sección, nos encontraremos con temas relacionados en otros lugares (deducciones lógicas en los sistemas de respuesta, problemas de álgebra en estudio de lenguajes naturales, y fundamentalmente la utilización de la lógica de predicados en variados dominios). Además del diseño de sistemas artificiales que realicen creaciones propias, el estudio de estas áreas está relacionado con la posibilidad de un comportamiento

interactivo entre seres humanos y máquinas. Un buen ejemplo de esto último está dado por los programas destinados a servir de ayuda a los matemáticos en sus trabajos, como los existentes en el M.I.T., donde un algebrista puede solicitar a un programa, desde una terminal de tiempo compartido, que éste le encuentre los isomorfismos existentes entre dos estructuras dadas, los subgrupos de un cierto grupo u otros datos similares, y donde el ingeniero puede entregar una complicada fórmula simbólica para que le sea simplificada o a fin de realizar una integración.

Respecto de la obtención de resultados aplicables a la lógica antes que el estudio de los métodos heurísticos, fue pionero en este tipo de trabajos Hao-Wang¹, quien obtuvo importantes "reglas de decisión" * para la resolución de problemas en la lógica proposicional y los generalizó a la de predicados de primer orden. Si bien para la lógica proposicional hay métodos directos ("tablas de verdad"), el interés radica en la generalización a áreas no decidibles (donde no existen propiamente algoritmos). Implementando su método en una IBM-704, Wang demostró los 350 teoremas de los primeros nueve capítulos de "Principia Matemática" (la obra básica de lógica matemática publicada a principio de siglo por Whitehead y Rusell) en ocho minutos y medio.

El trabajo de Wang se encuadra en el tipo de programas basados en el "método de expansión de Herbrand", que consiste en una búsqueda de soluciones en un universo dado, previa reducción del problema a una forma canónica. Los trabajos de Davis y Putnam y de J. A. Robinson también son significativos en el área, y en un excelente resumen T. P. Hart² explicó sus características sobresalientes, proponiendo además un programa que utilizaría dicho método en conjunción con métodos heurísticos que acortasen el proceso.

¹ Wang, Hao, *Towards mechanical mathematics*, IBM Journal of Research and Development, 4, 1, enero de 1960.

* Una regla de decisión es un algoritmo que determina, dada una cierta sentencia, si ésta es un teorema o no; un algoritmo de demostración (de teoremas), en cambio, dará una demostración en el caso positivo, pero en caso de ser la sentencia falsa no está determinado su comportamiento (así, el procedimiento enumerativo, al que nos referiremos luego, seguirá indefinidamente generando pruebas y comparando las conclusiones con el teorema en cuestión). Es claro que una regla de decisión puede ser en particular un algoritmo de demostración, pero no la recíproca, ya que la primera da siempre un resultado. La mayoría de las áreas interesantes no son decidibles (no existen reglas de decisión posibles), por lo que la investigación se concentra en la búsqueda de algoritmos de demostración eficientes.

² Hart, Timothy, *A proposal for a Geometry Theorem proving program*, A. I. Project Memo 56, M.I.T., 11-9-63.

En 1965 Robinson presentó su "principio de resolución" *, que significó un notable avance en el área. Los programas que se han llevado a cabo son tantos que es imposible siquiera su enumeración aquí. Respecto de programas que "interactúan" con el usuario, son importantes entre otros los proyectos del M.I.T. ("asistente matemático" de W. A. Martin en 1965, "experimentos de álgebra" de W. D. Maurer en el mismo año, "Mathlab", "laboratorio matemático", de Martin y Moses en 1968, etcétera), y de otras universidades (G. P. Huet en Ohio, 1971, etcétera).

En el presente libro sólo nos referiremos brevemente, por falta de espacio, a algunos casos que nos permitirán obtener resultados generales, especialmente en los métodos heurísticos de planificación ("Logic Theory Machine"), la importancia de la semántica (Gelernter) y la exploración en árboles.

En un sistema de lógica formal, un "problema" a resolver es simplemente la "demostración" de un teorema, que es una sentencia en cierto simbolismo. Esta demostración consiste en una secuencia de sentencias, cada una de las cuales es un teorema ya demostrado, un axioma, o una inferencia inmediata de la anterior, obtenida mediante el uso de una de las "reglas de inferencia" permitidas; la primera de esta sucesión de sentencias debe ser también un axioma o teorema previo, y la última debe ser la expresión del teorema a demostrar.

Existiendo siempre una prueba de si una secuencia es una demostración válida (se deben verificar simplemente las condiciones del párrafo anterior) el problema es un principio trivial, ya que se podrían generar todas las secuencias posibles e ir sometiendo a dicha prueba a cada una. Sin embargo, este método es tan inútil como el algoritmo de considerar todas las posibilidades hasta el fin del partido en el ajedrez; el problema combinatorio hace imposible su uso.

Una mejora notable a este esquema se obtiene al realizar un proceso similar al que lleva a cabo un matemático mentalmente al intentar solucionar un problema: el proceso invertido de partir del resultado deseado e intentar llegar a encontrar proposiciones que lo impliquen, haciendo el proceso recursivo hasta que sea implicado por un axioma o teorema ya válido.

* La *resolución* es una regla de inferencia general que ha sido adoptada en casi todos los programas contemporáneos. Varias modificaciones posteriores han tendido a optimizar su uso, al proveer estrategias para la búsqueda de la solución, y al considerar la "igualdad" como una relación especial ("resolución generalizada"). El extendernos en la explicación de los principios del método involucraría apartarnos del carácter de la exposición, más aún considerando su esencia "algorítmica" antes que "heurística".

Pero este método tiene también sus desventajas; si deseamos probar, por ejemplo, que "todo perro tiene cuatro patas", una proposición que lo implica es "todo ser viviente tiene cuatro patas y todo perro es un ser viviente", pero no siendo esta última proposición un axioma ni un teorema, debemos a su vez demostrarla, lo cual difícilmente podamos conseguir (a menos que tengamos en tan poca estima a los demás humanos). Es decir, mientras que la secuencia termina donde queríamos, la mayoría de las veces lo hará partiendo de suposiciones falsas, por lo que no estaremos probando nada. El método es, sin embargo, más útil que el "directo", ya que ahora se "parte" de una sola sentencia, pudiendo "llegar" (siempre hacia atrás) hasta cualquiera de varios resultados ya establecidos; en el método directo debemos además establecer de cuál partir.

El programa LT ("Logic Theorist")³ fue quizá el primer intento serio en el campo de las inteligencias artificiales en cuanto a grandes programas heurísticos; los trabajos posteriores de Newell y su grupo se han encuadrado en el mismo marco de selección de metas y subproblemas intermedios. Este programa, para la lógica proposicional, actúa, como todos, "hacia atrás" a partir del enunciado del teorema a demostrar, intentando llegar a axiomas o teoremas anteriores que lo implique. Para la generación de estos teoremas anteriores (que deben a su vez ser demostrados si aún no lo han sido, lo que significa postularlos como metas intermedias a resolver), se cuenta con criterios de "comparación" de expresiones, que se basan en una confrontación entre el resultado a demostrar y los axiomas que podrían ser útiles, deduciendo las transformaciones adecuadas, las cuales corresponden a las distintas reglas de inferencia (métodos de sustitución, encadenamiento, etcétera). Hay constantemente una "lista de subproblemas" que se atacan en orden, agregándose nuevos a medida que se van generando en el curso de las demostraciones, produciéndose así un "árbol de subproblemas", situación similar a la encontrada en los juegos. Los resultados no fueron muy importantes en sí, pero este programa se convirtió en el punto de partida de los esfuerzos posteriores.

En esta línea se encuentra el programa "SAINT" ("Symbolic Automatic Integrator") escrito en LISP en el M.I.T. por Slagle⁴; que permitía la integración simbólica (no numérica) de funciones, combinaciones de trigonométricas, potencias y logaritmos.

³ Newell, A., Shaw, J. C. y Simon, H. A., *Empirical explorations of the logic theory machine : A case study in heuristics*, Proc. 1957 Western Joint Comp. Conference, Vol. "11".

⁴ Slagle, J., *A heuristic program that solves symbolic integration problems in freshman calculus*, J.A.C.M. 10, 1963.

Contaba con rutinas de transformación del integrando por sustitución de variables, sustitución de subexpresiones cuya derivada divida al integrando, etcétera. Se forman dos tipos de árboles, los "conjuntivos", que indican que todos los subproblemas (ramas) deben ser resueltos para resolver el original (el nodo inicial), y los "disyuntivos", que indican que alcanzará resolver uno cualquiera de los problemas; la "poda" de los árboles se realiza a medida que se resuelven las etapas intermedias (por ejemplo, si se resuelve una de las alternativas en una disjunción, ya no es necesario atacar las demás, y se pueden eliminar). Resolvió gran cantidad de problemas de cálculo a nivel universitario, tardando desde 2 segundos hasta 18 minutos, según el caso. La organización del programa era similar al de "LT", jugando un papel importante el reconocimiento de la configuración de la expresión a integrar ("Pattern recognition", 3.2.), como medio para decidir las transformaciones a aplicar. Podía realizar además algunas integrales definidas e integrales dobles iteradas.

Más recientes en la misma área son el programa "INTEGRATE" de M. Manove (1965) para integración simbólica de funciones racionales a una variable, SIN ("A Symbolic Integrator") de Joel Moses (1966/67) que generaliza al de Slagle permitiendo resolver una amplia gama de problemas, llegando al "MATHLAB" (1969) como síntesis de todos éstos⁵, que permitió incluso descubrir un error en una tabla de integrales publicada en 1958. Los programas actuales se escriben mediante la utilización de lenguajes apropiados, que poseen ya, en forma de macro-instrucciones, rutinas para manejar los procesos más usuales (sustituciones, simplificaciones, etcétera), lo cual resulta en una mayor facilidad de diseño, y permite concentrarse en mayor medida en los problemas más importantes.

También basado en los principios de niveles de problemas se ubica el programa de demostración de teoremas de geometría elemental de Gelernter⁶, pero su importancia fundamental radica quizá en que por vez primera se mostró explícitamente la importancia de la semántica en la resolución de los problemas, a lo cual nos referiremos en detalle más adelante (3.3.). En la discusión previa, ya analizamos el inconveniente principal de las demostraciones "a la inversa", consistente en la ignorancia respecto de la validez de los sucesivos pasos. Cuando demos-

⁵ Aparecidos en general en distintas publicaciones del M.I.T. durante los años mencionados.

⁶ Gelernter, H., *Realization of a geometry theorem-proving machine*, Proc. of the International Conference on Information Processing, Paris, 1959.

tramos un teorema de geometría, sin embargo, tratamos de probar que dos ángulos son iguales o dos lados paralelos sólo si el estudio de un diagrama nos lo indica. Puesto de otro modo, el tener en consideración el *significado* o una *interpretación* intuitiva de las fórmulas en cuestión, nos permite resolver el problema; aun cuando no lo dibujamos en un papel, es bastante común el uso mental de un diagrama. Si el dibujo se efectúa de acuerdo con los datos, y proporcionado de tal modo de evitar igualdades accidentales, no incluidas en los datos, casi seguramente todas las relaciones observadas podrán ser objeto de una demostración formal.

El programa de Gelernter se basa en una estructura dividida en los procesos heurísticos, "sintácticos" y "gráficos". El programa heurístico se encarga de parte del control del proceso, comparando las sentencias propuestas como válidas por las rutinas sintácticas con los datos entregados por la parte encargada de los diagramas; se utiliza, además para reconocer ciertas "simetrías sintácticas" como las que llevan a un matemático a probar sólo una parte de un teorema, aclarando "el otro caso es similar". La parte sintáctica es equivalente a la de LT y otros programas, aunque se divide su actuación con la heurística. El diagrama se manejaba analíticamente, mediante el uso de las coordenadas de los puntos y la representación de los segmentos que los una (ver Evans más adelante).

También es posible por parte de este programa la construcción de segmentos auxiliares, tal como lo hace el estudiante ante un problema similar. El sistema explicado es más general de lo que parece, ya que modificando los axiomas y las rutinas "interpretadoras", se pueden demostrar teoremas de otras áreas que admitan algún modelo práctico; pero la eficiencia se puede aumentar en este caso agregando ciertas heurísticas muy particulares a la geometría, lo cual incrementa notablemente su capacidad. Con estas modificaciones es capaz de demostrar gran parte de los teoremas a nivel de escuela secundaria en pocos minutos.

P. C. Gilmore⁷ ha desarrollado un marco axiomático para el programa precedente, demostrando que si no se incluyen axiomas *constructivos*, el método es efectivamente una regla de decisión, ya que la demostración se puede efectuar sin usar más variables que en el enunciado del teorema, y el número de axiomas es finito, con lo que el universo de soluciones está perfectamente acotado: las heurísticas usadas sirven para guiar la búsqueda, tornando más eficiente el método.

⁷ Gilmore, P. C., *An examination of the geometry theorem machine*, *Artificial Intelligence*, 1, 3, 1970.

Al agregar el axioma "constructivo" *dos rectas no paralelas* (en el plano) *siempre se cortan*, el método indicado deja de ser una regla de decisión, reduciéndose a un algoritmo de demostración de teoremas, ya que se podrían generar un número infinito de puntos como sucesivas intersecciones de cada par de rectas determinadas por los puntos ya existentes en el paso anterior.

3.1.4. Un aprendizaje para la resolución "formal" de problemas.

La posibilidad de incluir el aprendizaje en los sistemas anteriores fue señalada en la mayoría de los casos por los propios autores como uno de los medios más promisorios de incrementar su efectividad. No se precisa en principio descubrir heurísticas especiales para el área considerada, sino que puede bastar adaptar las existentes, al menos en los primeros intentos. Es así que los métodos utilizados por Samuel, y otros a los cuales no hemos podido referirnos (programa de "regresión heurística" de Miller en 1967, trabajos de Tonge en 1965, Quinlan, Holden y Johnson, etcétera), pueden ser utilizados o tomados como base para el tema de interés.

Un trabajo que extiende las características comunes al programa de Gelernter, el "LT", y "SAINT", encontrados en la sección anterior, agregando una efectiva capacidad de aprendizaje, está dado por "MULTIPLE" ("programa multi-propósito que aprende") que consta de un programa encargado de las demostraciones de teoremas ("PP")¹ y uno del aprendizaje ("LP")².

Estos programas fueron escritos en LISP con el objeto de investigar las heurísticas aplicando sus técnicas de resolución a problemas de diversas áreas previamente formalizadas. El programa PP se basa en los principios anteriores, pero usando métodos avanzados de exploración en árboles, basado en evaluaciones de proposiciones de acuerdo a la "probabilidad" de que éstas lleven al resultado deseado y al costo asociado con dicha exploración. Estas evaluaciones están conceptualmente relacio-

¹ Slagle, J. y Bursky, P., *Experiments with a multipurpose, theorem-proving heuristic program*, J. ACM, 15, 1, enero de 1968.

² Slagle, J. y Farrell, C., *Experiments in automatic learning for a multipurpose heuristic program*, Comm. ACM, 14, 2, febrero de 1971.

Estos trabajos pueden ser considerados como ejemplos de la formalización a estudiarse en 3.4.3. (búsqueda heurística en grafos); se incluyen aquí con el fin de ilustrar la posibilidad de aprendizaje en las áreas consideradas.

nadas con las "comparaciones" del LT y SAINT y con el rechazo de proposiciones falsas en el programa de Gelernter; aquí un valor extremo de "cero" para la probabilidad indica la falsedad de la proposición, mientras que un "1" indica que es verdadera. Los valores intermedios indican una estimación de su valor de verdad, que se lleva a cabo de diversos modos, siendo éste el sitio en que se aplica el aprendizaje.

En el otro programa de Slagle ("SAINT") ya se hizo referencia a los árboles disyuntivos y conjuntivos, generalizándose aquí esta noción hasta abarcar todo tipo de función booleana. Es decir, el "valor de verdad" del origen es una función tal de sus ramas o subproblemas (aquí llamados "subproposiciones"). Dado un modo de transformar las proposiciones en sus subproposiciones (reglas de inferencia), el programa actúa alternadamente ramificando a partir de la proposición más "meritoria", y regresando los valores obtenidos a partir de ésta con fines de comparación. El "mérito" de cada proposición es inversamente proporcional al costo estimado de continuar la exploración a partir de ella (en tiempo, por ejemplo) y está relacionado directamente con la variación estimada de probabilidad de obtener una solución al nodo inicial, en comparación a la probabilidad anterior.

O sea, en lugar del "minimax" usual o del *equivalente* (desde el punto de vista teórico, no de eficiencia) alfa-beta, donde la exploración se lleva a cabo hasta varios niveles y luego se elige la mejor alternativa, en este caso la búsqueda de la solución se lleva a cabo de un modo realmente "dirigido" mediante el estudio en la dirección más promisoria. Una primera comparación entre el método "alfa-beta" y el de PP se realizó en el terreno de las proposiciones sobre finales de juegos; aquí las proposiciones son del tipo "el jugador A puede ganar desde esta posición", siendo las reglas de inferencia simplemente el generador de posibles movidas, y los "axiomas" las reglas del juego. Mediante una adaptación del alfa-beta al dominio de la demostración de "teoremas" se comparó ambos desempeños, obteniéndose resultados favorables al presente método. Con una pequeña modificación, se convirtió a los programas en efectivos "jugadores" (alcanza con que elija aquella posición que aumenta la probabilidad de la proposición "puedo ganar a partir de la situación actual"), notándose nuevamente la superioridad de PP.

Pero se obtiene un fabuloso incremento de efectividad (de acuerdo con las estadísticas publicadas en 1971) al funcionar el módulo de aprendizaje, "LP". Las probabilidades y el grado de confianza de cada proposición son consideradas funciones de las "características" de cada una. Una "característica" es una

función a valores reales de las metas propuestas, que indican la presencia o no de ciertos aspectos parciales; un ejemplo está dado por los términos del "polinomio" de Samuel. Se trata entonces de conocer la forma de la relación entre la probabilidad y el mérito de las proposiciones con las distintas "características", lo cual es perfectamente equivalente con el "aprendizaje por generalización" de Samuel, consistente en calcular los coeficientes de los distintos términos.

El aprendizaje es, sin embargo, mucho más completo, y se basa en la generación por parte del programa de "datos" de aprendizaje consistentes en las probabilidades calculadas para cada posición y los valores particulares de las "características", ponderando éstos y actuando sobre ellos el programa de aprendizaje mediante un procedimiento de regresión lineal para el cálculo de los coeficientes respectivos. De este modo se consiguen probar interesantes proposiciones de álgebra ("en un semigrupo la existencia de inversa e identidad a derecha implica la existencia de inversa a izquierda", etcétera), de química y de muchas otras áreas. El aprendizaje resultó rápido y dependiente de la cantidad de ejemplos presentados; además una práctica mediante la resolución de problemas en un tema particular aumentaba notablemente la efectividad en dichas áreas, pero contribuía también a un cierto conocimiento general. La razón para esta "generalidad" hay que buscarla en la notación unificada para los distintos tipos de problemas en el marco de la lógica de predicados, recordando que el aprendizaje se basa en la forma de las fórmulas y su similitud con la que se desea demostrar. Sin desmedro de reconocer la efectividad de estos métodos, señalemos sin embargo que se trata en realidad de un manipuleo "formal", y es difícil asignar algún tipo de "entendimiento" al programa respecto de los temas que maneja.

3.1.5. La solución de problemas.

El tema elegido para finalizar esta ejemplificación del método heurístico está en relación directa tanto con los métodos de demostración de teoremas como con el tema de la representación y la semántica, por lo que su ubicación es hasta cierto punto arbitraria. Por otra parte, los resultados de 3.2. y 3.3. también corresponden en gran parte a la "heurística", pero adquieren una importancia por sí mismos, mientras que, por ejemplo, la propuesta de McCarthy y su materialización por Black aunque relacionados con dichas áreas utilizan métodos similares en varios sentidos a lo anteriormente expuesto pero sin la utilización

de complicados esquemas de almacenamiento asociativo o de dispositivos de reconocimiento de configuraciones. Representan, sin embargo, un intento no ya de procesos de "ajuste de parámetros" sino un verdadero aprendizaje simbólico de conceptos y relaciones, lo cual ya había aparecido implícitamente en nuestra revisión si tenemos en consideración que para que "MULTIPLE" resolviese un problema era necesario "enseñarle" previamente los axiomas respectivos en los que basarse.

La razón de la misma investigación en inteligencias artificiales es la de obtener sistemas que ayuden al hombre en la resolución de sus problemas de modo tal que éste se limite a plantearlos, siendo función del programa establecer los pasos necesarios para su solución: en este sentido, los programas ya descritos han servido tanto para descubrir e implementar las capacidades necesarias de tales "resolucionadores de problemas" en áreas muy específicas (juegos, teoremas, etcétera) como para permitir abstraer las heurísticas más importantes a tener en cuenta en el curso de los procesos señalados.

Al diseñarse un tal "resolucionador", debe determinarse ante todo el lenguaje en que le serán presentados los problemas para su solución. Caben aquí varias posibilidades, según se desee orientar el sistema a la simpleza y generalidad del planteo o a la facilidad de solución. Si se elige un lenguaje irrestricto y amplio, como ser un lenguaje natural, será muy simple el enunciado de problemas de las más diversas áreas, mediante su expresión usual, pero la existencia de variados inconvenientes, los cuales analizaremos con más detalle en 3.3., hacen por el momento prácticamente imposible el liberarse del estricto marco de un área particular.

En el otro extremo se encuentran los sistemas basados en la lógica de predicados y otras formalizaciones similares, que si bien tienen la ventaja de la solución por medio de estudiadas técnicas de demostración de teoremas (como ser el principio de resolución), presentan en cambio una gran dificultad y complejidad en lo que respecta al enunciado de los problemas: si bien no nos dedicaremos en detalle en este libro a dichos intentos, tanto el "MULTIPLE" como el "Advice Taker", que veremos posteriormente, podrían ser encuadrados en dicho terreno. El GPS permite una mayor flexibilidad, aunque también el modo de presentar los problemas está dirigido al método de solución antes que a la facilidad de expresión.

Una variante interesante es la de los lenguajes "no determinísticos", en los que se modifica un lenguaje imperativo de programación (*lenguaje base*), permitiendo al sistema tomar decisiones propias en ciertos puntos indicados; el "REF-ARF"

de Fikes pertenece a esta clase. Este método tiene la ventaja de permitir el aprovechamiento de las unívocas estructuras de control y explicitación de los datos inherentes a un lenguaje de programación. Por último, podríamos citar un enfoque intermedio entre el anterior y el de la lógica, consistente en la demostración de la equivalencia entre la *existencia* de una solución para un programa no determinístico y la verdad o falsedad de un predicado que se le relacione, siendo entonces el método el de la expresión del programa no determinístico, la obtención del predicado asociado y la verificación de éste con métodos de demostración de teoremas, trazando los distintos pasos de la prueba con el fin de encontrar la solución ¹.

A partir de sus programas de ajedrez y de demostración de teoremas, Newell, Shaw y Simon crearon su ya famoso "GPS" ("solucionador general de problemas") dando forma a un esquema fundamentalmente recursivo de solución que ha servido de modelo a gran parte de los esfuerzos posteriores ². Este programa toma como fin adicional la simulación de los procesos análogos en los seres humanos, lo cual se efectúa mediante ciertas variaciones del modelo esencial.

El GPS actúa sobre *objetos* que constituyen su universo de acción; puede aplicar a éstos distintos *operadores*, de acuerdo con las *diferencias* observadas y a los planes trazados para la concreción de las distintas *metas*, las cuales pueden ser de diversos tipos ("transformar objeto A en B", "reducir la diferencia entre A y B", o "aplicar operador OP a objeto A"). No es difícil notar que prácticamente todos los problemas que debemos afrontar tienen dicha forma una vez realizada la abstracción apropiada: para ganar un juego, por ejemplo, nos presentamos el problema de pasar del "objeto" estado inicial al "objeto" estado ganador mediante aplicación de "operadores" consistentes en las posibles movidas y la concreción de diversas metas intermedias mediante la observación de las diferencias entre el estado actual y nuestro objetivo, y así con todas las demás áreas.

Con cada tipo de meta se encuentra asociado un método para lograrla. En el caso de la transformación de los objetos, el proceso consiste en comparar A y B y encontrar su diferencia, plantearse la meta de reducir ésta (aquí se usa el proceso respectivo), y una vez hecho esto repetir el proceso con el resul-

¹ Manna, Zohar, *The correctness of nondeterministic programs*, *Artificial Intelligence*, Vol. I, 1970.

² Newell, A., Shaw, J. C., Simon, H. A., *Report on a general problem-solving program*, Proc. of the International Conference on Information Processing, Paris, 1959.

tado obtenido (si bien desaparece la diferencia original, puede haberse introducido otra no intencionalmente). Para reducir la diferencia entre dos objetos, se busca el "operador" adecuado (de una lista a tal efecto, clasificada adecuadamente), y si se encuentra uno apropiado se establece la meta de "aplicar el operador". Esta última se resuelve mediante una comparación de las condiciones de aplicabilidad del operador con las características del objeto: si coinciden se aplica sin más, y si hay alguna diferencia se establece la submeta de reducir ésta y luego se aplica el operador. Existen además varias heurísticas para reducir la cantidad de metas que se plantean; en particular se dejan de lado todas aquellas cuya aparente "dificultad" es mayor que la del problema original a resolver.

En el esquema anterior se puede apreciar la necesidad, tantas veces postulada, de poseer buenos métodos de "reconocimiento" de formas y de conceptos, tanto para determinar el operador adecuado como para saber cuándo un problema presentará en su resolución más dificultad que otro. Los humanos actuamos guiados por una cierta "intuición", la cual puede tener quizá su análogo en un adecuado proceso de aprendizaje de los casos ya resueltos y en una adecuada clasificación de métodos; el análisis de "formas" puede ser útil al presentarse un problema visualmente o en un lenguaje simbólico, caso este último en que un adecuado análisis sintáctico puede darnos "pistas" para la resolución; mientras que el de "conceptos" podría ser considerado como más general y difícil de definir. En estos sentidos las secciones siguientes servirán para poner de manifiesto las tentativas efectuadas de realización.

Debido principalmente a la falta de respuesta a los principales inconvenientes que se presentan en cuanto al reconocimiento de los problemas según su tipo, condición necesaria para solucionar una situación que no hubiese ocurrido *exactamente* en las mismas condiciones, se llegó a un replanteo de los principios fundamentales de las investigaciones, arribándose a la conclusión de que es necesario, antes que la creación de métodos efectivos de resolución, el establecimiento de modos de transformar la información acumulada (la *experiencia*) en guías para el mejoramiento de los métodos existentes (incluyendo ahora reales mejoras *cualitativas*), así como el estudio de la forma óptima de acumular dicha información (la *representación* interna), y, fundamentalmente, del mismo proceso de comunicación de los datos pertinentes al sistema. Es entonces necesaria la capacidad de análisis de las experiencias de resolución de problemas y la categorización de los métodos que hayan resultado efectivos, generalizándolos adecuadamente si es necesario (cuando un

científico “reconoce” un problema como matemático, por ejemplo, recuerda las técnicas y los trucos adecuados).

Para permitir dichas generalizaciones y comparaciones se necesita poseer adecuados lenguajes *descriptivos* (obteniéndose resultados importantes como los de Evans) e inclusive buenos modelos interiores, lo cual nos lleva al tema de la *semántica*, al cual nos referiremos posteriormente. La dirección para una primera solución a la cuestión de la resolución de problemas por parte de sistemas con capacidad de modificación (en principio sobre la base de “consejos”) fue marcada ya en 1958 por McCarthy, sin acentuar demasiado explícitamente el aspecto del “entendimiento” y con una manipulación de carácter fundamentalmente formal.

El “Advice Taker” (literalmente, “toma-consejos”) propuesto por Mc Carthy representa un intento tal de programas con una capacidad de aprendizaje similarmente efectiva a la de los seres humanos, siendo el método propuesto de implementación la resolución de problemas manipulando sentencias en lenguajes formales, con capacidad de mejorar su rendimiento *no* mediante cambios en la estructura del programa sino a través de la comunicación directa del “entrenador” en el “AT”. La diferencia con los programas anteriores está en que las heurísticas y los distintos procesos forman parte más del lenguaje usado que del programa en sí; de este modo se da un gran paso hacia la obtención de máquinas inteligentes, mediante la posibilidad de fácil expresión de los cambios de conducta. Se trata con AT de estudiar *cómo enseñar* a un programa, como paso previo a un autoaprendizaje, y de considerar la posibilidad de que un programa tenga “sentido común”, entendiéndose por este término, como ya lo definimos, el que “automáticamente deduzca por sí mismo una amplia clase de consecuencias inmediatas de todo lo que se le diga y lo que ya sabe” (McCarthy).

El período de instruir al AT se diferencia de un esfuerzo de programación en que en este último se manejan instrucciones *imperativas*, mientras que en el primero las sentencias son de tipo *declarativas*, lo cual permite la deducción de consecuencias lógicas y la ignorancia del “profesor” de la estructura y los estados actuales del programa. Para su desempeño deberá poseer una regla de inferencia y rutinas de “deducciones inmediatas”, que le permitan manejar los objetos y sus relaciones, siendo la operación consistente en la aplicación de las rutinas de deducción a las sentencias que se presenten, y obedeciendo los imperativos que así resulten.

Para evitar la dificultad que presenta el manejo de un lenguaje natural (ver 3.3.), se optó en el proyecto para el AT por

un lenguaje definido recursivamente y con una estructura lógica muy coherente y manipulable. Así, la sentencia "en (yo, casa)" significa "yo estoy en casa", "poder (ir [casa, trabajo, caminando])" significaría "es factible ir caminando de casa al trabajo", etcétera; la ventaja de esta notación funcional es la facilidad del manejo de las expresiones así formadas. Ante una declaración tal como "estoy sentado ante mi escritorio en casa y deseo ir al aeropuerto. Mi automóvil está en casa también", formalizada adecuadamente, el AT debería —si previamente se le había hecho saber que es posible ir manejando desde la casa, etcétera—, contestar que "camine hasta el automóvil y maneje el coche hasta el aeropuerto", solucionando así un problema que, aunque trivial, da una idea de las posibilidades de los programas con "sentido común". Veremos cómo se la ingeniería para resolver estos problemas al estudiar el programa de Black.

A partir del anterior modelo propuesto por McCarthy se idearon gran cantidad de variantes, saliendo inclusive del contexto del lenguaje formal y permitiendo la comunicación en un inglés restringido; el esquema anterior tiene la virtud, sin embargo, de concentrar los esfuerzos más en el problema deductivo que en la resolución de ambigüedades y demás inconvenientes del lenguaje natural. En 1964 presentó F. Black una tesis de doctorado donde se describió un sistema bastante similar al del AT³.

El lenguaje usado por Black consiste en "palabras" (cadena de símbolos básicos y dígitos), "variables" (son: x, y, z, n, v, w); "constantes" (los verbos y sustantivos: tener, automóvil, etcétera) y las "frases complejas" (formadas por una constante seguida por una o más frases separadas por comas y entre paréntesis). Las sentencias se clasifican en "directas" y "condicionales"; las primeras son cualquier frase (como "en [yo, casa]") y las segundas consistentes en una o más frases (*antecedentes*) separadas por comas, una flecha y otra frase (*consecuente*). Una primera regla deductiva sería la siguiente: "si el antecedente de una condicional es igual a una sentencia directa (de las que ya se "aprendieron"), eliminar dicho antecedente". Si se sabe, por ejemplo, que "Juan es argentino" —simbolizado, por ejemplo, "argentino (Juan)"— y también que si Juan fuese argentino sería sudamericano-argentino (Juan) → sudamericano (Juan)— entonces es inmediato que Juan es sudamericano. Más importantes son las reglas que permiten sustituir las variables en ambos lados de la sentencia condicional, y la regla de colocar el

³ Black, Fischer, *A deductive question-answering system*, Ph.D. dissertation, Harvard, junio 1964. Reproducido en *Semantic Information Processing*, Minsky ed., 1968.

antecedente de una sentencia anterior en lugar del actual que coincida con el consecuente de aquella.

En el ejemplo anterior, si se sabe ya que "argentino (Juan)", que "argentino (x) \rightarrow sudamericano (x)" y que "sudamericano (x) \rightarrow americano (x)", aplicando el reemplazo de variables ("Juan" en lugar de "x") y luego la primera regla enunciada a las dos sentencias iniciales, se tendrá "sudamericano (Juan)" y "sudamericano (Juan) \rightarrow americano (Juan)". Las *preguntas* son de dos tipos: las que se contestan por sí o no, que consisten en presentar la sentencia en duda, y las que solicitan datos, en las que se colocan variables en representación de las incógnitas; la pregunta "Padre (Pedro, x)" correspondería a "¿quién es el padre de Pedro?"

Además de los datos del problema, se considera que el programa ya ha aprendido varios principios generales (como "argentino (x) \rightarrow sudamericano (x)" antes), inclusive respecto de las definiciones de las palabras (tal como "grande (x) \rightarrow no (pequeño (x)), etcétera). Esta "cultura general" es la que permite la existencia de un "sentido común". Para la resolución de los problemas planteados, el trabajo se hace "hacia atrás", intentando encontrar coincidencias entre la pregunta y las frases que sean semejantes, reemplazando las constantes por las variables de la pregunta, finalizando el proceso cuando se encuentra una sentencia directa en las condiciones requeridas, y usando los métodos de deducción apropiadamente para las condicionales.

Este programa solucionó el "problema del aeropuerto" ya mencionado, así como el del "mono" ("hay en una habitación bananas colgadas del techo, un mono que las quiere alcanzar y un cajón en el suelo", se pregunta si el mono puede alcanzarlas y se resuelve indicando que debe buscar el cajón, colocarlo debajo de las bananas y subir a él), e inclusive teoremas de lógica codificados adecuadamente (probó así uno que LT no pudo demostrar).

El ya mencionado sistema "REF-ARF"⁴ permite, a diferencia de los anteriores, un enunciado mucho más sencillo de los problemas. Como ya indicamos, un lenguaje no determinístico es uno tal que permite dejar a criterio del "solucionador" la elección de las acciones a tomar en ciertas circunstancias. En este caso se incluye un modo de indicar tales alternativas mediante la instrucción *seleccionar*, conjuntamente con la cláusula de *condición* que aclara los valores de interés; así el rango de *seleccionar* indica el espacio de soluciones, y *condición* da las restricciones que se pretenden satisfacer.

⁴ Fikes, R. E., *REF-ARF: A system for solving problems stated as procedures*, *Artificial Intelligence*, Vol. I, 1970.

Tomando un ejemplo de Fikes, con un lenguaje base del tipo del ALGOL, el enunciado "comenzar; enteros B, C; B ← seleccionar (0,9); C ← seleccionar (0,9); condición B + C = 15; fin" sería equivalente a "encontrar un par de enteros entre 0 y 9 tal que sumen 15".

REF-ARF resuelve los problemas así enunciados mediante un complejo método de obtención, a partir de dichos enunciados, de un sistema de ecuaciones sobre varias variables, creadas en el proceso de *interpretación*. Ante cada bifurcación condicional, se crean varias alternativas en cuanto a la continuación de la interpretación, realizándose un análisis de cada una de éstas. Por ejemplo, en un caso ideal, si debemos elegir, para resolver un problema, entre tomar las acciones A o B, y se nos aclara que éstas nos llevan a situaciones distintas, deberemos analizar cada una de las alternativas por separado para verificar cuál es la acción correcta a tomar, de acuerdo con nuestros fines. Y en esto es exactamente en lo que consiste el "análisis de casos" del REF-ARF.

Como analizar todos los casos es casi siempre imposible, se utilizan diversas heurísticas para *guiar* la *búsqueda* hacia las alternativas más prometedoras, existiendo para ello varias heurísticas, tales como calcular la "distancia" que habría que recorrer en cada caso, tomando la alternativa que la minimice (o sea, probar primero las soluciones más simples).

Una vez que se decidió la alternativa, se trata de resolver las ecuaciones que hayan resultado de su interpretación; si se obtiene alguna solución se finaliza el proceso, y caso contrario se considera la siguiente. Para la solución de las ecuaciones existen también varias heurísticas apropiadas. Acotemos aquí que el proceso mencionado puede hacerse más eficiente si se intercalan el análisis de casos y la solución de ecuaciones, lo que permite eliminar tempranamente las alternativas que ya se vean no conducen a una solución.

El sistema resolvió fácilmente el problema del "mono" ya mencionado, así como varios acertijos matemáticos. Con varias modificaciones, propuestas por el mismo autor, se podría incrementar el rango de aplicabilidad y la efectividad del sistema. Quizá se podría lograr un cierto "sentido común", por otra parte, mediante el uso de rutinas parametrizadas, aprendidas con anterioridad.

En 3.3. volveremos al tema de la resolución de problemas, en el caso en que éstos están expresados en un lenguaje natural.

3.2. "Pattern recognition".

El "pattern recognition", o "reconocimiento de configuraciones"¹ es quizá uno de los temas más importantes en la investigación en el área de inteligencias artificiales. Si bien en un principio se lo caracterizó como el reconocimiento de caracteres ópticos, hoy en día está asociado a temas fundamentales como la formación de hipótesis, generalización a partir de los casos individuales —inferencia inductiva—, etcétera. Su meta principal es la de realizar una abstracción que permita separar propiedades capaces de distinguir un concepto de otros, permitiendo así realizar una clasificación, ya sea de un problema (permitiendo así la aplicación de métodos específicos) o de un cierto "estímulo de entrada", buscando entonces la respuesta adecuada.

La razón primordial para esta necesidad, radica en lo que podríamos denominar "falta de perceptividad" de una máquina electrónica: si bien su capacidad de cálculo supera en velocidad y sofisticación a todo lo humanamente factible, tanto las instrucciones como los datos deben ser cuidadosamente preparados, ajustándose a un rígido formato. Y aquí radica una de las diferencias esenciales entre el comportamiento inteligente tal cual lo concebimos, y el actual de las computadoras. Aun en el caso de los organismos más elementales en la escala zoológica, se verifica que reaccionan ante el medio ambiente de un modo adecuado a sus necesidades (ya sea instintivamente o de un modo inteligente), pero su comportamiento difícilmente podrá especificarse mediante una tabla de decisión (a temperatura "x" aplicar acción "A", etcétera), ya que el medio actúa en una escala *continua*, no discreta.

Del mismo modo, cuando escribimos, producimos desviaciones de las verdaderas formas de las letras; no acostumbramos a "dibujarlas" con extremo cuidado ni a ceñirnos estrictamente a ciertos moldes y encasillarlas para permitir que una máquina las interprete automáticamente con ayuda de una célula fotoeléctrica. Al hablar no entonamos siempre exactamente del mismo modo cada letra y no nos cuidamos de separar adecuadamente entre palabras, lo que eventualmente permitiría a nuestro automática entendernos mediante un micrófono. Sin embargo, con nuestros semejantes podemos comunicarnos perfectamente: ellos reconocen la "configuración" (el *pattern*) subyacente a nues-

¹ La traducción de "pattern recognition" como "reconocimiento de configuraciones" es la que, a juicio de los autores, es la más adecuada. Con este término, que en adelante abreviaremos "RC", incluimos también a la percepción, discernimiento y clasificación de configuraciones.

tras palabras; en síntesis, es esencial para una comprensión cabal de las señales que obtenemos de nuestros órganos sensitivos la capacidad de reconocer un mismo concepto a través de distintas manifestaciones.

Por otra parte, es necesario para un comportamiento inteligente una cierta capacidad de "comprensión", concepto algo difícil de definir. Puede ser útil para nuestros fines considerar que "entender" significa "tener capacidad de descripción", y mientras más detallada sea esta última, mayor el "grado de comprensión" que se obtiene. Así, una de las metas de investigaciones en RC podría ser expresada como "obtener sistemas automáticos con una alta capacidad de descripción del medio ambiente".

Del mismo modo, de poco servirán los métodos heurísticos o los algoritmos que poseamos para resolver los diversos problemas si no contamos con un esquema que nos indique cuál usar en cada caso para obtener una real eficiencia. Y estas reglas para determinar procedimientos adecuados, las obtendremos justamente al tener un panorama claro de la estructura del problema en cuestión.

La misma capacidad de aprendizaje depende del RC: sin realizar inferencias y relacionar los ejemplos con conceptos más generales, el conocimiento que se adquiere consistirá en un mero "aprender de memoria", que difícilmente nos pueda brindar luego capacidad para realizaciones intelectuales propias.

La solución a los problemas que se plantean en relación al RC está lejos de ser conocida totalmente en la actualidad. Quizá la principal razón radique en que el mejor "reconocedor" de configuraciones (el ser humano), no puede explicarse conscientemente qué procesos ocurren en su mente al realizar una "percepción y clasificación". Si se nos pide describir un concepto, seguramente lo haremos basándonos en un ejemplo; pensaremos en él y trataremos de extraer las propiedades que posee en común con otras ocurrencias del mismo concepto (por ejemplo, diremos que la letra "A" consiste en dos líneas verticales y una horizontal) pero sólo al compararlo con otro distinto descubriremos que faltaba describir alguna propiedad fundamental (la "H" tiene también dos segmentos verticales y una barra que los cruza: entonces es esencial, para poder distinguir la "A" de la "H", agregar en el concepto de "A" que las líneas verticales deben unirse), y aun así nuestra descripción no será exacta, ya que podrán existir casos que no se ajusten exactamente al modelo, pero sean también clasificados como el mismo concepto (por ejemplo, interpretaremos los caracteres de la figura 1 como

letra "A", aun cuando los segmentos verticales no se unen exactamente o esta unión consiste en un segmento adicional).

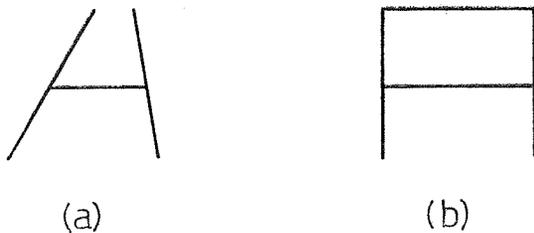


Figura 1.

Aún más importante es la relación de la clasificación con el contexto: no confundimos un nueve con un seis cuando estamos leyendo en una cierta posición, pero será difícil distinguirlos en una hoja en donde figuren aislados.

La inhabilidad del hombre para definir los conceptos (a la que se acaba de hacer referencia) se traduce en la imposibilidad de programar algoritmos efectivos de "reconocimiento" mediante el uso de descripciones precisas. Son muchos los casos en que diferencias imperceptibles para el humano (al menos conscientemente), no permiten la identificación de dos conceptos por parte de un autómata. La mejor solución para este problema es la de permitir que la máquina determine por su cuenta los factores (o al menos las respectivas ponderaciones) que deberá considerar para clasificar un ejemplo. En otras palabras, un método de RC sólo será efectivo si se admite la posibilidad del aprendizaje y la acumulación propia de la experiencia necesaria.

En vista de lo anterior, se ha trabajado intensamente en los últimos años en la obtención de modelos que permitan el desarrollo de la percepción y clasificación mediante el aprendizaje.

Los más generales son los que corresponden a las redes interconectadas al azar, o "redes nerviosas"², donde absolutamente todo conocimiento se adquiere de la experiencia. Se trata aquí, en líneas generales, de "discretizar" la información de entrada, y de asignar a cada par de "unidades de información" un factor de correlación o "peso" que indique la necesidad, o no, de su presencia conjunta en cierto concepto, a partir de los ejemplos que se recibieron durante el período de aprendizaje. Este enfoque, que corresponde en cierto modo a la biónica, será estudiado en más detalle en la sección 3.4. ("Modelos generales y

² "Random nets" y "neural nets", respectivamente.

enfoques formales”), pero cabe acotar que esta “generalidad” se obtiene al costo de una capacidad ínfima de aprendizaje y, por el momento, los resultados obtenidos por otros métodos superan ampliamente, en los problemas particulares, al comportamiento de estos sistemas “probabilísticos”. La razón de esto quizá debe ser buscada en el no aprovechamiento, en estos sistemas, de una rica estructura inicial.

En los otros casos, se trata de determinar ciertas “propiedades” o “características”, que deben estar presentes en los ejemplos, y podríamos englobar a la mayoría de estos procedimientos dentro de la categoría de métodos heurísticos. Señalemos que en algunos casos los factores son dados por el programador y la máquina sólo se concentra en decidir las ponderaciones adecuadas, mientras que en otros se generan las propias características a tenerse en cuenta (caso, por ejemplo, de Uhr y Vossler en reconocimiento visual). En ocasiones queremos determinar características generales (como distinguir los objetos presentes en un dibujo y diferenciarlos entre sí y con el “fondo”, el caso del programa SEE de Guzmán) u obtener conceptos a partir de las observaciones y la enseñanza (el programa de P. Winston, por ejemplo).

A su vez, los métodos heurísticos de RC se pueden clasificar según su carácter de “secuenciales” o “paralelos”. En el caso “secuencial” las propiedades o características se investigan una por una, y el seguir o no el proceso, o el paso siguiente, dependen del resultado anterior. Este tipo de programa se adapta naturalmente a las computadoras actuales, debido a que es en sí una secuencia de bifurcaciones; además posee una estructura de árbol, procesable fácilmente en LISP (ver capítulo 1). Por ejemplo, si se quieren diferenciar un cuadrado, un rectángulo, un triángulo equilátero, uno cualquiera y una circunferencia, alcanza con investigar adecuadamente la sucesión de propiedades “contornos lineales”, “cantidad de lados”, “igualdad de lados”. Con sólo una prueba (si se ordenan adecuadamente), se puede identificar a la circunferencia (ver figura 2).

Un ejemplo donde es útil este método es en los diversos casos de diagnóstico. Una característica sobresaliente es la posibilidad de aprovechar para el reconocimiento secuencial las técnicas de exploración en árboles que se utilizan en los juegos y otros programas heurísticos.

El proceso en “paralelo” es, en cambio el que parece ser más natural al ser humano. Se reciben todos los datos simultáneamente, y se combinan adecuadamente los detalles. Si consideramos que la retina ocular está compuesta por millones de “puntos” que reciben información simultánea enviándola al cerebro

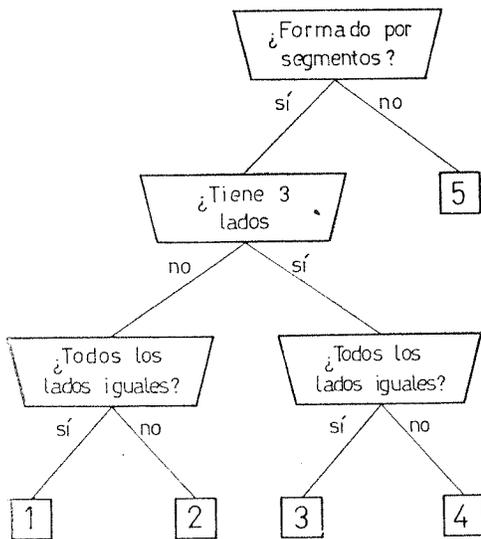
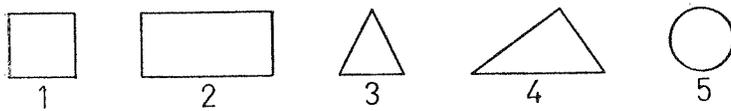


Figura 2. Posible árbol de decisiones para diferenciar las figuras 1, 2, 3, 4 y 5.

posteriormente, el modelo parece el más adecuado para representar la percepción humana. La analogía es más evidente si consideramos los estímulos provenientes de distintos órganos sensoriales: tacto, vista, oído, etcétera, al mismo tiempo. Los conceptos los formamos al relacionar toda esta información³.

Además de pesar en la adopción del sistema paralelo las mencionadas similitudes, existen otras ventajas sobre el método secuencial. Por ejemplo, algunas propiedades no son fáciles de decidir simplemente por un "sí" o un "no", sino que se debe efectuar alguna ponderación para facilitar el reconocimiento (por ejemplo, tanto un triángulo isóceles como el triángulo 4 de la figura 2, no tienen "todos los lados iguales", pero podemos decir que en cierto modo el isóceles los tiene "más iguales" que el 4,

³ Esta parece ser una característica esencial al comportamiento inteligente superior; algunas experiencias indicarían que en los animales no siempre se encuentra presente esta capacidad de coordinación.

y "menos iguales" que el 3). Otra ventaja del reconocimiento en paralelo es la posibilidad de modificar las reglas de decisión de un modo "continuo" con la finalidad del aprendizaje, mientras que en el caso secuencial, donde todas las decisiones son de tipo binario, los cambios deben ser más radicales (cambio de la estructura del árbol). Si los factores se evalúan numéricamente, por ejemplo, se puede variar el peso de una decisión en un cierto porcentaje de acuerdo a los resultados de la experiencia.

El caso más trivial (y en cierto modo más inútil) de reconocimiento en paralelo es el de las "comparaciones" o de los "moldes". Aquí se puede considerar que la "pregunta" a efectuar es una sola: no se investigan diversas propiedades sino que se compara el ejemplo con un "objeto-tipo" que aproxime adecuadamente al concepto. Para aclarar las ideas, tomemos el caso del reconocimiento de figuras geométricas en el plano. La idea será entonces la siguiente: se analiza el dibujo en cuestión, superponiéndolo a una "plantilla" que se posee a priori, buscando que coincidan. Si esto último sucede, se habrá identificado correctamente la figura, y el problema estará resuelto.

Sin embargo, en la mayoría de los casos se presentan varias dificultades con este método. El principal inconveniente es que las figuras no se dibujan siempre correctamente, en la misma posición, etcétera, con lo que un pequeño cambio en la orientación o en el tamaño destruye la posibilidad de coincidencia. Otras dificultades, se deben a diferencias en el contraste, borrones y todo otro tipo de "interferencias" que el humano pasa por alto en su lectura consciente, pero que debemos tener en cuenta al programar una máquina que aún no se ha "habituado" a ha-

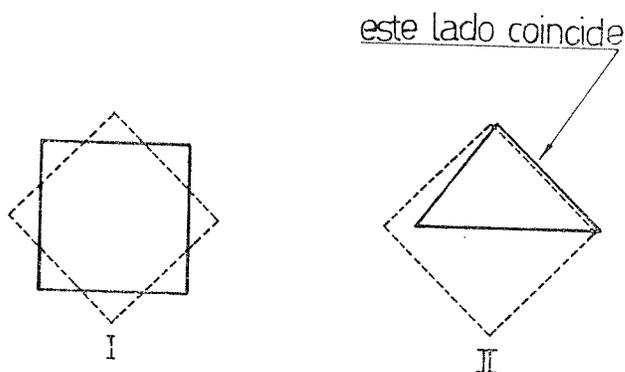


Figura 2. Se compara un cuadrado (inclinado) con los moldes I y II. Se decide que la asignación más acertada es la segunda, por encontrarse más coincidencias.

cerlo. Una primera solución podría ser la de efectuar varias comparaciones y luego elegir aquella que “anduvo mejor”; pero este método no es muy satisfactorio, ya que aún así se pueden adoptar decisiones erróneas. Por ejemplo, en la fig. 2, vemos cómo un cuadrado (dibujos con líneas de puntos) se identifica como un triángulo al compararlo con los moldes I y II.

Se puede aquí mejorar sensiblemente el sistema: se trata de “colocar en posición” la entrada y de reducirla o aumentarla de tamaño proporcionalmente hasta alcanzar dimensiones standard, lo cual se puede hacer de diversos modos. Uno de ellos, descrito años atrás por Minsky⁴, es esencialmente simple: se trata de inscribir la figura en un triángulo fijo (cuya elección determinará las elecciones finales), y variar sus medidas proporcionalmente hasta que toque a los tres lados del triángulo. Sin embargo, aun en estos casos la decisión será muy fácilmente errónea (ver fig. 3). En el caso de reconocimiento de caracteres alfabéticos, el problema es muy similar.

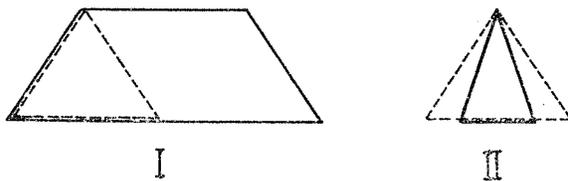


Figura 3. Ahora comparamos un triángulo (ya colocado en posición y tamaño correctos) con las figuras I y II: de nuevo elegimos erróneamente, ya que tomamos el I.

Como se puede apreciar, estos métodos carecen, en general, de la efectividad necesaria. La razón hay que buscarla en la idea subyacente de comparar ejemplos con un modelo que es a su vez un ejemplo. De este modo será siempre difícil llegar a una solución; es común incluir dentro de una única clasificación a objetos de los más variados tamaños, que tienen en común, sin embargo, ciertas propiedades globales (cantidad de lados o ángulos, etcétera), y ciertas características locales muy bien definidas y distinguibles: puntos de intersección de dos o tres líneas, ángulos agudos u obtusos, convexidades y otras. Y en ese sentido es que se encuentran los verdaderos logros en RC.

En efecto, el análisis detallado de las características globales y locales de la percepción (o más generalmente de un problema) es el que nos permite en definitiva obtener un recono-

⁴ Minsky, 1961, *op. cit.*

cimiento y una clasificación correctos. Es fundamental aquí el concepto de "propiedad". Minsky, en el artículo ya citado define "propiedad" como una función sobre el conjunto de objetos a clasificar (por ejemplo, figuras geométricas) que puede tomar dos valores: 1 ó 0 (interpretables como "verdadero" o "falso", respectivamente), de tal modo que dada una cierta cantidad de propiedades se determina para cada objeto un vector de ceros y unos que lo describe; a estos vectores los denomina "características" y sirven como denominación para las categorías en que se pueden dividir los objetos considerados, con la ventaja adicional de tener la estructura conveniente de listas. Una generalización está dada luego por el mismo Minsky, ante la necesidad de descripciones de elementos componentes y sus relaciones; se trata de especificar listas en forma de pares ordenados, donde el primer miembro indica una relación y el segundo es, a su vez, una lista ordenada que indica los elementos en esa relación. Por ejemplo: "(mira, [Juan, Ana])" indica el concepto "Juan mira a Ana" (comparar 3.1.5.).

La obtención de un método de generar propiedades que sirvan para una buena clasificación es un problema que se plantea inmediatamente. Una primera solución puede basarse en la propuesta por Selfridge y consiste básicamente en aplicar ciertas secuencias de transformaciones a la entrada de modo tal que de la última de éstas se obtenga una información numérica; si la secuencia de transformaciones termina en un número dado, se puede decir que la propiedad definida por esa sucesión particular se cumple o no. El conjunto de transformaciones posibles debe ser pre-especificado, pero la elección en cada caso puede ser obtenida al azar. Si al experimentar resulta que cierta secuencia es efectiva (por ejemplo, que un cierto número se obtiene en gran cantidad de casos), entonces generamos nuevas propiedades basándonos en éstas.

En general, resulta poco ventajoso el uso de las "listas de propiedades" de longitud fija si lo que se pretende es lograr la creación de nuevas propiedades y la generalización, objetivo para el cual es más apto el método "descriptivo" señalado. Sin embargo, su tolerancia ante pequeñas variaciones (debido a su paralelismo) unida a la facilidad de implementación y posibilidades de aprendizaje, hace de las "property-lists" un elemento indispensable en muchos sistemas de RC.

Por supuesto todo lo que estamos viendo se aplica (con las identificaciones apropiadas) a los diversos tipos de problemas de percepción, reconocimiento y clasificación que todos efectuamos constantemente. En la Universidad de Stanford, por ejemplo, se está desarrollando en la actualidad un programa

de reconocimiento de lenguaje oral, que incorpora las técnicas más modernas de aprendizaje; siendo factible que en un futuro quizás no muy lejano podamos comunicarnos con las computadoras hablando en nuestra lengua, o, más probablemente, en un lenguaje especial que no induzca posibilidades de error, mediante la ausencia de ambigüedades.

En el área de reconocimiento visual es sin embargo donde se comprenden más fácilmente los métodos y donde se han logrado importantes resultados, siendo ésta la razón por la cual le brindamos a este tipo de RC un lugar preponderante en nuestro estudio. En lo que sigue nos referiremos a experiencias concretas realizadas en el campo del "reconocimiento de configuraciones", en especial en percepciones visuales, y en la "formación de conceptos", mientras que en 3.3 profundizamos en lo referente a la formación y manejo de las representaciones interiores de los datos percibidos. Este tema está en cierto modo incluido en el actual, pero los estudios realizados al respecto implican generalmente procesos "cognoscitivos" de un nivel superior a los relacionados a las primeras etapas de la percepción. Esta parte inicial correspondiente al RC puede considerarse asociada principalmente con las relaciones de tipo sintáctico entre conceptos o "partes" de conceptos, mientras que la segunda se refiere más al aspecto semántico, el significado y las interrelaciones causales. Las distintas áreas no están sin embargo bien definidas, interactuando incluso los factores como en el caso del recurso a información semántica para los procesos de desambiguación, por lo que nuestra separación es en gran medida arbitraria.

3.2.1. Reconocimiento de caracteres alfabéticos y figuras geométricas.

Es inmensa la cantidad de trabajos que se han llevado a cabo en el área del reconocimiento de letras y dígitos. Prácticamente todos los programas efectivos se basan en el análisis local de "características", y en la actualidad se cuenta con buenos sistemas automáticos de lectura de caracteres ópticos, inclusive de uso comercial. La importancia de la obtención de dispositivos de lectura de información impresa no necesita ser remarcada; es suficiente considerar el costo en que se incurre al perforar los datos. Los sistemas comerciales, sin embargo, no poseen en general una capacidad de aprendizaje ni la posibilidad de interpretar símbolos bastante distorsionados o escritura cursiva. En lo que hace el reconocimiento de símbolos bien se-

parados y posicionados adecuadamente, se trabaja en estos dispositivos comerciales mediante la identificación de características tales como presencia de barras horizontales (en la "A" y la "H", por ejemplo) distancias relativas, curvaturas, etc. Este análisis local indica en oportunidades la realización de un nuevo análisis general mediante la normalización del tamaño del carácter en consideración. Otras técnicas usadas aprovechan diversos procesos de correlación con ejemplos o "moldes", aunque su uso es poco efectivo como ya vimos, en el caso de comparaciones generales, siendo útil principalmente para las características locales. Un modo de detectar estas últimas es mediante una "auto-correlación" consistente en desplazar el símbolo en consideración, superponiendo la imagen original con la obtenida. Como ya dijimos, la variedad de posibilidades es grande¹, y su estudio cae fuera del campo de la inteligencia artificial.

Lo que sí nos interesa de los procesos de reconocimiento es la posibilidad de generalizarlos a otros tipos de percepción, y las posibles técnicas de aprendizaje. Estas últimas se basan generalmente en un estudio estadístico de las propiedades en relación a las diversas configuraciones. Los métodos más "probabilísticos" los veremos posteriormente (3.4), pero acá nos interesan las heurísticas conectadas con el proceso de RC. La tarea de un programa con capacidad de aprendizaje para el reconocimiento de caracteres consiste en tabular para cada letra presentada en un período inicial de entrenamiento la frecuencia observada de cada característica; estas frecuencias no serán de 0 ó 100 % sino que variarán debido a las deformaciones e interferencias. Una "A" un poco inclinada, por ejemplo, puede dar lugar a la asignación de la característica "una línea horizontal por el centro la intersecciona en 3 puntos", lo cual no sería cierto en condiciones perfectas. Luego, al experimentarse con la clasificación obtenida, se analiza cada letra ahora presentada llevando a cabo los diversos "tests" correspondientes a las distintas propiedades obteniéndose así la "probabilidad" de cada símbolo y respondiendo con el más posible.

En lo anterior se suponían dadas las "características", siendo la tarea del programa la de ajustar los parámetros necesarios. Como en el caso del aprendizaje en los juegos al cual ya nos referimos, resulta también interesante la posibilidad de que el programa posea la habilidad de "crear" otros por su cuenta. Un ejemplo práctico fue presentado ya hace varios años

¹ Ver en general los diversos números de "Information and Control" y otras publicaciones especializadas.

por Leonard Uhr y C. Vossler², quienes se basaban en la generación de “operadores” consistentes en pequeñas matrices, las cuales se superponían luego con las diversas zonas del carácter a reconocer; éstos se generaban obteniéndolos de algunas de las figuras presentadas, al azar, o mediante operaciones lógicas sobre operadores ya existentes. Resultó bastante efectivo el método no sólo para símbolos numéricos o alfabéticos sino incluso para el reconocimiento “auditivo” mediante el análisis de fonogramas digitalizados.

Para el reconocimiento y clasificación de figuras el proceso es esencialmente similar, aunque en general se supone a éstas de contornos más definidos y con menor número de variaciones. Esto es debido al tipo de experiencias que se llevan a cabo con los resultados; en el caso de letras lo que se desea es una percepción de los símbolos tal como son escritos normalmente, lo que implica diversos grosores e “interferencias”, las cuales se compensan generalmente con el recurso de técnicas de “relleno” de figuras, “afinamiento” de líneas, etc., obligando a una etapa de pre-proceso, actuando luego el programa de reconocimiento propiamente dicho.

Para las figuras geométricas, más regulares en general que las letras, se puede obviar en general el paso previo mencionado, requiriéndose en cambio buenas “descripciones” de la imagen, no consistente ya de una de 27 letras (o de 10 dígitos) sino de una cantidad potencialmente infinita de figuras posibles. Estas descripciones, generalmente en forma de listas, deben contener información sobre los vértices y líneas presentes para permitir luego a un programa de RC la clasificación de las figuras existentes. Estas especificaciones previas pueden a su vez ser producidas por los “seguidores de líneas” los cuales constituyen un tipo especial de programas de reconocimiento óptico.

Sobre esta base realizó Evans su famoso programa “ANALOGY”³ dedicado a la solución de tests de inteligencia. Estos “problemas de analogía” son comunes en los Estados Unidos, formando parte incluso de los exámenes de ingreso a las universidades, tanto en los niveles inferiores como para graduados, y tienen en general la siguiente forma (ver Fig. 4): se presentan ocho figuras A, B, C, I, II, III, IV, V, y se pide al sujeto que “encuentre la regla por la cual la figura A se cam-

² Uhr, L. y Vossler, C., *A pattern-recognition program that generates, evaluates and adjusts its own operators*, Proc. of the Western Joint Computer Conference, 1961.

³ Evans, Thomas G., *A heuristic program to solve geometric-analogy problems*, Ph.D. Thesis, Dept. of Mathematics, M.I.T., 1963.

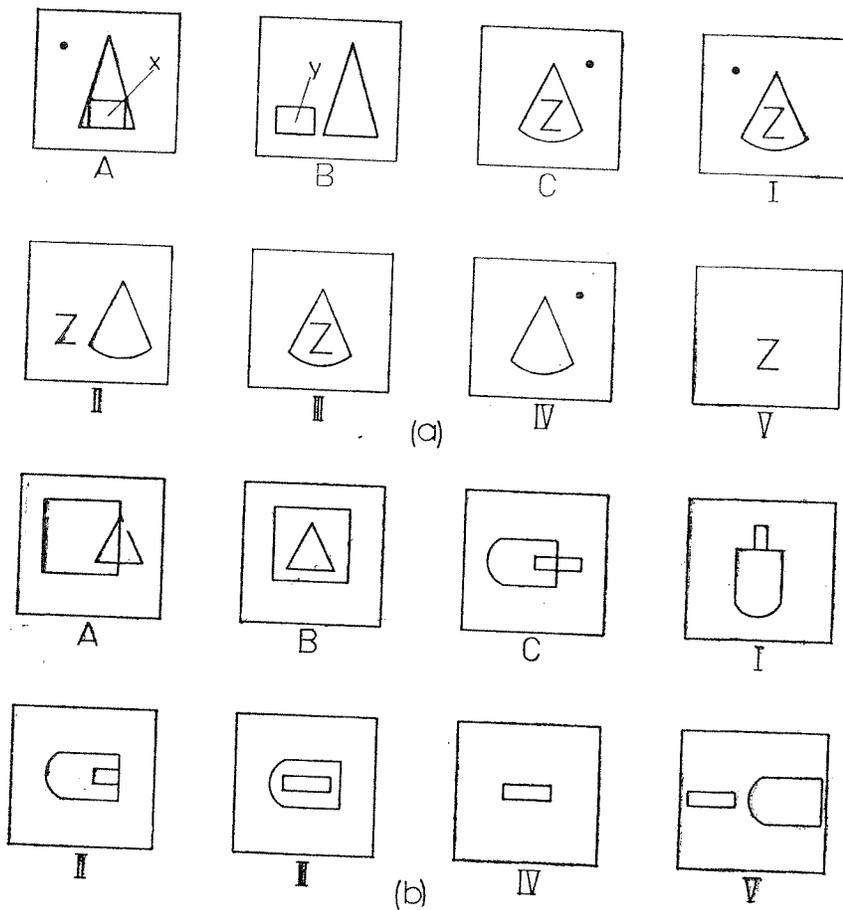


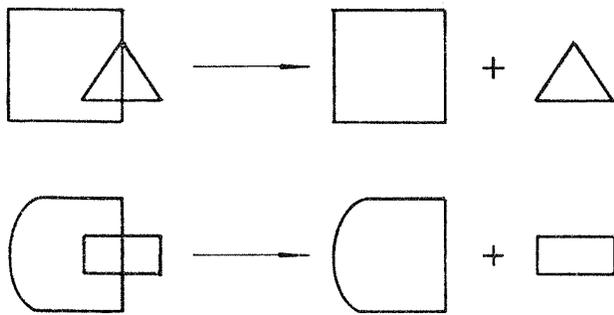
Figura 4.

bió en la B, aplicarla a la figura C y elegir la resultante de entre las figuras I a V”.

El propósito de Evans fue en realidad mucho más general que el del programa mencionado en sí, buscándose la posibilidad de implementar el concepto de “analogía”, tan importante en la inteligencia y simultáneamente difícil de precisar. Una parte de programa se ocupa de los factores relacionados específicamente con la visión, mientras que una segunda componente establece las analogías mediante las comparaciones necesarias. Esta última parte, independiente en cierto grado del tema (trabaja con descripciones, las cuales podrían adap-

tarse a otras áreas), juega el papel de una "elaboración de teorías" (las posibles transformaciones de A en B en este caso), tratando luego de "generalizarlas" para cubrir otros datos "experimentales" (las figuras C, I, V aquí), eligiendo luego la teoría más "fuerte" de acuerdo a un cierto criterio (la respuesta al test). En una palabra, una auténtica "formación de conceptos" mediante la experimentación.

Mediante la manipulación analítica de las descripciones de las figuras el programa inicial de reconocimiento calcula relaciones dentro de cada una, indicando por ejemplo que el rectángulo de A está adentro del triángulo (ver Fig. 4a), que el punto está "arriba y a la izquierda", etc. Para estas relaciones se basa en criterios muy simples, tales como trazar líneas desde la figura hasta el borde de la imagen y constatar las posibles intersecciones de estas líneas con otras subfiguras (un número impar de cruces indicará que está contenido en otra), o en dividir en cuadrantes a la imagen, decidiendo sobre la posición relativa en base a la comparación de las disposiciones de los centros. Las distintas subfiguras se distinguen de acuerdo a un criterio de búsqueda de objetos simples, aplicable cuando en otra figura aparecen separados; en Fig. 4.b. por ejemplo, decide las separaciones (en A y C):



en vista de que éstas aparecen disjuntas en B y III respectivamente, tratándolas en caso contrario como unidades indivisibles. Se trata de una técnica elemental de eliminación de posibles "ambigüedades" mediante el análisis de contexto.

Luego se establecen las "transformaciones" encontradas entre A y B, siendo éstas del tipo "el objeto x (en A de la Fig. 4.a.) se transforma en y (en B de la Fig. 4.a.) por medio de una variación nula de escala, ángulo de rotación y reflexiones", etc. Indicará además, en el ejemplo presentado, que el rectángulo ya no es más interior, y que el punto ha desaparecido. Y, final-

mente, se toman las posibles reglas de transformación de **C** en cada una de las restantes, descartando las que no coinciden con la original en cantidad de partes agregadas, retiradas, o coincidentes. Se "generaliza" ahora lo común de cada una de estas reglas con la que transforma A en B, obteniéndose entonces la más significativa. La figura a la cual corresponde esta regla es la entregada como respuesta al "test". En cuanto a los resultados, digamos que obtiene la respuesta correcta en alrededor de un 70 % de los casos (lo cual es bastante similar al rendimiento humano medio) con un tiempo aproximadamente igual al utilizado por las personas. Volveremos a referirnos a este programa en 3.4.

Hay gran cantidad de trabajos en estos temas, pero no tenemos aquí el espacio para dedicarnos a ellos, más aún, el reconocimiento de objetos en el espacio es un problema más apasionante, y tal que además de presentar gran parte de las dificultades del reconocimiento de caracteres, exhibe otras complicaciones propias.

3.2.2. El análisis de escenas y objetos tridimensionales.

El problema de la percepción del mundo exterior es investigado actualmente en el contexto de los "manipuladores mecánicos" o *robots*, autómatas con amplia capacidad de interacción con el medio ambiente, para los cuales es necesario contar con un adecuado sistema de RC. Esta es el área de inteligencia artificial donde más esfuerzos se encuentran concentrados, siendo preponderantes en estos momentos los grupos del Instituto de Tecnología de Massachusetts (M.I.T.), el Stanford Research Institute* y la Universidad de Stanford, todos en los Estados Unidos.

Siendo la presente sólo una introducción, nos limitaremos a señalar las características fundamentales de uno de los grupos (el del M.I.T.) como ejemplo representativo de los progresos alcanzados. El objetivo del grupo fue resumido como⁴ "extender el modo en que las computadoras pueden interactuar con el mundo real, específicamente el desarrollo de mejores equipos sensitivos y motores y de programas que los controlen". A tal efecto, los programas deberán analizar los objetivos propues-

* El robot del S.R.I., "Shakey", ha ganado gran popularidad, ocupándose de él inclusive la prensa no especializada.

⁴ *Status Report II: Research on Intelligent Automata*. Proyecto MAC, 1967.

tos y las escenas necesarias (mediante un equipo visual apropiado), formar un plan de acción para llevar a cabo dichas metas, controlar los órganos motores mediante el uso de la información visual y táctil y supervisar el cumplimiento de las tareas asignadas con una capacidad de interacción continua para superar los imprevistos. "Estos 'ojos y manos' deberán eventualmente realizar tareas razonablemente inteligentes pero ante todo, por cierto, ya es bastante difícil el conseguir que hagan cosas que son fáciles de realizar para un ser humano".

Para conseguir los mencionados fines, se cuenta en el laboratorio de Inteligencia Artificial del M.I.T. con un *robot* consistente de una computadora (el "cerebro"), una cámara de televisión (los "ojos") y un brazo y mano mecánicos, dividiéndose el trabajo entre el grupo de procesamiento de la información visual y el de control de brazo y mano. Como una discusión de los mecanismos de control de los "órganos motores" obligaría a entrar en detalles técnicos, nos referiremos principalmente a los resultados obtenidos por el primero de los grupos mencionados.

Los problemas que presenta el análisis de una escena son variados, destacándose además de los existentes en dos dimensiones (variaciones de tamaños y posiciones, etc.), los que aparecen debido a la perspectiva, a la obstrucción por otros objetos o partes de sí mismos y a los inconvenientes presentados por las variaciones de iluminación y las sombras. Se requieren facilidades para (A) "ubicarse espacialmente" (relacionar las posiciones del brazo, la cámara y la región de trabajo), (B) encontrar los objetos buscados, (C) reconocer y describir los objetos vistos y (D) construir un adecuado "modelo" interior para permitir los trabajos posteriores y debiendo para todo ello considerarse factores tan diversos como el enfoque de la cámara, la iluminación, la perspectiva, etc.

La solución a estos complicados problemas estará necesariamente en la interacción entre los distintos niveles perceptivos y los "conocimientos" que se poseen del mundo exterior, ya que es por ejemplo imposible describir completamente un objeto parcialmente oculto, a no ser que se recurra a la experiencia en cuanto a qué tipos de objetos podrían estar presentes. Claro está que aún así es inevitable cometer errores, como cuando creemos reconocer a una persona a quien observamos de perfil y luego notamos nuestra equivocación. Como la estructura del proceso es modular, nos referiremos ahora sólo a los programas referidos a la identificación de los objetos, dejando lo relacionado con la representación y el "modelo" interior para cuando hayamos discutido los problemas relacionados (3.3.3.).

Debe tenerse en cuenta, sin embargo, que no se trata de un procedimiento secuencial de aplicación de las diversas etapas de un proceso, sino de un constante intercambio de información entre los diversos programas en cuestión.

La imagen es entrada generalmente a través de una cámara que permite la observación local de características (mediante un "acceso" a los diversos puntos a voluntad), existiendo varios métodos de pre-proceso que permiten la formación de una descripción adecuada al manejo posterior. Un programa puede por ejemplo detectar los máximos y mínimos locales en la intensidad de la luz, permitiendo así identificar los bordes (recordando que casi nunca dos caras de un mismo objeto presentan exactamente la misma iluminación). Para determinar distancias se recurre ya sea a un análisis de las dos imágenes de una observación estereoscópica o el ajuste de la distancia focal, por ejemplo. Aclaremos que en general se trabaja con objetos de contornos lineales (prismas, cubos, etc.), lo cual simplifica gran parte de los cálculos, aunque los trabajos deberán ser por supuesto generalizados cuando se requieran análisis de escenas más reales que el de cuerpos geométricos sobre una mesa.

El primer análisis se realiza generalmente con un pequeño nivel de resolución, pero suficiente como para descubrir las características esenciales, debido al costo y tiempo que insumiría un análisis detallado. Pero una vez completado este estudio previo, la situación puede indicar la necesidad de investigar más a fondo una zona de la imagen, como en el caso de la Fig. 5, donde la línea marcada con puntos no es notada en una primera observación debido a la iluminación intensa desde encima, pero que es "propuesta" por uno de los programas y luego investigada por otro mediante un análisis más sensible.

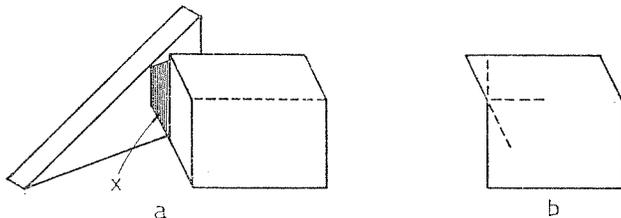


Figura 5.

En realidad el programa propone varias posibles continuaciones (ver 5.b.) intentando prolongar las líneas existentes o unir vértices separados, pero el "verificador" sólo detecta una de ellas. Los bordes son identificados además con la ayuda de

programas que intentan ajustar segmentos rectilíneos a las partes de la imagen propuestas como tales. Otro modo de determinar las caras de los objetos es mediante la detección de zonas homogéneamente iluminadas, separándolas entre sí. Complementando estos procesos, un programa ("ERASER") se encarga, basándose en diversos criterios heurísticos, de eliminar las sombras (ver marca "x" en Fig. 5.a.) calculando aproximadamente la dirección de la iluminación. La capacidad de "imaginar" líneas no observadas puede ayudar a identificar casos más difíciles, como el de la Fig. 6, donde se reconoce ante todo el cubo pequeño, luego se "retira" y continuando imaginariamente los bordes hasta que se corten, se deduce la forma del cubo que se encontraba oculto. Claro está que no existe ninguna posibilidad de saber si el resultado es correcto (el objeto inferior podría no ser exactamente un cubo), pero es la decisión que seguramente tomaríamos también nosotros.

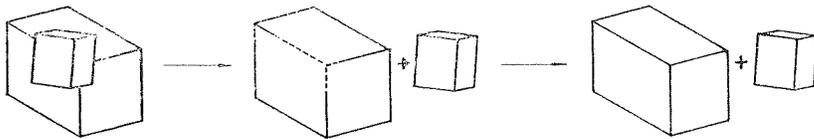


Figura 6.

Establecidos tentativamente los bordes se procura describir las regiones distinguidas, suponiéndolas de formas poligonales o no, ("Topologist" de T. Binford, "Polyseg" de A. Griffith, etc.) y se producen descripciones fijando las relaciones entre dichas áreas mediante indicaciones de posiciones relativas, etc. Posteriormente se comienza el "análisis de objetos" (la discriminación entre figuras y fondo) y el "reconocimiento de objetos", siendo esta última la parte más complicada del proceso. Comienza a actuar aquí el programa SEE del Dr. Guzmán, relatado en una interesantísima tesis publicada en 1968⁵.

El procedimiento de SEE consiste en basarse en diversos tipos de evidencias locales a fin de determinar si dos "caras" observadas pertenecen al mismo o a distintos objetos; los datos se obtienen principalmente mediante el análisis de los "vértices" (puntos de intersección de dos o más bordes de regiones),

⁵ Guzmán, Adolfo, *Computer recognition of three-dimensional objects in a visual scene*, M.I.T., 30 de diciembre de 1968. (Guzmán es latinoamericano, nació en México en 1943 y se graduó de ingeniero en comunicaciones y electrónica en su país antes de entrar en el proyecto M.A.C. del M.I.T.)

de acuerdo a los ángulos y otros detalles, con un proceso llevado a cabo por un programa de H. N. Mahabala. Los vértices se clasifican por dichos criterios (ver Fig. 7.), indicando una junta "Y", por ejemplo, la confluencia de tres caras del mismo cuerpo (Fig. 8 a.), las "T" un objeto que obstruye la visión de otro (Fig. 8 b.) y los del tipo "flecha" como indicación de dos lados del mismo objeto (I y II en Fig. 8 c.).

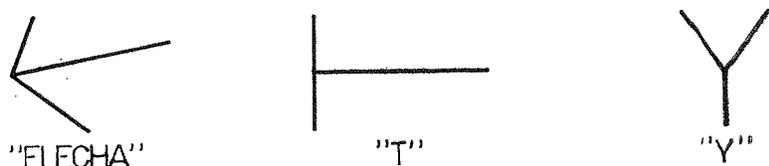


Figura 7. Algunos tipos de vértices.

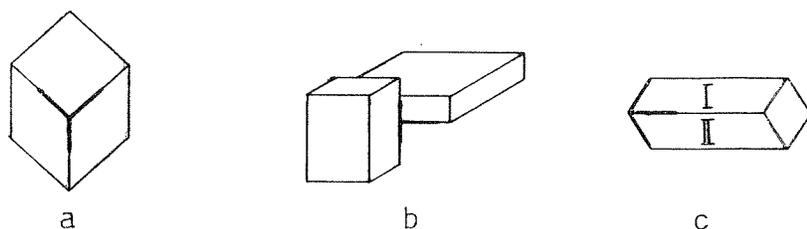
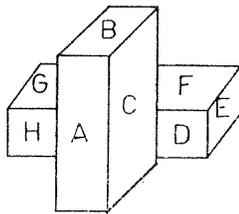


Figura 8.

A partir del análisis anterior, SEE traza "vínculos" entre regiones que podrían pertenecer al mismo cuerpo (como I y II en 8 c.), teniendo en cuenta tanto la información "a favor" como las indicaciones en contrario, que impiden ciertas vinculaciones. Se forman así "núcleos" de caras y vértices, decidiéndose por las elecciones más factibles. Es claro que un análisis local no será infalible, y por ello el recurso a un modelo interior de los posibles objetos puede ser necesario para dilucidar una situación ambigua. El programa resuelve figuras extremadamente complicadas, separando adecuadamente los objetos, comparando Guzmán sus resultados con los humanos en condiciones similares, especialmente si confrontados con objetos no vistos con anterioridad. Es realmente interesante también la discusión de Guzmán sobre las "ilusiones ópticas" y su enfoque por "SEE", pero no podemos extendernos acá con el tema.

El programa de Guzmán analiza figuras como la siguiente:

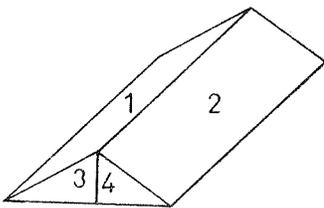


entregando un resultado del tipo:

“CUERPO 1 CONSISTE DE A, B, C.

CUERPO 2 CONSISTE DE D, E, F, G, H”.

Si bien su potencia le permite distinguir los 12 cuerpos de la figura 9, en algunos casos la información no le es suficiente para decidir, como en:



en que da como respuesta “el cuerpo (1, 2, 3, 4)” en lugar de “los dos prismas adyacentes (1, 3) y (2, 4)”. Estas fallas son en gran parte remediadas por una modificación del programa SEE, llevada a cabo por M. H. Rattner⁶, el cual no sólo establece “vínculos” sino que tiene en cuenta toda información en el sentido de *diferenciar* los cuerpos presentes; provee además la interacción con programas superiores de “formación de conceptos” y “aprendizaje” y la posibilidad de dar descripciones alternativas de una escena ambigua, permitiendo que un pro-

⁶ Martin H. Rattner, *Extending Guzmán's SEE Program*, Project M.A.C., Artificial Intelligence Laboratory Memo 204, julio de 1970.

grama, con conocimientos del tipo de objetos que es dable esperar, decida el reconocimiento definitivo.

Estos programas u otros similares permiten la acción de los "robots" experimentales, capaces de recibir instrucciones de formar figuras con bloques, etc., o inclusive de observar una cantidad de objetos, tomarlos uno a uno y cambiar sus posiciones. En Stanford, usando programas de reconocimiento óptico sensibles a los colores y mediante un lenguaje de comunicación

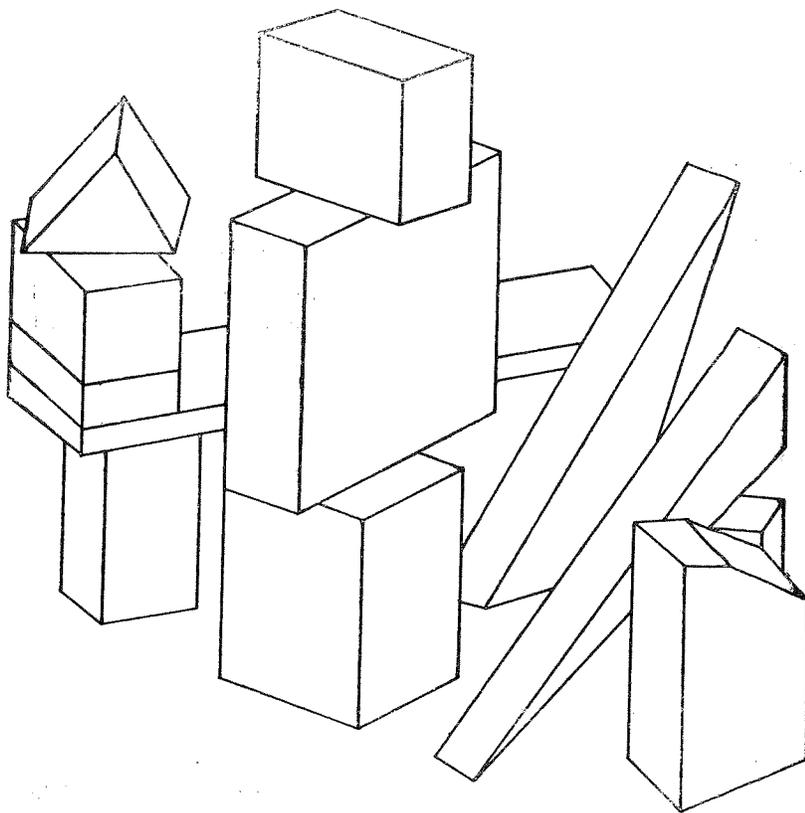


Figura 9.

adecuado, se pueden especificar tareas tales como "apilar todos los bloques azules pequeños a la derecha de la mesa y sobre el prisma rojo".

Como ya indicamos repetidamente, es fundamental la existencia de modelos de los objetos (aunque más no sea, para en-

tender qué es un "prisma" al recibir las instrucciones), pero estas estructuras de información son demasiado complicadas para ser compatibles con los sencillos árboles binarios del RC secuencial, por lo que recién volveremos al tema en (3.3.3.).

3.2.3. Conceptos y estructuras arborescentes.

El significado del término "concepto" es bastante difícil de precisar, y podemos asociarlo a la regla que nos permite determinar si un objeto dado pertenece a una cierta clase (su *concepto*). Podrá observarse la similitud con la idea de "configuración" tal como la hemos venido utilizando, lo cual indica la conveniencia de un estudio conjunto del RC y el aprendizaje de conceptos, ambos como el problema de elegir una respuesta ante un determinado estímulo, de acuerdo a cierta información o conocimientos.

Un clásico en la implementación de los procesos señalados lo constituye el "EPAM" de E. Feigenbaum¹, propuesto por su autor para la simulación del aprendizaje verbal, pero además modelo interesante para los trabajos en inteligencias artificiales (Feigenbaum y Simon, 1961). La noción fundamental del trabajo consiste en la disposición de una serie de tests o "discriminadores" como nodos en un árbol binario, dirigiendo éstos la búsqueda de la respuesta adecuada. Se simuló en particular el experimento consistente en la memorización de sílabas sin sentido de a pares (tal como "XUT-PIL", etc.), en el cual se presentan al sujeto dichos pares y luego se le solicita que recuerde, ante la presentación del primer miembro del par, cuál era la segunda sílaba.

A medida que se le presentan los pares al programa, éste va componiendo una "red de discriminación" tal que en cada nodo intermedio se encuentra explicitado un factor que permite diferenciar los caminos, mientras que en los nodos terminales se encuentran las sílabas o parte de ellas. Se almacena la forma completa de una sílaba si ésta es la segunda de un par (ya que se la necesitará para dar la respuesta correcta), pero en caso de ser la primera esto no es estrictamente necesario ya que si por ejemplo sólo se han presentado "XUT-PIL" y "DER-LAT", con almacenar "X" y "D" será suficiente para poder decidir la respuesta adecuada.

¹ Feigenbaum, E. A., *The simulation of verbal learning behavior*, Proc. of the Western Joint Comp. Conference, 1961 (reprod. en *Computers and Thought*, 1963).

El hecho de que sólo se almacene parte de la información es fundamental para comprender el comportamiento del programa, cuyo propósito esencial era, lo recordamos, simular al ser humano. Es así como debido a la imperfección de la discriminación se producen en ocasiones respuestas equivocadas, especialmente si las entradas respectivas son similares; continuando con el ejemplo anterior, si se hubiese aprendido también "XIT-GIL", una discriminación incompleta podría implicar la réplica "GIL" ante la sílaba "XUT" o "PIL" ante el estímulo "XIT" (fenómenos de "generalización" de estímulo y de respuesta, respectivamente). Además se puede observar en el programa un tipo de "olvido", no adjudicable a la destrucción de información sino, aunque resulte paradójico a primera vista, debido al aumento en la cantidad de conocimientos almacenados.

Los procesos de pérdida de capacidad de discriminación son de esperar en general en el RC o el aprendizaje de conceptos, ya que aunque en el caso de "EPAM" se podrían haber resuelto mediante la memorización de las sílabas enteras, en el caso de configuraciones más complejas deberemos limitarnos a un número reducido de características aparentemente distintivas, lo que puede acarrear confusión en caso de existir otro concepto que las comparta. Lo que sí puede mejorarse, en cambio, es la flexibilidad del modelo, de modo de permitir procesos más complejos que el estudiado, tales como la *comparación* de conceptos, etc. Pero para ello, es necesaria una estructura más rica que la mencionada en lo que respecta a la asociatividad y las *descripciones*.

Más recientemente, Minsky² se refirió a este punto indicando que "poseer una descripción (antes que sólo un test) para un concepto parece ser crucial. Las ventajas incluyen: 1) la posibilidad de obtener deducciones a partir del concepto, incluida su consistencia con el esquema de detección propuesto; 2) combinar varias descripciones de modos no triviales; 3) comparar y contrastar descripciones, como en el programa (de analogías) de Evans; 4) usar la descripción para generar, en lugar que meramente seleccionar, los próximos pasos de un programa de exploración". El programa de Winston (3.3.3.) servirá para demostrar la eficacia de estas ideas.

Para problemas más simples de RC secuencial, sin embargo, las estructuras simples de árbol son suficientemente adecuadas para el proceso de clasificación. Estos modelos, por otra parte, permiten aplicar técnicas avanzadas de exploración para su di-

² Project M.A.C., Progress Report VI, M.I.T., julio de 1969.

seño. El problema que se presenta es el de establecer una regla de decisión que implique la comprobación del menor número posible de características de la configuración; en la confección de un diagnóstico médico, por ejemplo, cada "test" se corresponde a un proceso al que se debe someter al paciente, lo cual puede ser molesto, costoso e inclusive perjudicial^{3,4}.

Se trata entonces de minimizar un "costo" (el de que la decisión tomada sea errónea), basándose en la aplicación de diversos tests, siendo la secuencia dependiente de los ya llevados a cabo. Este proceso es similar en ciertos sentidos al minimax de la teoría de juegos, aunque acá se trata de minimizar y promediar (de acuerdo con los valores de ciertas probabilidades), lo que podríamos denominar "minipromedio". Slagle y Lee⁵ describen el método "gamma", un equivalente del alfa-beta en el caso presente, y discuten las relaciones de estos métodos con la programación dinámica y los de "branch and bound". Veremos algo más de RC y teoría de la decisión en (3.4.2.).

3.3. Semántica y representación.

En las secciones anteriores hemos podido apreciar las dificultades y las soluciones propuestas para superarlas, que se presentan en los más variados campos de la inteligencia artificial. Es así como estudiamos máquinas que juegan, prueban teoremas, demuestran una restringida capacidad de "razonamiento" con el uso de dichas técnicas (3.1.5.), perciben y clasifican configuraciones y adquieren una capacidad elemental de discriminación conceptual.

En los últimos años, sin embargo, se produce una tendencia hacia el estudio de cuestiones en cierto modo más profundas y fundamentales. El hombre basa su inteligencia en factores distintos de una simple capacidad de memoria, alta velocidad y sofisticación de cálculo: es esencial su posibilidad de *entender* el *significado* de los diversos eventos. Minsky, líder en el área, acota: "creo que en nuestro estado actual sería más productivo entender cómo la gente entiende tan bien lo que se les dice que

³ Winston, Patrick, *A heuristic program that constructs decision trees*, A. I. Memo 173, marzo de 1969.

⁴ Banks, E. R., *Construction of decision trees*, A. I. Memo 189, febrero de 1970.

⁵ Slagle J. R. y Lee, Richard C. T., *Application of game tree searching techniques to sequential pattern recognition*, Comm. del A.C.M., 14, 2, febrero de 1971.

concentrarse exclusivamente en lo que descubren por sí mismos; a la larga puede que sea lo mismo, pero los dos enfoques llevan a distintas estrategias de investigación”.

Es por otra parte difícil definir “significado” y “entendimiento”; el primero puede, en principio, ser interpretado de un modo subjetivo, como la “idea”, que cada uno posee sobre un concepto; así “1810” será significativo para un argentino, pero probablemente sólo sea un año más para un africano. Con “idea”, a su vez, nos estamos refiriendo al *modelo* interior del mundo que poseemos. Este debe ser suficientemente completo, ya que “sólo en la medida en que estos ‘conocimientos’ (del medio ambiente) se representen adecuadamente podrá un programa utilizarlos apropiadamente en el sentido de aparentar *entender* la información, en lugar de utilizarla simplemente como parte de una búsqueda combinatoria exhaustiva” (Minsky).

La representación interna es fundamental para la resolución de problemas en general, siendo un requerimiento indispensable la adecuada modelación de los tipos de problemas y los diversos métodos de solución; es de esperar que los futuros programas con capacidad de aprendizaje basen ésta en una real modificación de dichos modelos, adecuando su comportamiento de un modo realmente cualitativo. Los conocimientos representados deberán referirse a las más diversas áreas, si bien es posible que una cantidad bastante limitada de elementos, *relacionados* adecuadamente, sea suficiente para la mayor parte del “sentido común” necesario.

El estudio de la percepción del lenguaje es un área que, además de tener un interés intrínseco, es ampliamente adecuada a la experimentación en la utilización de modelos. Por otra parte, se ha llegado a la conclusión de que gran parte de los procesos que lo involucran no pueden llevarse a cabo sin un adecuado “entendimiento”. El problema de las ambigüedades, por ejemplo, se presenta en los intentos de “traducción mecánica”; es grande la cantidad de palabras que tienen varios significados en una lengua particular, pero que al traducirse deberán determinarse. Esto fue lo que llevó a Bar-Hillel a su “demostración” de la imposibilidad de máquinas de traducir (1960), usando como ejemplo las sentencias en inglés “the pen is in the box” y “the box is in the pen”, y teniendo en cuenta que “pen” puede denotar tanto una lapicera como un corral; sólo “sabiendo” que una caja no puede estar dentro de una lapicera es que es posible traducir la primera sentencia, y con la segunda similarmente. Con la posibilidad de que la máquina posea esos conocimientos, sin embargo, podría, en principio, resolverse el problema.

En lenguajes naturales se encuentra, por otra parte, la mayoría de la información que sería interesante que la computadora tuviera, además de ser el medio más cómodo para comunicarse con ésta, en particular para especificar los problemas que nos agradaría que resolviera, y el único modo de interacción posible entre una máquina y una persona que desconozca su funcionamiento, pero que necesita efectuarle algún tipo de consulta. En fin, "la habilidad del hombre para usar símbolos y lenguajes es un factor primordial en su inteligencia, y si podemos aprender cómo hacer que una computadora entienda el lenguaje natural, habremos dado un gran paso hacia la creación de una inteligencia artificial" (Bobrow) ¹.

Finalmente, podemos considerar el "entendimiento" del lenguaje como aquel proceso consistente en reconocer la estructura de las relaciones entre las palabras del texto, identificar las palabras a las que hacen referencia los pronombres y el sentido de aquellas con múltiples interpretaciones posibles, produciendo en definitiva una codificación del significado del texto en cierta forma similar a la que toman los conocimientos ya almacenados, permitiéndose así las comparaciones y adiciones necesarias (Quillian) ². Lo último forma parte, a su vez, del "contexto global" necesario para la comprensión del lenguaje, mientras que el estudio de las frases adyacentes y pronombres es parte del análisis del contexto "local" implicado.

Es claro que no es de nuestro interés la "semántica" de los sistemas formales lógico-matemáticos consistentes en el estudio de los valores de verdad ("semántica pura"), sino como ya dijimos, el significado de las sentencias de acuerdo con lo empíricamente "razonable". Esta capacidad de entendimiento debe ser comprobada mediante la experimentación y deberá englobar la capacidad de la máquina de dialogar inteligentemente, recordando lo que se le diga, identificando inclusive los conocimientos pertinentes para cada tema y las implicancias lógicas de los hechos que se le presenten.

En cuanto a una formalización de estos conceptos, ésta puede estar dada por el agregado de "categorías sintácticas" más ricas y detalladas, de modo tal de diferenciar las propiedades semánticas (el significado), o en suponer una cierta cantidad de "escalas de valores", ubicando cada concepto de acuerdo con su índice en cada una de éstas (o sea, considerar los valores de los *atributos*). El estudio y determinación de las *relaciones* entre

¹ Bobrow, Daniel G., *Natural language input for a computer problem-solving system*, tesis de doctorado, M.I.T., septiembre de 1964.

² Quillian, M. Ross, *Semantic memory*, tesis de doctorado, Instituto de Tecnología de Carnegie, octubre de 1966.

los conceptos ayuda también a sus definiciones respectivas, como en el modelo que propusiera Quillian; es razonable considerar que el entendimiento de un concepto lo llevamos a cabo al ubicarlo en relación con los demás.

El problema de los bancos de datos y la "recuperación de información" ("information retrieval") es otro de los importantes por sí mismos, pero que necesitan de una adecuada estructura de la información. Además del interés de poder recabar datos mediante sentencias en nuestra propia lengua (con todos los problemas que acarrea su utilización), nos interesa el caso del "fact retrieval", en que la información que se nos hace llegar no consiste en el texto original ("document retrieval") sino en los *hechos* en sí; para ello es necesario que el sistema entienda nuestro pedido hasta el punto de identificar las referencias requeridas. El "fact retrieval" se puede caracterizar como aquel que implica un cierto tipo de proceso sobre el texto original, siendo las respuestas consistentes en textos *sobre* los archivos y no simplemente partes de éstos. El tema no está en general encuadrado en el de la inteligencia artificial, pero nos interesa en el caso particular de su utilización de lenguajes naturales y de los sistemas de información³ (en adelante, S.I.) diseñados específicamente para responder a preguntas.

Veremos a continuación algunos ejemplos de S.I. en relación a la percepción del lenguaje y a la solución de problemas, con estructuras de diversa complejidad, y un caso de utilización de éstas en el análisis de la percepción visual.

3.3.1. Lenguaje y comunicación.

Los procesos que forman parte de lo que denominamos como "inteligencia" son sin duda de una extrema complejidad. Se hace entonces necesario el manejo de grandes volúmenes de información interrelacionada, lo cual es, en gran parte, imposible si nos limitamos a la utilización de los "lenguajes de máquina" para la comunicación entre el hombre y la computadora. Es así como surge naturalmente la necesidad de utilizar lenguajes "de alto nivel" e inclusive de los lenguajes naturales en que expresamos habitualmente nuestras ideas.

El lenguaje puede ser considerado en sí como un código mediante el cual un individuo ("emisor") transmite a otro ("receptor") ciertas características de su propio modelo interior

³ "Question-Answering Systems", literalmente "Sistemas de Respuesta a Preguntas".

(ideas), proceso que culmina con la formación en el receptor de un cierto modelo que representa su entendimiento de la situación. El estudio de la relación entre los signos que constituyen el lenguaje (sintaxis) se ha realizado mediante la exposición de diversos modelos formales, algunos de los cuales vimos ya en el capítulo 1. Pero, además del análisis sintáctico, es necesario un estudio semántico, consistente en el examen de las posibles formas de representación del conocimiento adquirido y en el modo en que se realiza la “decodificación” del significado del mensaje recibido.

En lo que hace al procesamiento del lenguaje por sistemas artificiales, podemos en principio diferenciar entre el estudio lingüístico por sí mismo y el procesamiento de información lingüística, campo este último en que la “ciencia del lenguaje” juega un papel auxiliar y el objeto del estudio es otro (traducción automática, S.I., etcétera). Este procesamiento puede, por su parte, limitarse en lo semántico al *reconocimiento* del contenido (en algunos casos simples de traducción, por ejemplo) o incluir también una *evaluación* de éste (como para responder en los S.I.).

Los estudios lingüísticos son una ayuda fundamental en el procesamiento de información en lenguajes naturales. Como ejemplo, veremos brevemente su papel en la traducción mecánica, cuyo fin es el reconocimiento del contenido de una expresión en el lenguaje *fuentes* obteniendo su codificación en un lenguaje *objetos*, mediante una computadora. El recurso a la traducción “literal”, palabra por palabra, mediante una búsqueda en diccionario es generalmente inútil, debido no sólo a los giros idiomáticos y a las diferentes construcciones gramaticales de los diversos lenguajes sino también a la influencia que tiene la relación sintáctica entre las palabras sobre el significado de la sentencia completa respecto del de cada una.

Se hace entonces necesario un detallado análisis sintáctico de la entrada, para lo cual es, a su vez, imprescindible un adecuado conocimiento de la estructura del lenguaje *fuentes*. En algunos lenguajes el orden de las palabras es fundamental para la identificación del sujeto y objeto de la acción, mientras que en otros (ruso, por ejemplo) son importantes las terminaciones; las *rutinas sintácticas* son entonces una parte integral del proceso, permitiendo tanto el reconocimiento como la reordenación de las palabras obtenidas en el lenguaje *objetos* con el fin de dar el sentido apropiado al resultado, y el agregado o eliminación de términos auxiliares (preposiciones, etcétera) reconociendo las palabras con verdadero contenido semántico.

Pero si bien reconocemos la importancia de este análisis de la disposición de los signos y estructura de la sentencia, en ocasiones la existencia de “parsings” alternativos no permite la traducción correcta, aun en el caso en que no existe una real ambigüedad en el significado de las palabras, sino más bien en su tipo gramatical. Un caso interesante es, por ejemplo, el de la frase en inglés “They are flying planes”, que según se analice como en la Fig. 1 (a) o como en la Fig. 1 (b) significará respec-

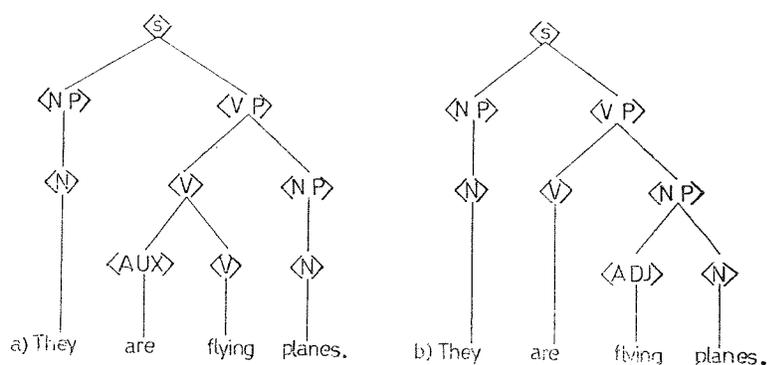


Figura 1. Basado en ejemplo de Chomsky, 1957; gramática de Harrison, op. citado.

tivamente en castellano. “Ellos están tripulando aviones” o “Son aviones voladores”. Se hace entonces necesaria una capacidad de entendimiento basada en un mayor análisis del contexto y en conocimientos generales: en “apoyé la escalera contra la pared, pero resbaló”, el modo en que “intuyo” inmediatamente el sujeto tácito de “resbaló” es mediante el conocimiento de que es poco probable que la pared resbale (comparar si no con “le presté el libro y no lo volví a ver”, en que no es claro a quién no encontré nuevamente).

Los estudios relacionados con el desarrollo de los S.I. son en gran parte paralelos (en cuanto al análisis del lenguaje y su significado) a los de traductores mecánicos, por lo que nos dedicaremos a los primeros que corresponden más al tema de inteligencia artificial. Para comparar los diversos esfuerzos podemos basarnos en variadas características: su capacidad de interacción con el usuario, la complejidad de datos y preguntas que admite como entrada, el tipo y calidad de información que puede entregar, su capacidad de deducción en lo que respecta a responder combinando distintos conocimientos, etcétera. Cabe aclarar que si bien nos referiremos en general a programas que utilizan el

inglés (debido a la procedencia de los principales ejemplos), es claro que los resultados se aplicarían prácticamente sin modificaciones ya sea en castellano o en cualquier otro idioma, siempre y cuando se provean las estructuras lingüísticas respectivas.

En 1960, Phillips¹ presentó un sistema capaz de contestar preguntas simples en un inglés restringido definido por una gramática libre del contexto y sin ambigüedades. Si bien no tenía capacidad para sacar deducciones por sí mismo, el programa, escrito en LISP, podía expandir sus conocimientos y poseía un modelo con los conocimientos necesarios para responder a las preguntas, consistente de una lista de cinco elementos (sujeto, verbo, objeto, lugar y tiempo), transformando a ese formato tanto los datos como las preguntas, estas últimas con un signo indicando el dato que se desea averiguar. El proceso de responder la pregunta se reduce a encontrar una lista coincidente y devolver la palabra ubicada en el lugar correspondiente al interrogante, de lo que se deduce que sólo daría la respuesta si se usaron las mismas palabras que al "instruirlo".

El programa Baseball², escrito en I.P.L., fue un intento de estudiar los S.I. en áreas particulares, lo cual disminuye el problema de la interpretación. En concreto se diseñó un sistema capaz de contestar preguntas en un inglés no demasiado restringido referidas a estadísticas de partidos de béisbol. Con rutinas especiales determinaba el significado de las posibles ambigüedades y giros idiomáticos e inclusive realizaba deducciones en cuanto a contar el número de partidos, etcétera. El modelo interior consistía en una "lista de especificaciones" consistente en pares formados por un atributo y su valor (por ejemplo "[mes, julio]"), y ordenados en una estructura jerárquica. Podía contestar a preguntas tales como "¿contra quién perdió el equipo X el 13 de julio?", indicando el nombre del equipo, o buscar los datos y contarlos dando sólo un número de respuesta como en "¿en cuántos sitios jugó Y en julio?". Formaba parte además de su estructura un diccionario con el significado y parte de la oración y una indicación sobre si indica pregunta ("qué", "cómo", etcétera) o si forma parte de algún modismo, para cada

¹ Phillips, *A question-answering routine*, Master thesis, Depart. de Matemáticas, M.I.T., 1960.

En un interesante artículo (Simmons, R. F., *Answering english questions by computer: A survey*, Comm. A.C.M. 8, 1, enero de 1965), se exponen los principios fundamentales y varios de los ejemplos aquí presentados sobre S.I.

² Green, B. F. et alii, *Baseball: An automatic question answerer*, Proc. West. Joint Comp. Conference, 1961 (Reproducido en *Computers and Thought*, 1963).

palabra. En este caso, el "significado" de una oración está dado por la lista en que ésta es transformada por el programa; así, por ejemplo, la pregunta "¿qué equipos ganaron 10 partidos en julio?" sería "entendido" luego del respectivo análisis como "(Equipo [ganador] = ?, Partidos [número] = 10, Mes = Julio)", y una rutina de búsqueda se encargará después de seleccionar los que cumplan. La respuesta se entregaba sin editar simplemente como una lista con los datos solicitados. Gran parte de su éxito debe adjudicarse a lo restringido del tema del que se ocupaba.

También se basa en datos estructurados como listas el programa "DEACON", de J. Craig y F. Thompson (1963/64), que realiza un análisis sintáctico de la entrada en inglés efectuando simultáneamente una cierta interpretación semántica, representada por el reconocimiento de las listas y sublistas referidas por la pregunta. Se generan varias alternativas, decidiéndose luego la correcta. Por ejemplo, si se han almacenado datos referidos a estadísticas de exportación, indexadas por país y envío, ante una pregunta tal como "¿cuáles embarques a Uruguay sobrepasan los 10 millones de pesos?", el sistema reconoce "embarques" como una lista principal y "Uruguay" como un modificador, y luego de generar (debido a la acción de las rutinas asociadas a la palabra "cuáles") todos los embarques que figuran, el "sobrepasan" indica una comparación de los valores obtenidos con "10 millones", eliminándose los que no cumplen.

El traducir a listas se puede generalizar a la transformación de las preguntas a equivalentes en "lenguajes de información" artificiales creados especialmente para la exploración de archivos. Desde el momento en que esta "traducción" requiere conocimientos del tema e inclusive una interacción con el usuario (en el caso de palabras desconocidas, estructura gramatical complicada de la pregunta, etcétera), pueden interesar sus resultados en lo que se refiere a la inteligencia artificial. Entre los sistemas más recientes se encuentra, por ejemplo, "CONVERSE" de C. Kellogg³, en el que se resuelven las ambigüedades sintácticas y semánticas mediante la interpretación de los "parsings" alternativos y el estudio de pares de palabras en las distintas sentencias.

Otros estudios interesantes fueron el de Darlinton para traducción de razonamientos en inglés a una notación de lógica formal que luego de analizar la proposición expresada lógica-

³ Kellogg, Charles H., *On line translation of natural language questions into artificial language queries*, Information Storage and Retrieval, agosto de 1968. En la mencionada publicación se encuentran a menudo otros ejemplos de sistemas como los aquí expuestos.

mente indica si representa un argumento válido, o el "Synthex" de Simmons (1962) para contestar preguntas sobre el contenido de una enciclopedia infantil previamente codificada y almacenada en una cinta magnética. Este último se basaba en una clasificación de las palabras en "funcionales" (que sólo poseen un "significado sintáctico") y aquellas con "contenido" semántico; ante una pregunta se buscaban entre las frases almacenadas, aquellas con mayor cantidad de palabras "con contenido" coincidentes, realizando luego un análisis sintáctico para elegir la respuesta adecuada. Si la pregunta era, por ejemplo, "¿qué comen las lombrices?" y en la enciclopedia se tiene "las aves comen lombrices" y "la hierba es comida por las lombrices", la primera respuesta es dejada de lado por no ser "lombrices" el sujeto sino el objeto de "comer", y la segunda, luego de verificarse que sirve como respuesta, es transformada en voz activa (como la pregunta) contestándose "las lombrices comen hierba". También simple era el modelo utilizado por Bennet para relaciones entre palabras; usaba 5 formatos con sentencias del tipo "avión es un transporte" y preguntas en los mismos con un "qué" u otro interrogante adelante. Se realizaron muchas otras experiencias con programas de estos tipos, pero acá nos interesan más bien aquellos en los cuales la información es manejada de acuerdo con su significado antes que con su forma.

Una interesante representación era la correspondiente a los datos del programa "Sad Sam" de Lindsay⁴. Aquí se trata de obtener, a partir de un texto, una estructura representando todo tipo de relación familiar que se pueda inferir del análisis de su contenido. En este área el modelo es bastante natural y está dado por los conocidos árboles genealógicos. Las sentencias analizadas corresponden a las expresables en el "Basic English", un subconjunto del inglés de alrededor de 1.700 palabras que contiene la mayor parte de las expresiones normalmente utilizadas, incluyendo oraciones bastante complicadas, las cuales son objeto de un análisis sintáctico del cual se obtiene en primer término un diagrama que represente su estructura.

Sobre este diagrama se produce el "análisis semántico", consistente en varios pasos. Se construye ante todo una lista con todos los sustantivos, y con ayuda del árbol sintáctico se determinan aquellos pares sujeto/objeto unidos por un verbo del tipo "ser", indicando su *equivalencia*, y agrupando los modificadores de sustantivos equivalentes. Y finalmente, sobre la base de la ocurrencia de las palabras "padre", "hermana", etcé-

⁴ Lindsay, R. K., *Inferential memory as the basis of machines which understand natural language*, en *Computers and Thought*, 1963.

tera, se identifican las relaciones existentes, produciéndose en definitiva una estructura asociativa similar a la de la Fig. 2.

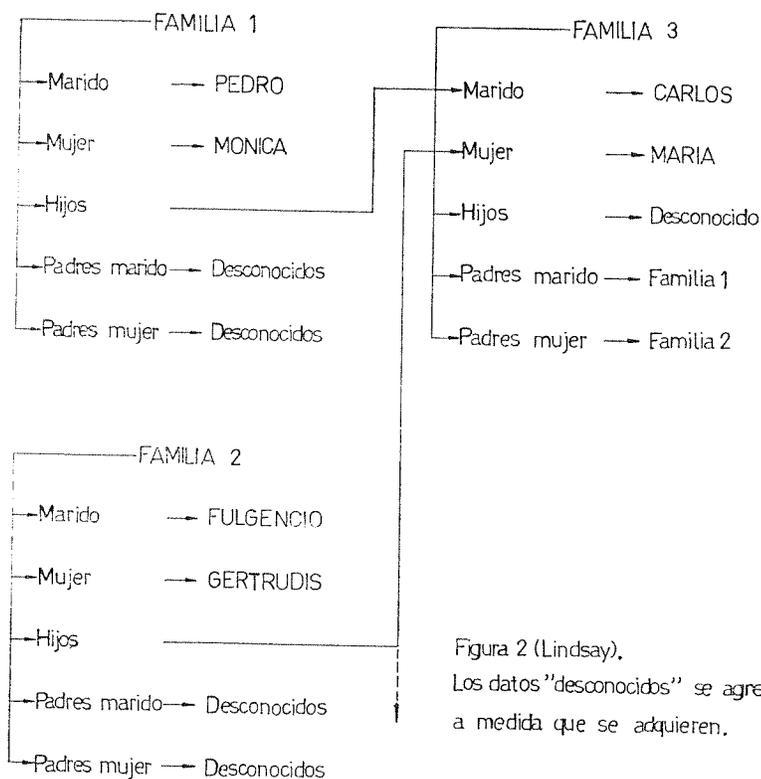


Figura 2 (Lindsay).
Los datos "desconocidos" se agregan a medida que se adquieren.

El modelo se sigue completando a medida que se leen nuevas sentencias, y a partir de su análisis se pueden obtener datos que, si bien implicados por el texto original, no aparecieron explícitamente en él. El mismo Lindsay propuso una modificación del esquema original con el fin de permitir el almacenamiento de "posibles" relaciones implicadas por los datos iniciales, como en la oración "Julio estaba con su tía Ana cuando lo llamó su madre", en la cual no es claro si Ana es la hermana o la cuñada de la madre de Julio. El lenguaje usado fue el I.P.L.-V, y el propósito del programa el estudio de los métodos de modelar interiormente la información semántica y sintáctica, lo cual se

cumplió en gran medida debido fundamentalmente a la estrechez del área atacada.

La "comprensión" del lenguaje en los casos anteriores ha sido en general satisfactoria, debido a que al restringir los dominios de los diversos programas, es también posible evitar en gran parte el afrontar la complicada cuestión de las ambigüedades. Como ya dijimos, éstas pueden ser de varios tipos, y en dos ejemplos más, que veremos a continuación, nos encontraremos con algunas de las más usuales, tales como las palabras con doble significado y la identificación de las referencias de los pronombres.

El programa "S.I.R." de B. Raphael⁵, estaba motivado por el estudio de la implementación de un sistema con capacidad de entender el lenguaje natural con la ayuda de un adecuado modelo formado a partir de la experiencia. El *significado* de las palabras es aproximado por una *descripción* consistente en explicitar sus propiedades y, en particular, su relación con otros conceptos.

Con el fin de simplificar el esfuerzo, sólo se tuvieron en cuenta algunas de las posibles relaciones entre conceptos en el programa considerado. Se trataba de: 1) inclusión de conjuntos, que sería una de las más naturales al humano; 2) relación entre las partes y el todo; 3) ponderaciones de la anterior; 4) pertenencia a conjuntos; 5) relaciones espaciales, y 6) relaciones de propiedad. De este modo es posible expresar conocimientos tales como "todo argentino es inteligente" (o sea, el *conjunto* de los argentinos está *incluido* en el de los inteligentes, con lo que corresponde a una relación del tipo 1) o "Patricia tiene un anillo", que indica la *propiedad o posesión* de un objeto por parte del sujeto (relación del tipo 6).

El análisis sintáctico del texto original no es en sí el objetivo del estudio, por lo que se optó por reemplazarlo por un sistema mucho más simple aunque limitado. Sería en principio posible, sin embargo, utilizar una rutina de análisis similar a las demás existentes como paso previo, y luego el resto del programa casi sin variaciones. El método utilizado consiste en comparar la sentencia contra ciertos "moldes" constituidos por constantes y variables. En caso de que las primeras coincidan, se toma una acción de acuerdo al valor que toman las variables. Por ejemplo, si uno de los formatos es "(X es un Y)" y la entrada ha sido "(Este perro es un boxer)", se adjudicarán los valores "este perro" a X y "boxer" a Y, con lo que se indica el significado como una relación de pertenencia a conjuntos: este perro particular es un

⁵ Raphael, Bertram, *SIR, A Computer Program for Semantic Information Retrieval*, tesis de doctorado, M.I.T., junio de 1964.

boxer, mientras que con "(Todo perro es un animal)" el "todo" correspondiente a la variable X indicará que en el modelo se deberá explicitar que el conjunto de los perros es una parte del de los animales.

El programa está escrito como una serie de listas, cada una de las cuales indica un formato, identifica las variables y su posible contenido y expresa la acción a tomar. Ante cada entrada se comparan los formatos hasta encontrar el adecuado; si se trata de una información nueva, agrega el conocimiento al modelo y responde con un "entendido"; si es una pregunta la contesta o, si no sabe la respuesta, pide más datos. En otros casos, debe resolver las ambigüedades que se presenten.

Se diseñó de tal modo de presentar un problema de ambigüedad que es representativo de los que suelen ocurrir normalmente, y que sirve para demostrar cómo una computadora con sentido común podría también resolverlos. Se trata del múltiple significado del verbo "tener", que se puede usar tanto para indicar que se es dueño de un objeto (relación de tipo 6) o que dicho objeto *forma parte* del sujeto (2). Esta distinción es tan obvia que la realizamos a menudo inconscientemente, pero en un programa debe aparecer explícitamente, ya que es conceptualmente distinto el que "Juan tenga una mano" indique que ésta es parte de su cuerpo, del caso en que haya adquirido una ajena; por el contrario, si se dice "Juan tiene un automóvil", nadie inferirá que haya uno en su persona.

El método que utiliza es relativamente complejo, ya que en caso de no haber ya tenido el mismo conocimiento, pero aclarado en cuanto a la ambigüedad presentada, realiza razonamientos tales como el siguiente: si se le dijo en una ocasión que "Juan es una persona", que Pedro también lo es y que Pedro es *propietario* de un automóvil, cuando se le diga "Juan tiene un automóvil" notará que Juan es en particular una persona y que, por otra parte, algunas personas tienen automóviles en propiedad, por lo que el "tiene" será correctamente entendido.

Por otra parte, las respuestas son entregadas en un formato preestablecido, y pueden incluir información deducida. Así, si se le dijo al programa en algún momento de su aprendizaje que "un hombre es una persona", al interrogársele "¿es una persona un hombre?" la máquina responderá "A veces", reflejando el hecho de que puede eventualmente existir otro sexo. En el siguiente pasaje de una conversación entre el operador y S.I.R., traducido del artículo de Raphael, podemos apreciar algunas características del programa ("E" indica lo expresado por el experimentador, "S", las respuestas de S.I.R.):

"E: Un muchacho es una persona. S: entiendo. E: Juan es

un muchacho. S: entiendo. E: Un dedo es parte de una mano. S: entiendo. E: ¿Cuántos dedos tiene Juan?. S: La sentencia es ambigua, pero supongo que 'tiene' significa 'tiene como parte'; no sé si 'dedo' es parte de Juan. E: Hay una mano en cada brazo. S: Entiendo. E: Hay dos brazos en una persona. S: Entiendo. E: ¿Cuántos dedos tiene Juan?. S: La sentencia es ambigua, pero supongo que 'tiene' significa 'tiene como parte', ¿cuántos dedos hay en cada mano? (notar que pide más información). E: Una mano tiene 5 dedos. S: Entiendo. E: ¿Cuántos dedos tiene Juan? S: Diez." Más adelante veremos cómo se almacena la información necesaria.

Una interesante experiencia fue también la del programa STUDENT de D. Bobrow⁶, que resolvía problemas de álgebra elemental expresados en inglés corriente. La entrada se analizaba de un modo bastante similar al anterior, utilizando el interpretador "METEOR", una adaptación de LISP y COMIT. El enfoque utilizado es el de las gramáticas transformacionales (1.3.2.), convirtiéndose las sentencias "núcleo" en ecuaciones a resolver. Si los datos no son suficientes, utiliza diversas heurísticas tales como identificar variables, encontrar giros idiomáticos y utilizar información general del tipo "un kilómetro equivale a mil metros", que posee en su memoria. Si esto no es suficiente, solicita ayuda. Los pronombres los reconoce buscando una frase similar a aquella en que aparecen, y la operación en general consiste en la identificación de "operadores" y considerando como variables a las frases que no los contengan.

Por ejemplo, para el problema: "La distancia de Nueva York a Los Angeles es de 3.000 millas. Si la velocidad promedio de un avión jet es 600 millas por hora, encontrar el tiempo que toma viajar de Nueva York a Los Angeles en jet", lo primero que hace es establecer las variables "tiempo que toma viajar . . . jet", "velocidad promedio de un avión jet" y "distancia de Nueva York a Los Angeles", estableciendo las ecuaciones respectivas. Como no es posible resolverlas así, "recuerda" además la relación "distancia = velocidad x tiempo", identifica "velocidad" (de esta relación) con "velocidad promedio de un avión jet", "tiempo" con la variable "tiempo que toma viajar de Nueva York a Los Angeles por jet" y "distancia" con "distancia de Nueva York a Los Angeles", respondiendo entonces correctamente.

Los resultados demuestran que es, efectivamente, más simple resolver un problema cuando se tienen "a mano" los conocimientos referidos a ese tema, con lo que un "reconocimiento"

⁶ Bobrow, Daniel G., *Natural language input for a computer problem solving system*, tesis de doctorado, M.I.T, setiembre de 1964.

del área se hace necesario al intentar atacar una cuestión. Aun dentro de Student, cuando se detectan palabras como "tiene la misma edad que", etcétera, se indica la presencia de un problema "de edades", utilizándose conocimientos aún más especializados, lo que permite interpretar correctamente problemas como: "María tiene el doble de la edad que Ana tenía cuando María era de la misma edad que Ana ahora. Si María tiene 24 años, ¿qué edad tiene Ana?", y resolverlo.

Un programa más actual, "CARPS"⁷, resuelve problemas más complicados que los de Student, los que se limitaban a ecuaciones simultáneas. Este programa, en cambio, resuelve problemas de cálculo elemental. Aunque su comportamiento recuerda en muchos sentidos a Student, en lo que hace a identificación de variables, el dominio del que se ocupa lo hace usar heurísticas especiales. Para resolver los problemas una vez entendido el texto, usa parte del "Mathlab" (3.1.3.).

El sistema "ELIZA" de J. Weizenbaum⁸ permite una eficiente comunicación en lenguaje natural e inclusive la resolución de problemas simples. Sobre la base de este sistema, se diseñó un programa que realiza un interrogatorio a pacientes con problemas de conducta, clasificando las palabras y frases y solicitando más detalles, cuando se hace necesario. Resulta interesante que efectuando una variante del "test de Turing" (introducción a este capítulo), luego de estar una hora charlando con el programa (vía una teletipo), el 62 % de los sujetos creyeron haber estado comunicándose con una *persona*, el 21 % no sabía y sólo el 17 % notó que su interlocutor era una *máquina*. Un factor que pesó en estas decisiones fue seguramente el hecho de que el programa imprimía sus respuestas de un modo vacilante y no uniforme, contrariamente a la rapidez y perfección de las computadoras en su uso normal.

En lo que respecta a los grupos de Inteligencia Artificial de la actualidad, se investiga⁹ la posibilidad de comunicación oral entre el hombre y la máquina, haciéndose necesario un análisis gramatical de la entrada para permitir una correcta decodificación de los fonogramas y detectar los límites de las palabras y frases. Para el control del robot, por otra parte, es necesario un adecuado lenguaje de comunicación.

El grupo del M.I.T., por ejemplo, cuenta con "PROGRAMMAR"¹⁰, un interpretador para la ejecución de los diversos tipos de análisis relacionados con el entendimiento del lenguaje

⁷ Charniak, E., *Loosely stated mathematical problems*, en Project M.A.C. Progress Report VI, julio de 1969.

⁸ Weizenbaum, Joseph, *Eliza*, y Stone, P., *An interactive inquirer*, Project M.A.C. Progress Report IV, julio de 1957.

natural. Es parte de un sistema más general, cuyo propósito es contestar preguntas, ejecutar las órdenes y aceptar información transformándola al formato necesario por otros programas, como "PLANNER", de C. Hewitt.

En lugar de usar "moldes" o de realizar un análisis sintáctico previo al "entendimiento" de la sentencia, su estructura permite que los diversos análisis se realicen simultáneamente, ayudando la información semántica y el contexto en el proceso de identificar las referencias de los pronombres y las ambigüedades. Los conocimientos generales se completan con una capacidad deductiva, permitiendo esto no sólo la desambiguación sino inclusive sirviendo como guía para el análisis sintáctico.

La idea fundamental y distintiva de este esfuerzo es el hecho de que tanto la gramática como los conocimientos necesarios se encuentran en forma de *programas*, antes que reglas que deban ser experimentadas sucesivamente como con el análisis sintáctico usual. Además utiliza una gramática algo distinta de los tradicionales enfoques de "estructura de frase" o "transformaciones". El uso de esta gramática permite resumir muchos de los procesos y partes del análisis "dependientes del contexto", verificando la coincidencia de tiempos de verbos, etcétera. Los programas tienen una forma del tipo siguiente (análisis de un predicado, por ejemplo): "buscar verbo, si es transitivo buscar objeto y analizarlo, caso contrario si es intransitivo devolver el resultado, si no lo es indicarlo", donde el análisis del objeto sería a su vez una rutina similar. Posee además rutinas para reconocer palabras que se derivan de otras (por prefijos, sufijos, etcétera) que están en su diccionario, con lo que éste se puede reducir en gran medida. Un ejemplo de frase analizada correctamente es: "tomar cualquier objeto verde, al menos tres de las piezas, y ya sea una caja o una esfera que sea mayor que cualquier ladrillo de los ubicados sobre la mesa".

En síntesis, como hemos podido apreciar a través de varios ejemplos, la única posibilidad viable para el entendimiento del lenguaje natural estaría aparentemente apoyada, de un modo necesario, sobre un adecuado modelo interior, único modo de resolver los variados problemas que se presentan.

3.3.2. Estructuras.

En la sección 3.1.5. vimos cómo es posible, en principio, la

⁹ McCarthy et alii, Stanford University, 1970.

¹⁰ Winograd, Terry, *Programmar: A language for writing grammars*, A.I. Memo 181, M.I.T., noviembre de 1969.

obtención de un comportamiento "razonable" utilizando una estructura relativamente simple como la del programa de F. Black, donde los conocimientos se almacenaban en forma de sentencias "directas" y "condicionales". Sin embargo, el programa resultaba así bastante ineficiente, y es difícil adjudicarle algún tipo de "entendimiento" a su conducta.

Motivado por su "Advice Taker", McCarthy ideó el sistema LISP, un formalismo y lenguaje de programación útil para el manejo de expresiones simbólicas, o *listas*. La memoria de la computadora se dispone de tal modo de representarse *árboles* de longitud arbitraria, los cuales pueden extenderse sin deber ser especificado su almacenamiento de antemano. A los *átomos*, las unidades de información, se les pueden asociar diversas propiedades o atributos. El sistema IPL-V (Newell et alii, 1959) también se basa en las estructuras de listas, y una "lista de descripciones" asociada a cada una, donde se pueden especificar los diversos atributos y sus valores.

En los últimos casos hemos visto ya comportamientos más complejos basados en estructuras más ricas que las simples memorias asociativas. Así, mientras que en el caso de Black se usaba una notación similar a la lógica formal tanto para la comunicación como para el manejo interno, y en otros casos *árboles* de discriminación para el reconocimiento de sílabas o de configuraciones, en los programas de 3.3.1. ya se tenían representaciones del significado de los términos, sea mediante relaciones especificadas como en "SIR", en ecuaciones algebraicas o en *árboles* de familia, indicando en cada caso los conocimientos distinguidos respecto de cada área particular.

El programa de Raphael utilizaba listas de propiedades asociadas a los diversos objetos, consistentes en pares de elementos, tales como (COLOR, VERDE), etcétera. Como se observó al analizarlo, era fundamental la posibilidad de expresar las relaciones de los 6 tipos señalados; esto se llevó a cabo mediante el establecimiento de "vínculos" de diversos tipos en la lista correspondiente al elemento. Así se tienen los que toman como valor del atributo el nombre del objeto señalado, aquellos casos en que se da a su vez una lista, o las propiedades que se califican además cuantitativamente. Si se le hubiese dado como información al programa que "una persona tiene dos manos" y que posee dedos (sin especificar cuántos), la estructura creada sería tal que en la lista correspondiente a "persona" se tenga: "(PARTES (NOMBRE, MANO, NUMERO, 2), (NOMBRE, DE-DO))" indicando así los conocimientos adquiridos.

Esta estructura corresponde a la de asociaciones con "rótulos" que indiquen su tipo. En la anterior, por ejemplo, la co-

respondencia de "persona" con "dedo" indicará que la segunda es parte de la primera. Y si a su vez en la lista descriptiva de "dedo" se incluye "uñas" como *parte* de "dedos", mediante una simple búsqueda siguiendo la dirección de las asociaciones se podrá contestar afirmativamente a la pregunta sobre si las uñas son parte de una persona. En esto consiste fundamentalmente el proceso deductivo capaz de llevar a cabo SIR, y un adecuado procedimiento de exploración para agilizar las búsquedas puede ser ventajosamente empleado con el fin de aumentar la eficiencia del sistema.

Una interesante representación del significado de los conceptos está dada por una propuesta de Quillian¹. Cada palabra queda definida por un "plano" de memoria, compuesto cada uno por una serie de "nodos" con términos que sirven para especificarla y que están a su vez conectados con los lugares donde son precisados. Cada uno de estos planos se corresponde con una de-

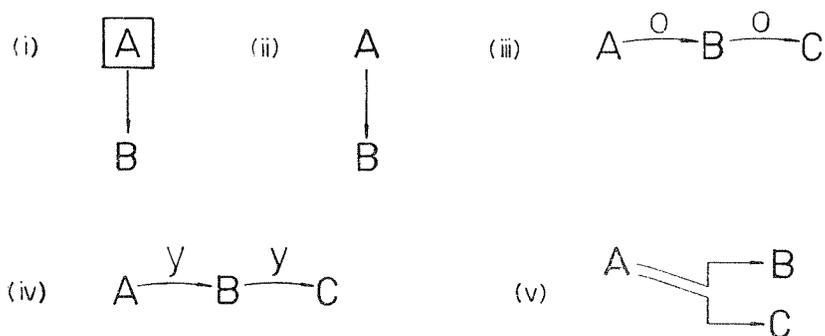


Figura 1.

finición de diccionario, siendo una especie de representación gráfica de la explicación. Así, mediante el uso de indicadores de diversos tipos (Fig. 1) se puede indicar (i) que "A" es un objeto del tipo "B", (ii) que "B" es un modificador de "A", (iii) que puede darse cualquiera de los casos A, B, C, (iv) que se deben dar todos y (v) que el sujeto B está relacionado con el objeto C de la manera especificada por A, siendo tácito en algunos casos B o C.

Por ejemplo, si definiésemos "planta" como "estructura viviente que no puede trasladarse, con hojas y que obtiene su alimento del aire, la tierra o el agua", la representación podría

¹ Quillian, M. Ross, *Semantic memory*, op. cit.

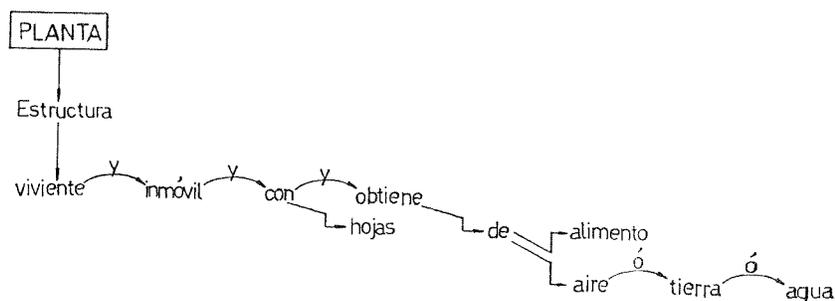


Figura 2.

ser esquematizada (simplificadamente) como en la Fig. 2. Esta es la forma de los “planos”, entendiéndose que cada término está “conectado” con el correspondiente a su definición. Por otra parte, se pueden indicar distintos significados de una palabra con indicadores del tipo (iii).

Si bien la estructura es mucho más complicada en cuanto a los detalles, lo expuesto es suficiente para ilustrar sus principios. Una interesante aplicación del modelo está en el modo como se comparan dos conceptos. Todo el proceso puede consistir en partir de ambos planos de definición simultáneamente, recorriendo las flechas en las direcciones indicadas, hasta que se encuentren. Una vez unidos, se habrá formado un camino que puede servir de “comparación” si se leen los nodos que lo forman. Así, por ejemplo, si se considera la definición indicada de “planta” y una para “hojas” en la cual se indica que pueden ser verdes, al comparar “planta” y “verde” se establecería una “conexión” entre ambos, a través del término “hojas”, explicando que “las plantas son seres vivos con hojas que son, a veces, verdes”. Para comprobar su efectividad, Quillian programó el modelo con algunas palabras y un programa de comparación, usando el lenguaje IPL, obteniendo interesantes resultados. Esta estructura podría ser apropiada también para los procesos de entendimiento del lenguaje, particularmente para la resolución de ambigüedades.

Los modelos existentes son muchos, adaptándose a las diversas situaciones y temas, quedando aún un gran camino por delante en cuanto a la determinación de los más adecuados; entre éstos deberán encontrarse esquemas de representación tanto para los conocimientos de las relaciones entre objetos como para los métodos de resolución de diversos problemas. La posibilidad de basarse en procesos dinámicos o programas (como

se hizo con 'PROGRAMMAR', por ejemplo) representa una alternativa interesante.

También se puede pensar en estructuras diferenciadas de acuerdo con la procedencia del conocimiento. Así, lo adquirido visualmente se podría quizá almacenar de un modo más directo, con lo que no se perderían detalles debido a la capacidad de entendimiento en el momento de la percepción. Su eficiencia puede inferirse de la consideración de nuestro propio comportamiento, cuando "imaginamos" una situación que nos ocurrió intentando descubrir algún detalle que en su momento pasó inadvertido. Para el proceso posterior de estos conocimientos se necesitarían analizadores adecuados; ya hace varios años se comenzaron estudios al respecto, y un resultado interesante fue el del grupo del NBS², en Estados Unidos, que diseñó programas capaces de estudiar propiedades de una figura. Su interés radica en gran medida en que permitía en cierto modo la comparación de conocimientos de diverso origen. En efecto, admitiendo como entrada una sentencia en inglés (restringido) y una imagen, se generaban, luego de un análisis de ambas por separado, dos predicados de lógica formal, los cuales eran luego comparados verificándose así si la sentencia se correspondía con la imagen.

El criterio en el cual basar la elección del modo de almacenamiento adecuado puede ser el de velocidad de acceso (como en las aplicaciones en las cuales los datos son usados por otros programas) o de economía en el uso de la memoria (aplicaciones comerciales con grandes volúmenes de datos). También la posibilidad de rápida codificación de la información puede ser un factor importante. Esta última condición, la de la facilidad de transformación, puede ser verificada mediante una representación mediante una estructura *natural* al tema en cuestión, sea ésta una lista con estadísticas o un árbol con datos genealógicos.

Pero para un sistema más general, capaz de demostrar una conducta inteligente, se requiere la posibilidad de representar todo tipo de concepto, ya sean objetos o relaciones entre ellos, e inclusive el significado de estas relaciones. El enfoque de Quillian, definiendo un concepto de modo similar al de los diccionarios, es un paso en tal sentido. Pero las palabras no son lo único que se puede representar efectivamente por una "red" del tipo mencionado. La descripción de escenas tridimensionales por medio de tales redes, en particular, fue el objeto del que quizá sea uno de los programas más importantes, en cuanto a métodos y principios básicos, de los que se hayan llevado a cabo en el área de la inteligencia artificial.

² Ver, por ejemplo, Cohen, D., *Picture processing in a picture language machine*, National Bureau of Standards, Washington, 1962.

3.3.3. Un modelo de aplicación.

El trabajo al que nos referiremos se encuadra dentro del sistema de percepción y manipulación del M.I.T. (3.2.2.), y constituye una contribución fundamental en el área del aprendizaje de conceptos. Siendo parte de un programa de investigación actualmente en curso, surgen constantes modificaciones y mejoras, por lo que no entraremos en la consideración de detalles técnicos. Sus principios, sin embargo, son suficientemente generales como para inducir a su consideración como un auténtico modelo para los futuros autómatas inteligentes. Seguiremos el tratamiento de su autor, el doctor P. Winston¹.

Entre los antecedentes de la serie de programas que componen el trabajo, merecen destacarse el de Evans de analogías (3.2.1.) en cuanto al manejo de descripciones, y en menor medida a EPAM de Feigenbaum y Simon (3.2.3.) en lo que se refiere al aprendizaje y discriminación. Sin embargo, se diferencia de ambos tanto en cuanto al enfoque como a los métodos utilizados.

El propósito del sistema es, según Winston, la clarificación de los interrogantes referidos al reconocimiento de ejemplos de conceptos y aprendizaje de las técnicas implicadas, en el hombre y la máquina, y la determinación de la importancia de un adiestramiento cuidadoso. El dominio utilizado para la experiencia es el del análisis de escenas en las que aparecen cuerpos geométricos regulares y simples, tales como prismas, ladrillos, etcétera. Son esenciales en su desempeño las descripciones abstractas, construidas a partir de las imágenes, las cuales serán más o menos similares entre sí de acuerdo con la relación existente entre las escenas originales. Gracias a esto último, las comparaciones se pueden llevar a cabo por medio de programas en gran medida independientes del área elegida, por lo que serían en principio aplicables los resultados a todo campo en el que el "entendimiento" puede ser formalizado por medio de adecuadas descripciones.

Entre los principios fundamentales que influyen sobre el aprendizaje llevado a cabo por el programa se encuentran la existencia de una adecuada secuencia de entrenamiento y la noción del "error por poco". En efecto, el orden en que son presentados los conocimientos resulta condicionante de la formación de ciertos "prejuicios", por lo que éstos se deben estructurar adecuadamente; un contraejemplo para un concepto, por

¹ Winston, Patrick H., *Learning structural descriptions from examples*, Project MAC TR-76, M.I.T., setiembre de 1970.

otra parte, si es tal que difiere en pequeños detalles, indicará qué es lo significativo en lo que se desea aprender.

Uno de los programas, operando sobre la salida del "SEE" de Guzmán (3.2.2.), describe la escena presentada, en términos de las relaciones "ARRIBA-DE", "A LA DERECHA DE", etcétera. Para esto se basa en varios tipos de evidencia local (vértices, intersección de líneas correspondientes a figuras distintas, etcétera) y otras algo más globales (objetos que interfieren la visual, etcétera). Como ejemplo digamos que describe adecuadamente las relaciones entre los objetos de la Fig. 9 de (3.2.2.).

También es posible la comparación de dos escenas, mediante un programa encargado de la contrastación de las descripciones, el cual *describe* a su vez las diferencias y similitudes encontradas. De este modo se puede implementar el aprendizaje de los conceptos a partir de una serie de ejemplos: se van formando las descripciones a medida que éstos se presentan, y se *abstraen* las propiedades en común, indicando la aparente necesidad de ciertas características (las que aparecen repetidamente) y lo accidental de otras (aquellas que forman parte de algunos ejemplos pero no de otros). Si se presenta un contraejemplo, el programa determinará el detalle que a su juicio influye en que lo pueda considerarse un caso del mismo concepto.

Programas de identificación permiten hallar en una imagen el cuerpo buscado, o comparar una figura con modelos almacenados con el fin de reconocerla. Si se *busca* un objeto particular, es más probable "hallarlo" que en el caso de listar directamente los observados, ya que en caso de ambigüedad (debido a una oclusión parcial, por ejemplo), se decidirá en lugar de otro que se sabe no se encuentra presente.

El hecho de representar las comparaciones entre descripciones de un modo adecuado (a su vez como *descripción*), se puede interpretar como un segundo nivel de abstracción, y permite, por ejemplo, resolver problemas como los de ANALOGY (3.2.1.). En efecto, si notamos por c ($A : B$) la comparación de A con B , y c ($C : n$) las de C con las figuras I a V , alcanza ahora comparar a su vez c ($A : B$) con cada una de las c ($C : n$), dando como respuesta la que coincida en mayor grado. Y es interesante especular sobre su aplicación en el GPS (3.1.5.), en cuanto a la elección de "operadores" de acuerdo a la "diferencia" que se desea eliminar: si cada operador es definido a su vez por la diferencia que causa, una "comparación" de las descripciones respectivas sería a su vez útil para determinar el adecuado.

La estructura de estas descripciones puede ilustrarse como una red o grafo con rótulos donde cada nodo corresponde a un concepto, mientras que las flechas indican el tipo de relación.

Así es posible describir la escena de la Fig. 1 indicando que el cuerpo "ABC" (lo denotamos por los nombres de las caras) es un ladrillo que se encuentra parado y a la derecha del "DEF", del modo mostrado por la Fig. 2.

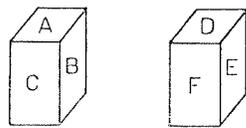


Figura 1.

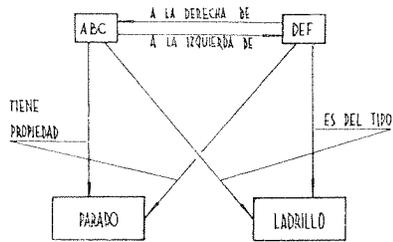


Figura 2.

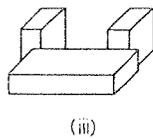
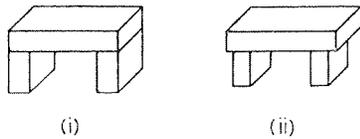


Figura 3.

Es interesante el hecho de que estas "relaciones", siendo a su vez *conceptos* pueden ser descriptos de una manera similar, apareciendo entonces como nodos, como en el caso en que se representaría que "a la derecha" es lo opuesto de "a la izquierda", etcétera. De este modo, además de obtenerse una cierta simetría, se permite la existencia de unas pocas relaciones básicas en los programas en sí, encontrándose las demás definidas junto al resto de los conocimientos.

Para finalizar, veamos con un ejemplo muy simple el proceso del aprendizaje del concepto de "arco" a través de una secuencia de cinco figuras. Con la primera (ver Fig. 3) le informamos que es un ejemplo de arco. Asimismo, con la siguiente, con lo cual la computadora deduce que la coincidencia de los bordes entre los pilares y la parte superior no es esencial; en cambio sí lo es el que los dos inferiores lo sostengan, como lo indica (iii), que no es un arco. Del siguiente contraejemplo aprende que los pilares no se deben tocar entre sí, ya que es la única diferencia con la figura inicial, y por lo tanto única causa posible de que (iv) no sea un arco. Si hubiesen existido más diferencias se formarían descripciones alternativas hasta tanto quede en claro (con ejemplos posteriores) cuál o cuáles son determinantes de que no lo sea. Una vez aprendida la necesidad del hueco, sólo resta hacer notar que la parte superior puede ser cualquier objeto (no necesariamente un ladrillo), para lo cual es suficiente la presentación del ejemplo (v). Los procesos son en general más complejos que el descripto, pero ésta es la idea subyacente.

3.4. Modelos generales y enfoques formales.

En lo precedente, hemos resumido algunos de los principales estudios heurísticos relacionados con la inteligencia artificial. En casi todos los casos nos encontramos con sistemas diseñados para cumplir una función específica, utilizando a tal fin diversos procedimientos, creados "ad-hoc" por el programador. Sin embargo, fue común también a la mayoría el propósito de utilizar esos métodos en casos más generales.

En ocasiones los autores han señalado explícitamente, o al menos indicado los lineamientos globales para una implementación de sus ideas en conjunción con las de otros investigadores, con el fin de ampliar su campo de aplicación. Un ejemplo está dado por B. Raphael, quien propuso una axiomatización de su "S.I.R.", la que permitiría extender indefinidamente el tipo de relaciones capaz de manejar, mediante su definición sobre la

base de unas pocas primitivas; las ventajas de este modelo ("SIR-I") incluyen también la posibilidad de utilización de métodos de demostración de teoremas en el curso del proceso deductivo. En particular se concretó recientemente un trabajo en ese sentido, utilizando un programa basado en el "principio de resolución" de J. A. Robinson (3.1.3.) en el Stanford Research Institute (B. Raphael y C. Green, 1968). Raphael propone la resolución del problema del "Advice Taker" (3.1.5.) mediante la utilización de su esquema, en un proceso de abstracción a partir de los datos almacenados: "esencialmente, la máquina 'inteligente' deberá poder abstraer lo necesario de la información en su modelo, 'notar' la necesidad de realizar acciones adicionales y crear las instrucciones necesarias por sí misma. El diseño de una tal 'inteligencia artificial' aguarda el desarrollo de la formación automática de conceptos y los sistemas de inferencia inductiva así como a las generalizaciones de SIR descriptas".

Algo más ideal es el esquema propuesto por Evans sobre la base de su programa de analogías, involucrando el uso de gramáticas con estructura de frase y el "GPS". Se trataría en principio de formalizar, en un lenguaje adecuado, las descripciones de los diversos temas, utilizando una gramática generativa del tipo de las que hemos ya estudiado. El "GPS" actuaría del modo usual, sólo que ahora, en el momento de aplicar una "transformación" (recordar 3.1.5.), en lugar de basarse en tablas limitadas y fijas, se trataría de investigar la *estructura* de la submeta en cuestión mediante un "análisis sintáctico" en conjunción con las rutinas de comparación de "ANALOGY", las cuales, además, introducirían un elemento de aprendizaje en el sistema, al abstraer y generalizar la forma de las diversas reglas de la gramática.

Como podemos apreciar, la generalización está prácticamente siempre unida a un cierto grado de "formalización" de los conceptos involucrados. Esto permite un estudio más profundo de los métodos y de sus limitaciones, como así también la aplicación de procedimientos lógico-matemáticos adecuados. A su vez, esto implica, en gran cantidad de casos, la formulación de *algoritmos* en reemplazo de previos intentos heurísticos, como así también la creación de varios procedimientos estadísticos, todo lo cual deja en parte de incluirse en los imprecisos dominios de la "inteligencia artificial". La formalización de los procesos generales de aprendizaje reviste un carácter sumamente teórico, siendo difícil asignarles un papel importante en el desarrollo futuro, al menos a corto plazo.

La emergencia de comportamiento inteligente a partir de modelos biológicos es otro de los enfoques que difícilmente lle-

ven a resultados satisfactorios en un futuro cercano, aunque su valor para el entendimiento de ciertos principios fundamentales y su importancia histórica los hacen merecedores de un análisis más detenido. La aseveración de McCulloch y Pitts (1943) de que sería posible representar y simular con una red de "neuronas" formales todo proceso biológico unívocamente definido, es uno de los supuestos básicos de la biónica, la ciencia del aprendizaje de los principios de la naturaleza y su aplicación al diseño de dispositivos electrónicos.

Los modelos basados en estas "redes nerviosas" consisten en conmutadores ("células" o "neuronas") dispuestos de diversos modos, los cuales se organizan de acuerdo con la experiencia. En el modelo de Hebb (1949), las células se encuentran asociadas inicialmente al azar, y si las decisiones tomadas son correctas, las conexiones que las permitieron son ponderadas en mayor grado en las oportunidades siguientes ("aprendizaje sináptico"); el "Perceptrón" de Rosenblatt es una variación de este esquema. La cantidad de modelos es demasiado extensa para que nos ocupemos de ellos aquí, y un ejemplo alcanzará para formarse una idea de su esencia.

Como ya hicimos notar, la aplicación de estos métodos a la creación de máquinas inteligentes está, por el momento, sumamente restringida, lo cual se debe fundamentalmente a un problema de complejidad: en el cerebro humano existen alrededor de diez mil millones de neuronas, cada una de las cuales posee varios estados y grados posibles de respuesta. Sin embargo, con la eventual creación de dispositivos técnicos más completos y ricos en interrelaciones, es difícil predecir si no llegaremos a obtener un "sistema nervioso" artificial. Según W. Wieser¹, teniendo en cuenta solamente la diferencia en el modo de comportamiento, aun cualitativo, que se obtiene al pasar de sistemas con unos pocos elementos conmutadores a otros con el doble de componentes (y sus correspondientes interconexiones), podemos tener una idea de la posible razón por la cual puede el cerebro comportarse mucho más complejamente que una computadora actual: aquí la relación no es de "dos a uno", sino que se mide en elevadas potencias de diez (entre 10^6 y 10^7 según Wieser, aunque esta estimación es quizá todavía demasiado conservadora, teniendo en cuenta los resultados obtenidos últimamente sobre la variedad de grados de respuesta posibles en las neuronas); en definitiva, tal como dice Wieser: "un aumen-

¹ Wieser, Wolfgang, *Organismos, estructuras, máquinas*, traducción de Editorial Universitaria de Buenos Aires, 3ª edición, diciembre de 1970.

to en la complejidad puede traer aparejado un cambio cualitativo del sistema de magnitudes insospechadas².

La variedad de comportamientos posibles con sistemas aun de pocos elementos, puede ser verificada con simples dispositivos, como ya en 1953 lo demostró G. Walter² con su "machina specular" compuesta de dos elementos que pueden interactuar de siete modos distintos. Sobre la base de unos pocos circuitos binarios como los descritos en la Capítulo 1, podemos ver un simple pero significativo ejemplo de estos tipos de sistemas. En

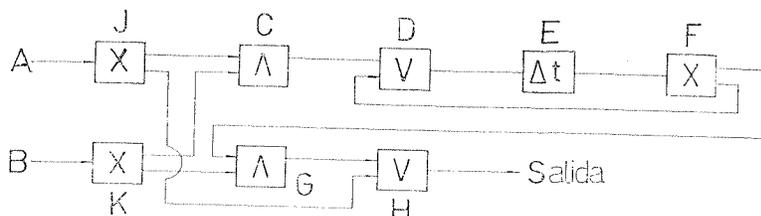


Figura 1.

la Fig. 1 vemos un circuito formado por tres duplicadores (J, F, K) un retardador (E), dos circuitos de disyunción inclusiva (D, H) y dos de conjunción (C, G).

Dejamos al lector la comprobación de que, suponiendo la salida del retardador en el instante inicial igual a cero, la salida será "cero" si así lo fueron las entradas en A, independientemente de la presencia, o no, de un "1" en B; y ante un "1" en A se produce igual salida, no influyendo el valor de B. Este comportamiento se puede describir por las ecuaciones que se obtendrían al determinar H (la salida) y por D, que, como lo podrá verificar el lector a quien le interese, resultan:

$$H_t = A_t + D_{t-1} \cdot B_t$$

$$D_t = A_t \cdot B_t + D_{t-1}$$

Lo interesante de este sistema radica en lo que sucede cuando se dan conjuntamente un "1" en A y otro en B. En este caso, se emitirá un "1", pero con la consecuencia adicional de que el retardador emitirá en adelante "1" en forma constante, lo que dará como resultado que cada vez que se presente ahora

² Walter. W. Grey, *The living brain*, N. York, Norton & Comp., 1953, y recientemente en castellano en *Sobre modelos cibernéticos de comportamiento y aprendizaje*, Cibernética Hoy, Editorial Tiempo Nuevo, Venezuela, marzo de 1971.

una señal positiva en B, ésta será la respuesta del sistema, cualquiera sea el valor de A.

Si ahora interpretamos la entrada A como la vista (o el olfato) de la comida, B como el sonido de la campana, la salida de H como el aumento de la salivación y todo nuestro sistema como esquematizando parte del sistema nervioso de un perro, tendremos el clásico experimento de Pavlov, con la variante (debido a las simplificaciones efectuadas) de que el reflejo es adquirido con una sola experiencia. Este pequeño ejemplo debería ser suficiente para ilustrar la posibilidad de modelar procesos básicos con sistemas extremadamente simples. Más adelante, veremos (3.4.1.) cómo se pueden llegar a representar algunos comportamientos adaptativos mucho más complicados gracias al aumento de la complejidad del sistema.

Subyacente a todos los mecanismos “adaptativos” se encuentra un proceso que tiene por finalidad la identificación de los factores que estuvieron presentes en cada “éxito” o “fracaso” de su actuación, y la determinación adecuada del peso relativo de cada uno sobre las decisiones finales. Estos procesos de “refuerzo” pueden actuar sobre un sistema a modo de “instructor” que, de acuerdo con el grado de éxito obtenido, lo “premia” o “castiga”, con lo que se acentúan algunos comportamientos y se dejan otros de lado. El “instructor” es, claro está, un ente ideal, ya que puede ser tanto un humano como la propia máquina, la cual puede valorar su propia actuación sobre la base de los resultados obtenidos con su accionar.

Una interesante interpretación es la que toma a la naturaleza como “instructor”, no ya del individuo sino de su raza. Esta influye sobre el aprendizaje “filogenético” (de la raza) en contraposición con el “ontogenético” del individuo, “enseñanza” que podríamos interpretar como el proceso de selección natural. Sobre la base de este enfoque hay una interesante proposición en lo que respecta a la inteligencia artificial: la “programación evolutiva”³, a la cual nos referiremos solamente como una curiosidad, debido al poco desarrollo del tema.

La programación evolutiva ofrece un método para obtener sistemas con una gran capacidad de adaptación al medio ambiente. La idea esencial del método consiste en simular, con la ayuda de modernas computadoras digitales, un proceso de evolución, similar en su concepto, al ocurrido en nuestro planeta durante millones de años. Como sabemos, de acuerdo con la teoría evolucionista, la influencia de factores externos —principalmente

³ Fogel, L. J., Owens, A. J., Walsh, M. J., *Artificial intelligence through simulated evolution*, John Wiley & Sons, abril de 1966.

la radiación cósmica— ha producido mutaciones en los genes de los seres vivos, originando así descendencias con características algo distintas a las de sus antepasados. La lucha por la vida hizo que sólo los mejores adaptados pudieran sobrevivir. Actualmente, con la ayuda de calculadoras capaces de trabajar en millonésimas de segundo, procesos similares pueden ser simulados —aunque de un modo simplista, claro está— en tiempos ínfimos en comparación con los reales. En vista de que la adaptación a gran variedad de situaciones es una de las características más importantes de la inteligencia, es útil investigar qué resultados se pueden obtener de ese modo.

Se trata en la práctica de considerar un “modelo”, un “algoritmo”, enfrentado a ciertas entradas (la acción que ejerce sobre él el medio ambiente), y evaluar su comportamiento mediante un análisis de sus respuestas. En caso de no ser perfecto, se introducen modificaciones (“mutaciones”) en su estructura, y si de este modo se obtienen uno o varios algoritmos más eficaces, se eliminará al primero. En este último caso diremos que nos hallamos en presencia de una nueva “generación” y que se ha obtenido una “descendencia”. El paralelo con la evolución es claro: todo organismo puede considerarse en cierto modo como un algoritmo que reaccionará de un modo determinado ante el medio ambiente, y que sucumbirá (por “selección natural”) ante la competencia de “algoritmos” más adaptados.

Si bien se puede, en principio, considerar cualquier tipo de modelo, de acuerdo con lo visto en el Capítulo 1, podemos limitarnos a las funciones recursivas o, equivalentemente, a las máquinas de Turing. Como existe una cantidad numerable de tales algoritmos, en principio nuestro problema tendría siempre solución: sólo haría falta considerar sucesivamente todas las máquinas e ir probándolas en la situación deseada. Claro que prácticamente esta idea es —por el momento— irrealizable, ya que se deberían generar millones de posibilidades antes de llegar a la que resultase eficaz. Debido principalmente a esto, y en parte a ventajas que se ven más claras con un análisis más detallado (posibilidades de reconocer regularidades en el medio ambiente, etcétera), los autores de la propuesta original (Fogel et alii) se limitaron a experimentar con un tipo muy especial de algoritmos: aquellos que se pueden representar mediante autómatas finitos, un subconjunto de los expresables por máquinas de Turing. Las “mutaciones” serán ahora sencillamente cambios en las funciones de transición y de salida.

Los resultados obtenidos fueron muy interesantes, fundamentalmente en lo que respecta a la posibilidad de obtener un sistema que “prediga” el comportamiento del medio ambiente;

en otras palabras, que realice una *inferencia inductiva* sobre la base de la experiencia. El proceso de abstracción, predicción y contrastación de resultados llevado así a cabo es esencialmente el *método científico*, que le ha permitido al hombre avanzar tanto en sus conocimientos, por lo que este enfoque de la inteligencia artificial podría ser relativamente natural. La insuficiencia de los medios técnicos actuales (en capacidad y velocidad), torna a este método —al igual que los de la biónica—, poco eficaz. Se han alcanzado, sin embargo, algunos resultados de consideración, y el lector podrá hallar abundantes ejemplos, estadísticas y aplicaciones en la bibliografía antes citada.

3.4.1. Ejemplos de la biónica. El Perceptrón.

En varias ocasiones nos hemos referido al enfoque del problema de la inteligencia artificial por parte de la biónica, punto de vista que ahora discutiremos en lo que se refiere a los modelos de “redes nerviosas” mediante un ejemplo clásico: el “Perceptrón” de Rosenblatt¹. En los estudios de este tipo, uno de los problemas más atacados es el R.C. sensorial, y ésta es la razón por la cual los términos usados son generalmente pertenecientes a la fisiología.

El Perceptrón consta de 3 tipos de componentes. En primer lugar consideraremos que posee un conjunto de elementos *receptores*, agrupados de tal forma de constituir una “retina”. Cada receptor puede enviar un tipo de señal continua (“perceptrones continuos”) o discreta, en cuyo caso se consideran generalmente sólo dos valores.

El segundo tipo de componente es la “neurona” (también llamado “elemento — A”) denominado así por simular en cierto modo el comportamiento de las células nerviosas. Estos elementos poseen dos tipos de entradas: “excitatorias” e “inhibitorias”, siendo varias las entradas y una única salida. Asociadas a cada una se tiene un factor de ponderación o “peso” y una regla que determina la señal de salida y su peso a partir de las entradas provenientes de los receptores. Cabe aquí una clasificación entre neuronas continuas y discretas. Para las primeras, Rosenblatt estableció una función de salida igual a la diferencia entre el valor total de las señales excitatorias e inhibitorias, multiplicadas por su peso. En el caso de las discretas, se supone dado también un valor límite o “umbral”; si la diferencia antes

¹ Se podrán encontrar estudios al respecto en la mayoría de los libros sobre aprendizaje y en particular RC en sistemas artificiales.

señalada es menor que esta cota, emitirá un cero, mientras que si es igual o la supera, la señal de salida tendrá como valor el "peso" de la neurona.

Un último tipo de elementos está constituido por los acumuladores (o "elementos-R"), a los cuales se conectan las salidas de las neuronas, separadas en ciertos grupos. La salida de los elementos-R es simplemente la suma de las entradas correspondientes, y la "respuesta" de todo el sistema ante una señal determinada consistirá en identificar al acumulador que dio un resultado mayor.

Para resolver ahora un cierto problema de RC, se deberán exponer a la "retina" las imágenes a clasificar y, de acuerdo a la respuesta del sistema, modificar los "pesos" de las neuronas (en una cantidad fija, o en proporción a su valor), aumentando los de aquellas que influyeron correctamente y disminuyendo aque-

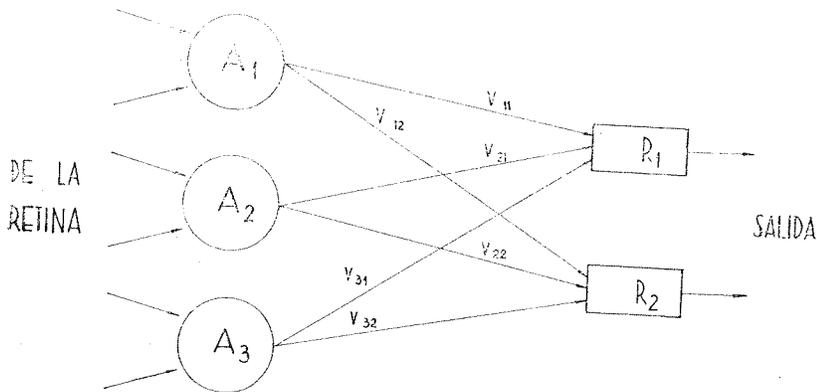


Figura 1.

llas conectadas al acumulador que da la respuesta errónea. En un principio, las conexiones y asignación de pesos pueden ser hechas totalmente al azar, y luego el sistema efectuará los ajustes necesarios de acuerdo con los "premios" y "castigos" recibidos.

El estudio teórico del funcionamiento se puede llevar a cabo desde varios puntos de vista, introduciendo conceptos tales como las "secuencias de aprendizaje" y la distribución de probabilidades de aparición de las diversas configuraciones². Señalemos,

² Glushkov, Viktor M., *Introduction to cybernetics* (Academia de Ciencias de Kiev, Unión Soviética); en inglés: Academic Press, 2ª edición, N. York, 1968.

por otra parte, que se pueden efectuar muchas variaciones sobre el esquema planteado, tales como la interconexión de las unidades "A" o conexión de cada neurona con varios acumuladores, asignando distintos "pesos" a cada una, lo cual determina variaciones en los desarrollos teórico y experimental del tema. Un ejemplo es el Perceptrón de la Fig. 1, y donde A_n indica el n-ésimo elemento-A, R_1 y R_2 los acumuladores y V_{ij} el "peso" de la conexión A_i y R_j .

Si se le quiere enseñar a distinguir entre un cuadrado y un triángulo, por ejemplo, éstos se "proyectarán" adecuadamente sobre la retina; si lo que se desea es que $R_1 > R_2$ cuando la imagen sea un cuadrado, y $R_1 < R_2$ en caso contrario, cada vez que se presente un ejemplo y se obtenga una decisión errónea se reducirá el valor de los V_{ij} correspondientes, y se aumentan si la identificación es correcta. Como se podrá notar, el Perceptrón es en sí un dispositivo de RC "en paralelo", por lo que una adecuada normalización previa de la imagen, por ejemplo, resulta de gran eficacia.

Para un estudio más detallado de estos tipos de dispositivos, remitimos al lector a la amplia bibliografía sobre el tema. En una interesante variación Harth³ propone un sistema que demuestra una considerable capacidad de asociación de ideas. En general, podemos decir que, debido fundamentalmente al señalado problema de complejidad, no es posible esperar la emergencia de "inteligencia" de estos dispositivos, en parte debido a la poca estructura inicial en comparación a los programas heurísticos. Sin embargo, para algunas experiencias de RC y modelación teórica, es interesante el estudio de estos sistemas.

3.4.2. Estudios probabilísticos y estadísticos.

Como ya indicamos en variadas oportunidades, es posible un estudio matemático de algunos modelos particulares de procesos de aprendizaje. No es, sin embargo, nuestro propósito dedicarnos a ellos, lo cual está fuera del alcance del presente trabajo, además de requerir una cierta base teórica. En lo que se refiere a la variación del comportamiento en los autómatas (dejando de lado las teorías estadísticas de la psicología matemática), se puede encontrar un interesante estudio en la ya citada obra del soviético Glushkov: el aprendizaje consiste en principio en considerar cierta secuencia de entrada (la "experiencia") a

³ Harth, Erich M., *Brain models and thought processes*, en *Automata theory*, A. Press, N. York, 1966.

la que es sometida la máquina a partir de su estado inicial, llegando ésta a un estado distinto, con lo que una nueva señal le hará comportarse de un modo diferente al que hubiese tenido lugar sin la experiencia. Se definen luego los tipos de "preguntas" y "respuestas" a que se expondrá el autómata ("ciclado de la información") y se obtiene, luego de introducir algunos conceptos auxiliares, un teorema relacionando la posibilidad de modificación temporal de las respuestas con la existencia o no de una "memoria". Glushkov aplica posteriormente sus resultados al Perceptrón y en general al RC.

Un estudio más avanzado se puede llevar a cabo mediante la utilización de resultados sobre "cadenas de Markov" de la Teoría de las Probabilidades, mediante la definición, por ejemplo, de los "autómatas probabilísticos", donde una tabla indica las funciones de transición y de salida no de un modo determinado sino dando la distribución de probabilidades de las posibles respuestas. También se pueden definir una entrada "aceptada" por un autómata como aquella con una alta probabilidad de llevar a un "estado final"¹. Los resultados parecen ser en este caso más importantes, sin embargo, para el estudio de circuitos con dados "coeficientes de confiabilidad" (aquí el factor probabilístico) que para una investigación teórica en inteligencia artificial.

Respecto de los métodos estadísticos aplicados a ciertas áreas particulares, una de las más desarrolladas ha sido la del RC, obteniéndose variados algoritmos en lo que se refiere tanto a los procesos de clasificación como de reconocimiento. El problema puede resolverse bajo ciertas condiciones mediante la aplicación de resultados de la investigación operativa y de la teoría estadística de la decisión². Aquí se trata no de diseñar métodos heurísticos para el reconocimiento de objetos particulares (figuras geométricas, letras, etcétera), sino que se trata de un enfoque mucho más general, encontrando incluso la propia máquina las "características" a tener en cuenta, y abstraendo en cierto modo las particularidades que distinguen los conceptos de una dada clase ("detección de la configuración").

La aplicación de estos métodos supone la previa codificación de los datos de entrada como vectores n-dimensionales, donde cada uno de los valores indica una dada medición. Es clara la

¹ Rabin, Michael O., *Lectures on classical and probabilistic automata* (Universidad Hebrea de Jerusalén, Israel), en *Automata theory*, 1966.

² Tema que se puede encontrar expuesto en los textos usuales de estadística, ver por ejemplo algunas formalizaciones en Ríos, Sixto, *Métodos estadísticos*, McGraw Hill, Madrid, 1967.

generalidad de este proceso, ya que podemos suponer, si se quiere, que cada coordenada corresponde a la salida de uno de los puntos de una imagen de televisión. Así cada señal tendría asignada un vector con su descripción; pero se debe notar aquí que si bien los estudios teóricos se aplican perfectamente a este caso (no se especifica el valor de "n"), es prácticamente imposible implementarlos, ya que tal cosa supondría trabajar con vectores con miles de coordenadas. Es por eso que se debe poseer un número limitado de parámetros suficientemente representativos, lo cual es justamente el factor que limita el uso de estos procedimientos.

El proceso del RC se enuncia entonces como el de decidir la pertenencia de un ejemplo a una de varias clases, las cuales se conocen, a su vez, por otros ejemplos para los que se especificó en una etapa previa (aprendizaje) la clasificación correcta. El método consiste en hallar una regla de decisión óptima de acuerdo con ciertos criterios, fundamentalmente el minimizar la posibilidad de error en la posterior etapa de reconocimiento. En esta regla de decisión incluimos una distribución de probabilidades de las diversas clases, la cual se puede obtener mediante la determinación de líneas que las separen (en el espacio n-dimensional en cuestión), resultado que puede, a su vez, ser logrado mediante la aplicación de transformaciones de coordenadas que resulten en un agrupamiento de los ejemplos de la misma configuración (en inglés, "clustering").

En un interesante trabajo de Sebestyen³ se dan gran cantidad de métodos para llevar a cabo estas transformaciones, incluyendo una aplicación al reconocimiento de la voz humana diferenciándola de otros sonidos, sobre la base de mediciones efectuadas sobre fonogramas.

Además de estos métodos, existen otros "aproximados", que son más económicos y simples, aunque menos confiables, tales como el almacenamiento de los ejemplos aparecidos durante el período de aprendizaje y el reconocimiento mediante la identificación del más "próximo" al caso a decidir. Inclusive es posible un estudio analítico del Perceptrón en cuanto a su papel en el RC, concluyéndose que es un dispositivo cuyo comportamiento es equivalente al de la aplicación de transformaciones *lineales* sobre los datos; y sobre éstas se demuestra a su vez que sólo pueden distinguir las clases cuando éstas forman agrupamientos muy regulares y bien separados (matemáticamente, deben ser convexas).

³ Sebestyen, George S., *Decision-Making processes in pattern recognition*, McMillan Comp., N. York, 1962. Más recientemente, *Perceptrons*, de Minsky y S. Papert, Cap. 12, M.I.T. Press, 1969.

El proceso de realizar predicciones sobre los símbolos que pueden aparecer en la entrada, basándose en los ya recibidos, puede interpretarse como una abstracción del proceso de inferencia inductiva, de generalización a partir de la experiencia.

Uno de los primeros trabajos respecto de la formalización del concepto de inferencia inductiva y métodos de implementarla lo realizó Solomonoff a principios de la década del 60. En 1964 se publicaron gran parte de sus resultados⁴; el problema se puede, sobre la base de la discusión anterior, plantear así: "dada una secuencia T , ¿cuál es la probabilidad de que su continuación sea la secuencia a ?"

Lo que se requiere es, en principio, el cálculo de las probabilidades con que cada secuencia de símbolos podría aparecer; una vez hecho esto, si se tiene la sucesión T , la probabilidad de que ahora aparezca a será:

$$P(a/T) = \frac{P(Ta)}{P(T)}$$

o sea el cociente entre la probabilidad de la secuencia entera Ta y la probabilidad de aparición de T .

El cálculo de dichas probabilidades se debe efectuar, por supuesto, de manera aproximada, ya que para conocerlas exactamente necesitaríamos conocer la regla que las genera. Las más eficientes heurísticas se obtienen de la suposición simplista de que el comportamiento del medio debe poder explicarse con la hipótesis más "económica" en un cierto modo. Notemos que si una teoría "explica" una secuencia sobre la base de las regularidades (ciclos, etcétera), en ella observadas, la explicación será más simple que en el caso de sucesos completamente aleatorios, en donde el único modo de describirlos exactamente debería consistir en la enumeración de los sucesos, la cual podría ser extremadamente extensa. Para formalizar el concepto de "descripción" más corto, se requiere la utilización de las máquinas de Turing universales.

Si M es una máquina de Turing arbitraria (dicho de otro modo, un *algoritmo* cualquiera), una máquina de Turing N es "universal" si existe una cadena " m " (que depende sólo de M y N) de tal modo que para cualquier secuencia x se verifique:

$$N(mx) = M(x)$$

dicho de otro modo, al leer " mx ", la máquina universal inter-

⁴ Solomonoff, R. J., *A formal theory of inductive inference*, Parte I: *Information and control* 7, 1 (marzo 1964), y Parte II: *Information and Control* 7, 2 (junio 1964).

preta "m" como las instrucciones a seguir y las aplica a x . Otra interpretación de la definición consiste en considerar "m" como un programa de simulación para la máquina N, de modo que reaccione ante cualquier entrada como lo haría M. En cualquiera de los casos, el hecho que nos interesa es que, en definitiva, todos los comportamientos que nos interesen pueden ser producidos por N, dada una entrada adecuada.

Entonces, para una secuencia T, si se tiene la cadena "S" tal que $N(S) = T$ (o sea, al aplicarle como entrada a N, se obtiene una salida T), se dirá que "S" es una "descripción" de T. Notar que diferentes entradas pueden llevar a la misma salida, por lo que la descripción no es única. Sobre la base de este esquema, entonces, se asignan las probabilidades de tal modo que se suponga mayor la posibilidad de obtener un evento cuyas descripciones sean más cortas. Esta regla se puede mejorar teniendo en cuenta las posibles continuaciones de T y sus descripciones, ya que la observación fue sólo parcial; de este modo se toman en consideración casos como el de "abcdabcdabcdab", que si bien tienen en principio una descripción larga puede, teniendo en cuenta la continuación "cd", ser descrita como 4 "abcd" mediante la indicación de la regularidad observada. Otra heurística que se puede aplicar consiste en aumentar la probabilidad "a priori" de las secuencias que admiten mayor número de descripciones, lo que se podría interpretar considerando a los "S" como "causas" de T, lo que equivaldría a postular como más probable aquel evento que se puede producir por mayor cantidad de causas.

Si bien las justificaciones pueden ser convincentes en mayor o menor grado, la efectividad de los métodos sólo puede ser probada a partir de la experimentación, como bien lo señala Solomonoff. En particular, éste aplicó sus resultados a varios ejemplos (ver los artículos citados), en especial a la inferencia gramatical. Para esta última, se basó en un continuo "ajuste" a los datos mediante sucesivos cambios o mutaciones en las reglas, intentando obtener la gramática correcta mediante aproximaciones sucesivas.

Si bien la eficiencia de ésta y otras formalizaciones de la inferencia inductiva pueden ciertamente cuestionarse, específicamente en vista del uso cada vez mayor de descripciones simbólicas y modelos adecuados, estos métodos "estadísticos" tienen un interés propio desde el punto de vista de la generalidad formal y de su aplicación a ciertas áreas particulares (ajuste de curvas, etcétera).

El problema de las descripciones más cortas y las regularidades está estrechamente vinculado con el de la "complejidad"

y las definiciones de "secuencias al azar"⁵, y el tema de la inferencia gramatical es actualmente motivo de varios estudios (en particular Stanford, 1970).

3.4.3. Otras formalizaciones.

En los últimos años es posible apreciar una creciente tendencia hacia la formalización de los diversos conceptos referidos a la inteligencia artificial, tomando caminos distintos de los enfoques previos, en la mayoría de los casos basándose en la lógica matemática y en la Teoría de Grafos.

Algunos autores enfatizan la importancia de las adecuadas descripciones de las situaciones¹, como indispensables para una adecuada capacidad de resolución de problemas, poniendo así de manifiesto el papel del reconocimiento de configuraciones en la inteligencia artificial.

Mediante la adecuada formalización en términos de la lógica de predicados, se espera poder implementar adecuados sistemas capaces de razonar sobre las más diversas áreas, aprovechando los grandes adelantos conseguidos últimamente en cuanto a la obtención de algoritmos de demostración.

También la inferencia inductiva podría ser formalizada, con un enfoque radicalmente distinto del indicado en la sección anterior. Se trataría ahora de manejarla de un modo similar a la deducción; si se considera dada una cierta "evidencia experimental", una sentencia será una válida *inferencia inductiva* de otra *respecto de esta evidencia* si la implica en sentido lógico y es consistente con dichos datos. Meltzer² propone un sistema automático de inducción usando un sistema axiomático deductivo e *invirtiendo* todas sus reglas, obteniendo así todas las sentencias que implicarían a una dada, y realizando luego las respectivas pruebas de consistencia.

Respecto de los juegos y las estrategias de evaluación, D. S. Johnson³ formaliza un tipo especial de juego secuencial, en el

⁵ Se puede encontrar una discusión en castellano de estos temas en *Hacia una definición matemática de "vida"*, Greg Chaitin, Universidad Tecnológica Nacional, Buenos Aires, 1969.

¹ Banerji, R. B., *Theory of problem solving, an approach to Artificial Intelligence*, Elsevier, 1969.

² Meltzer, B., *The semantics of induction and the possibility of complete systems of inductive inference*, *Artificial Intelligence*, Vol. I, 1970, y *Power amplification for Theorem-Provers*, *Machine Intelligence* 5, 1970.

³ Johnson, D. S., *Look-ahead strategies in one person games with randomly generated game trees*, A. I. Memo 205, M.I.T., julio de 1970.

cual la probabilidad de que un estado sea ganador se corresponde al valor de la función de evaluación estática de los juegos usuales. Sobre la base de este modelo, investiga las diversas estrategias de análisis, demostrando cuáles resultan óptimas respecto de los juegos en cuestión.

El anterior se puede ver como un caso particular del enfoque de la *exploración heurística* como la búsqueda de caminos en un grafo orientado, una de las formalizaciones más interesantes del tema (Doran y Michie, 1966/69), inclusive con aplicaciones directas tales como la consideración del medio ambiente como un grafo y la utilización de los algoritmos así encontrados para guiar al robot (Hart, Nilsson y Raphael, 1968), usando como "función de evaluación" una estimación de las distancias observadas por el sistema.

A estos temas y a su modelo particular se refiere Ira Pohl ⁴, quien estudia la efectividad de los métodos de búsqueda en grafos suponiendo dada una cierta función de estimación, que en cada caso particular dependerá de las características del dominio, y que en general se supondrá aproximante de la distancia que separa el punto en cuestión del *nodo terminal* del problema. Tomado como "función de evaluación", a una suma ponderada del estimador y de la cantidad de nodos ya recorridos, demuestra diversos resultados respecto de la capacidad del algoritmo para encontrar caminos óptimos, especialmente cuando se conoce una cota del error del estimador. Un programa que ya estudiamos, MULTIPLE (3.1.4.) se basa en un enfoque similar en varios aspectos.

En la sección de juegos nos referimos a un programa que juega Poker diseñado por D. A. Waterman ⁵, y con éste finalizaremos esta corta mención de algunas de las recientes formalizaciones, en las que no nos es posible extendernos por motivos de espacio. Se trata de una aplicación del autor, a un área específica, de su método de representar heurísticas, y su aprendizaje por un programa. Luego de destacar la necesidad de la separación de las heurísticas del resto del programa, con el fin de permitir su fácil modificación, Waterman propone una formalización de éstas como "producciones" análogas a las utilizadas en las gramáticas generativas (capítulo 1).

Así se pueden formalizar cómodamente los conceptos de *definición heurística* y de *regla heurística*. El modelo consiste en una serie de producciones que indican la acción a tomar de

⁴ Pohl, Ira, *Heuristic search viewed as path finding in a graph*, *Artificial Intelligence*, Vol. I, 1970.

⁵ Waterman, D. A., 1970, *op. cit.*

acuerdo con las variables del programa. Así, una regla del tipo

$$“(A1, \&, C2) \rightarrow (\&, b - 5, \&)”$$

indicaría: “si la variable A está en un valor del rango simbolizado por $A1$, y si C está en el rango $C2$, restar 5 al valor actual de B (simbolizado por “ b ”), independientemente del rango actual de B (por ello el “ $\&$ ” a la izquierda), y sin afectar los valores de A y C ”; por otra parte, una definición como “ $A1 \rightarrow A, A \geq 13$ ” indica que $A1$ representa el rango de valores de A mayores o iguales que 13.

Dada ahora una situación (representada por un “vector de estado” en el que constan todas las variables significativas del sistema), se realiza un “parsing”, obteniéndose el vector simbólico correspondiente. Si por ejemplo el vector de estado es “(3, 2, 1)” y 3 está en el rango denotado por $A1$, 2 en el de $B1$ y 1 en el de $C1$ y el de $C2$, el resultado del análisis será una lista tal como $(A1, B1, [C1, C2])$, decidiéndose ahora la acción a tomar de acuerdo con la regla heurística correspondiente a dichos rangos.

La utilidad del uso de “vectores simbólicos” para el rango de las variables es clara: caso contrario, habría que especificar una regla para cada combinación de valores de las variables. Así, por ejemplo, si definimos “ $A\text{-Grande} \rightarrow A, A \geq 100$ ”, cada vez que A tenga un valor mayor o igual que 100 (por ejemplo, A puede ser el valor de la apuesta en una mano de poker), el programa “reconoce” que A es “grande” y actúa correspondientemente.

El elemento de aprendizaje se introduce a partir de los resultados obtenidos con la experiencia, y puede consistir en la modificación de las reglas, ampliando o disminuyendo un rango (en el ejemplo anterior, si se decide que un valor de A de 100 no era tan importante como para tomar las acciones correspondientes a una apuesta “grande”, digamos abandonar, se puede cambiar la definición anterior a “ $A\text{-Grande} \rightarrow A, A \geq 500$ ”), creando en otros casos nuevas reglas, o eliminando otras redundantes.

El problema de la “asignación de responsabilidades” se resuelve considerando como causas de un fracaso a la regla que determinó la decisión y a las superiores (para tomar una decisión, una vez que se obtuvo el “vector simbólico” mediante el análisis ya indicado a partir del vector de estado, se prueban sucesivamente las reglas hasta encontrar una coincidente en su parte izquierda); la acción tomada una vez detectada la “culpable” es la de modificarla suficientemente como para que el error no se repita.

Waterman implementó un modo de aprendizaje “explícito” en que se le hace saber al programa las variables significativas,

cuál hubiese sido la decisión correcta, y la razón exacta por la cual se debió haberla tomado, y también un aprendizaje “implícito” en el cual, gracias a una axiomatización y un procedimiento de demostración de teoremas, el programa (experimentado en el caso del poker) realiza las deducciones necesarias por su cuenta. En ambos casos se logra una sorprendente capacidad de aprendizaje, con lo que el enfoque resulta altamente prometedor.

Los trabajos en el área de la computación teórica y la inteligencia artificial están en pleno desarrollo, no existiendo hasta el presente un enfoque unificado de los diversos problemas, razón por la cual es difícil señalar con precisión cuáles están destinados a servir como base para la solución de los grandes problemas que se plantean. Es así que nuestra meta a lo largo de este estudio no ha sido ni remotamente un enfoque de tipo general sino solamente la indicación de algunas de las tendencias más sobresalientes, con el objeto de adentrar al lector en la formulación de algunos de los problemas, la enunciación de los inconvenientes que se presentan para su resolución y las líneas de ataque propuestas por los investigadores.

Capítulo 4. OBSERVACIONES FINALES

En el capítulo 2 vimos algunos de los modos en que pueden simularse y estudiarse diversas características del comportamiento humano por medio de computadoras, lo cual nos sirvió como introducción al capítulo 3, donde analizamos ya los problemas esenciales que se plantean al intentar obtener un comportamiento "inteligente" en un sistema artificial. Sin embargo, nuestro estudio dista de ser completo por varias razones. Ante todo, el hecho de que nos hayamos concentrado en los principales proyectos (principalmente los desarrollados en América del Norte), dejando de lado muchas contribuciones, también importantes, por parte de científicos de otras universidades y de otros países. Además, aún dentro de las áreas estudiadas, no fue posible repasar algunas soluciones propuestas recientemente. Existe también cierta falta de sistematización expositiva y no se han analizado las consecuencias que los estudios en curso podrán tener sobre el futuro de la humanidad, incluso las posibles aplicaciones concretas a corto y mediano plazo.

La falta de una adecuada referencia en lo que respecta a los dos primeros puntos señalados la atribuimos fundamentalmente a la restricción de espacio a que nos debimos someter para no llegar a un volumen enciclopédico; esto se ha traducido también en la tendencia a resumir las explicaciones y no extendernos en detalles.

Respecto a la falta de sistematización expositiva, debemos señalar que no existe un método tal en un área tan compleja y amplia como la I.A., por lo que consideramos que un estudio histórico y motivado constituía el modo más adecuado de enfrentar el problema de exposición. No obstante, en algunas áreas particulares, se pueden establecer tentativamente ciertos enfoques formales más o menos coherentes, como en el caso de la "resolución de problemas".

El estudio detallado de las formalizaciones existentes en campos restringidos estuvo fuera del alcance de este trabajo, debido a los requerimientos matemáticos que implican y en

cuanto al último punto, preferimos omitirlo por su carácter fundamentalmente especulativo.

En lo que resta nos limitaremos a señalar muy someramente algunos aspectos relevantes de los puntos mencionados. Son prácticamente nulas las referencias que hayamos hecho a los trabajos soviéticos en el área; lo que no significa que carezcan de importancia o que los científicos los consideren carentes de interés. Por el contrario, las apariencias indican que el ritmo de investigación es muy intenso, aunque la falta de documentación y dificultad de acceso es notoria. En ciertos campos tales como los enfoques matemáticos de ciertos temas como R.C., los trabajos soviéticos han alcanzado gran nivel y relativa difusión. Asimismo serían varias las experiencias en el campo de las "redes neurales", inclusive en lo que respecta a su simulación por computadora, habiéndose experimentado con modelos de este tipo en la computadora M-20 del Instituto de Cibernética de Kiev¹, bajo la dirección del destacado científico N. M. Amosov, a quien ya nos referimos en el capítulo 2.

Sin embargo, al igual que los científicos occidentales, aceptan la poca utilidad de estos métodos en la obtención de una verdadera inteligencia; en la revista de la Academia de Ciencias de Ucrania se define la posición²: "los procesos cerebrales han interesado siempre a los científicos, pero los resultados de las investigaciones en esta área han sido demasiado variados como para su uso directo en el diseño de dispositivos o programas para computadoras digitales... los neurobionistas³ han creado neuronas artificiales y sintetizado redes neurales, tratando de aplicarlas a los problemas de reconocimiento visual. Estas neuronas artificiales, sin embargo, sólo duplican ciertas propiedades superficiales de las neuronas naturales, y la interacción entre los subsistemas del cerebro es aún poco clara. Los intentos de modelar el cerebro usando neuronas artificiales, por lo tanto, han sido infructuosos". Aparentemente, los proyectos soviéticos reflejarían las actitudes estadounidenses, pero encontramos un elemento quizá novedoso en la aplicación de la biónica a la inteligencia artificial: "...en el futuro, se usarán probablemente redes neurales vivas como elementos adaptativos en sistemas autorregulables; la factibilidad de adaptar un cerebro de gato para controlar cohetes aire-aire ha sido estudiada. Al respecto, se ha sugerido que un dispositivo óptico

¹ De *Computers in Soviet Medicine* en *Soviet Cybernetics Review*, Rand Corp., California, 1970.

² De *Modern Problems of Neurobionics*, extractado en *SCR*, 1970.

³ Quienes estudian la *neurobiónica*, rama de la biónica dedicada al cerebro.

electrónico se coloque en el cohete, el cual transmitirá señales directamente al cerebro del animal, donde debido a un condicionamiento previo, se producirían ciertas reacciones mentales, permitiendo al proyectil localizar su objetivo”.

En cuanto al área de la programación heurística, se desarrollaron algunos programas de ajedrez. Uno de éstos, del Instituto para la Física Teórica y Aplicada de Moscú, ganó una partida contra el programa de A. Kotok (3.1.1.). Ultimamente, el famoso ajedrecista M. Botvinnik dio a conocer el resultado de siete años de investigaciones en el desarrollo de un “algoritmo que juegue al ajedrez” basado en la observación de que el buen ajedrecista fija su atención casi exclusivamente en un pequeño número de piezas y una parte muy restringida del tablero, analizando la posibilidad de una movida en particular e incluyendo las demás regiones del tablero en un proceso incremental, a medida que se reconocen como relevantes. En base a esto, la estrategia consiste en una continua ofensiva; si las trayectorias que se deciden para las piezas en el curso del ataque prueban ser riesgosas, se intenta buscar la intervención de otras piezas con el fin de mejorar la posición. La implementación del algoritmo como programa de computadora fue asignada al matemático V. I. Butenko, usando una M-220. Para quienes no estén al tanto de los modelos de computadoras soviéticas, aclaremos que la M-20 es una máquina binaria tridireccional, con una memoria de 4K palabras de 45 bits cada una y una capacidad de realizar 20.000 operaciones por segundo; se la considera algo inferior a la IBM 7090. La M-220 es una versión transistorizada de la anterior, pero con una memoria expandible a 16K y palabras de 45 bits.

En base principalmente a los estudios del ya nombrado Asomov, se han diseñado varios modelos de programas con capacidades “emotivas” además de intelectuales, basados en memorias de tipo asociativo; se reporta al respecto la influencia de la interacción de los factores intelectuales y emocionales en un programa de comportamiento verbal y un modelo de toma de decisión en un dominio restringido. El programa EMIK⁴, en Kiev, contesta de acuerdo a la “evaluación emocional” que efectúa a las preguntas, y es un ejemplo de la aplicación de las teorías de Asomov, quien afirma que para entender a los hombres la computadora deberá estar “armada” con emociones, ya que en su opinión, el proceso pensante es interacción de lo intelectual con lo emotivo.

⁴ Karkvosky, *Una Computadora “Emocional”*, trad. al inglés en SCR, Rand Corp., 1970.

En otras áreas los soviéticos se muestran igualmente activos; en el Instituto de Cibernética, por ejemplo, se experimentó en el RC auditivo, lográndose un sistema capaz de reconocer palabras emitidas por diversas personas, con una eficacia del 93 % usando 73 palabras. Para realizar en general las experiencias de RC, se construyó en el mismo sitio un sistema constituido por lectores de papel y películas y un dispositivo de control ⁵. Mediante el ajuste de diversos parámetros, se puede utilizar al sistema en la experimentación en RC estadístico, iterativo (tipo Perceptrón, etc.), estructural (“lingüístico”), etc. Esto se lleva a cabo gracias a la existencia de algoritmos de reconocimiento “standard” (capaces por ejemplo de transformaciones geométricas de la entrada), los que se disponen de acuerdo a la investigación a realizar. Su fácil modificación lo hace apto para tales estudios, y ya se aplicó en reconocimiento de la escritura normal y otras experiencias. Con el fin de estudiar posibles métodos de almacenamiento de la información en las futuras computadoras, utilizando las estructuras que a tal fin posee el hombre, se estudiaron en la Universidad de Leningrado modelos de memorias de corto ⁶ y largo ⁷ plazo. El primero incluye el almacenamiento de la información como un proceso sucesivamente dinámico y la formación de estructuras estáticas (transformación espacio-temporal de la información), y el segundo fue usado en experiencias de reconocimiento visual de letras.

Finalmente, señalemos que el campo de la demostración automática de teoremas también ha sido objeto del interés soviético, como lo demuestran el Primero (1964) y el Segundo (Lituania, agosto de 1969) Simposios sobre el tema. El principio de resolución resulta también en este caso base del estudio en el área, aunque se experimenten también otros enfoques.

También se ha experimentado con nuevos dispositivos de memoria en el Japón, basándose en una estructura asociativa y homogénea, capaz de permitir un comportamiento “inteligente” de un modo muy simple. El “Asociatrón” ⁸ combina los estudios “micro” (biónica) y “macro” (psicología), al tener una estructura similar a la neural, pero más general, al permitir modelar comportamientos superiores. Se trata de “neuronas” que reci-

⁵ Proc. del IJCAI/71.

⁶ Bereznaya, I. J., Granovskaya, R. M.: *The Model of Human Short-Term Memory*, Proc. IJCAI/71.

⁷ Granovskaya, R. M.: *The Model of Human Visual Long-Term Memory with Ability for Generalization*, Proc. IJCAI/71.

⁸ Nakano, L., Uaguno, J.: *Information Processing Using a Model of Associative Memory*, IJCAI/71.

ben la información y unidades de memoria que las interconectan de a pares. Las percepciones ("entes") se almacenan sucesivamente en la misma zona, pero luego la presencia de una parte de un ente es suficiente para la identificación completa, con una cierta probabilidad de error, debida a la superposición de percepciones. Esta indeterminación es útil sin embargo, ya que permite el descubrimiento de similitudes entre "entes" parecidos. El Asociatrón fue simulado en computadora y se construyó también un modelo físico, lográndose algunos resultados interesantes en el aprendizaje de juegos y en una "formación de conceptos" elemental.

También en el campo de los robots han experimentado los japoneses, exponiendo en particular un modelo capaz de construir un objeto con cubos, prismas y ladrillos, basándose en un diagrama que se le presente. Se trata en este caso de que el robot "imagine" la forma del objeto a partir de varias vistas laterales, y que lo analice, determinando las partes necesarias para su armado con las piezas de que dispone. Es claramente interesante la posibilidad de incorporar en el futuro mecanismos de este tipo en la producción de artículos en pequeñas cantidades, donde no se justifican máquinas especiales.

Volviendo a las memorias asociativas, digamos que las "redes semánticas" basadas en modelos como el de Quillian (3.3.2.) han sido experimentadas en varias áreas relacionadas con la percepción del lenguaje natural. Se ha afirmado⁹ que es preferible el uso de tales estructuras antes que las basadas en formalismos de la lógica de predicados en los S.I.; en el segundo caso se debe demostrar un teorema al contestar una pregunta, lo cual resultaría más complicado y menos natural que la utilización de una red semántica. En todo caso se dan algoritmos que permiten transformar una representación en otra, permitiendo así usar métodos lógicos en caso de que una respuesta no pueda ser hallada mediante una simple búsqueda en la red.

Las reglas o "producciones" usadas por Waterman en su programa de aprendizaje de heurísticas (3.4.3.) han tenido también gran aplicación como estructuras, habiéndoselas aplicado por ejemplo en el análisis automático de protocolos (una importante tarea en la psicología), para modelar al sujeto de la experiencia. También se usan para representar las "teorías" creadas por el programa Meta-Dendral, una interesante experiencia que se realiza en Stanford en el área de la formulación automática de hipótesis y teorías (*inducción*) en la química.

⁹ Simmons, R. F.; Bruce, B. C.: *Some Relations between Predicate Calculus and Semantic Net Representations of Discourse*, Proc. IJCAI/71.

La inducción es también estudiada en Escocia, en base al esquema de Meltzer (3.4.3.), y recientemente se diseñó un programa¹⁰ capaz de generalizar y realizar "deducciones inversas" en base al principio de resolución. Se ha experimentado con la teoría de grupos: dados ciertos productos de pares de elementos del grupo, el programa induce leyes de conmutatividad, asociatividad y existencia de identidad.

El desarrollo de técnicas de demostración de teoremas efectivas como refinamientos del principio de resolución, ha dado lugar a varios intentos de aplicación de estos algoritmos, aparte de los ya enunciados. Así, representando los datos como axiomas (recordar el AT de McCarthy y el sistema de F. Black), se pueden diseñar S.I. y "solucionadores de problemas" más eficientes; el primero en llevarlos a la práctica (usando el enunciado principio) fue C. Green¹¹, quien en 1969 presentó su programa "QA3" como tesis de doctorado en Stanford y en la Primera Conferencia Internacional de Inteligencia Artificial. Parte del trabajo de Green se refirió a los procesos necesarios para extraer la respuesta a partir de la demostración del teorema usado como pregunta. Este no es un proceso trivial, debido a la forma en que se demuestran los teoremas por resolución; si bien podría el programa contestarnos que "existe un elemento que cumple con la condición pedida", para encontrar uno en particular habrá que estudiar los pasos que le permitieron al programa afirmar la existencia. Aún cuando puede considerarse a este trabajo como una contribución fundamental, debe tenerse en cuenta su relativamente baja eficiencia, debida a la ineficacia de la *resolución* para utilizar conocimientos *semánticos* en la solución de un problema. Es posible que con la introducción de nuevos métodos, se incremente la eficiencia de los métodos de demostración, ya sea permitiendo la adecuada utilización de la experiencia anterior (ver Kling, 4.1.), la introducción de "macropredicados" (Meltzer, 1970, op. cit.) o el uso de otras heurísticas especiales como la reducción a subproblemas¹². En ese caso, programas como el de Green aumentarían también su eficacia.

Otra aplicación interesante de los métodos lógicos está en la *síntesis automática de programas*, problema muy parecido al

¹⁰ Wilson, C. E.: "Inverse Deduction and the Generation of Inductive Hypothesis", Depto. de Inteligencia de Máquinas y Percepción, Univ. de Edimburgo, Dic. 1971.

¹¹ Green, C.: *Theorem Proving by Resolution as the Basis for Question - Answering Systems*, en *Machine Intelligence* 4, 1969.

¹² Ernest, G. W.: "The Utility of Independent Subgoals in Theorem Proving", en *Information and Control*, Abril 1971.

de la resolución de problemas; se trata de describir lo que se desea que un programa efectúe, dando la relación necesaria entre entrada y salida, y que un sistema automático lo construya a partir de tales especificaciones.

En un interesante artículo¹³ se demuestra cómo puede esto lograrse, basándose en la demostración del teorema que dice "para toda entrada existe una salida que cumple la relación pedida"; si en la demostración se emplea el *principio de inducción*, el programa sintetizado incluirá ciclos ("loops").

La *solución de problemas* (S.P.) puede considerarse, junto con el *reconocimiento de configuraciones*, la *recuperación de información* (en sistemas complejos) y el estudio de los *lenguajes naturales* como una de las grandes áreas en las que se puede dividir a la *inteligencia artificial*; áreas por cierto no del todo disjuntas, ya que frecuentemente es necesario aplicar en cada una de ellas métodos provenientes de las restantes. En particular, la S.P. puede estudiarse en dos aspectos: el problema de la *representación*, consistente en la determinación del modo más efectivo de presentar los problemas, y el de la *búsqueda* de la solución. Este último tema puede ser presentado con un enfoque y vocabulario unificado, como lo hace Nilsson¹⁴ en un interesante libro.

Un *problema* puede ser descripto por los estados *inicial*, *intermedios* (aquellos por los que se puede pasar en el curso de la solución) y *finales* (a los que se desea llegar), contándose con *operadores*, que transforman un estado en otro. La solución consistirá entonces, partiendo del estado inicial, en aplicar los operadores necesarios para llegar, mediante sucesivos estados, a la meta deseada. Formalmente, este problema es el de encontrar el *camino* entre dos nodos de un grafo orientado, y para su solución se puede usar un "estimador heurístico" que guíe la búsqueda, existiendo teoremas que aseguran en qué casos será posible hallar soluciones óptimas (3.4.3.). Una gran cantidad de acertijos matemáticos, por ejemplo, se pueden resolver así.

El enfoque anterior es el del *espacio de estados*, y se puede considerar aun un caso particular del de *reducción de problemas*, consistente en la descomposición de un problema en varios sub-problemas, y la repetición del proceso con éstos, hasta llegar a problemas *primitivos*, cuya solución es trivial. En términos de grafos, se trataría de encontrar una sucesión de nodos

¹³ Manna, Z.; Waldinger, R. J.: *Toward Automatic Program Synthesis*, en *Comm. ACM*, Marzo 1971.

¹⁴ Nilsson, N. J.: *Problem Solving Methods in Artificial Intelligence*, Mc Graw Hill, 1971.

intermedios, y reducir el problema a encontrar el camino entre cada par de nodos sucesivos. Claro está que no será fácil determinar cuáles deben ser dichos "escalones", y ese es justamente el lugar donde se deben aplicar las heurísticas apropiadas. Para solucionar un problema de este modo, se plantea una estructura de árbol disyuntivo/conjuntivo: en cada paso hay varias posibles elecciones para determinar la sucesión de submetas, y con solo resolver los problemas de *una* de esas sucesiones, el problema original está resuelto (disyunción), mientras que *en* cada sucesión se deben resolver *todos* los subproblemas que la componen (conjunción). Se hace uso entonces de métodos heurísticos de búsqueda en árboles "y/o" tales como el ya estudiado MULTIPLE (3.1.4.). Es claro que esta descripción corresponde al método del GPS (3.1.5.); y con una generalización se pueden considerar los métodos de exploración en árboles de juegos (Alfa/Beta, etc.) en este contexto, ya que en un juego buscamos *alguna* jugada buena ("o") tal que *todas* ("y") las respuestas de nuestro oponente sean malas.

4.1. ¿Y ahora qué?

No es fácil distinguir en un área como la que nos ocupa, los trabajos que podrán llegar a tener una significación especial en el futuro, y cualquier intento de hacerlo por nuestra parte sería un signo de extrema pedantería. Es por ello que nos limitamos a señalar las nuevas ideas, sin juzgarlas demasiado. En mayo de 1969 se realizó en Washington la Primera Conferencia Internacional Conjunta sobre Inteligencia Artificial (IJCAI) con la presentación de sesenta y tres trabajos por parte de autores de nueve países, y entre el 1 y el 3 de setiembre de 1971 se realizó la Segunda IJCAI en Londres, esta vez con sesenta y dos trabajos representando a once países.

Se pudo observar en la última conferencia la presentación de algunos trabajos considerables "de base" para investigaciones posteriores. En esta categoría incluimos a un sistema de programación sueco para el manejo de datos estructurados en forma lógica, útil para la construcción de S.I., y el Planner de Carl Hewitt, también para la demostración de teoremas. Asimismo, una "memoria semántica" ("Mens"), y un programa de almacenamiento y recuperación ("Mental"), que la utiliza, puede servir de sustrato a un S.I., un programa para demostración de teoremas o uno para S.P.

Nuevas ideas para implementar el *razonamiento por analo-*

gía fueron presentadas por el hindú S. Ramani, de Bombay, quien propone un sistema de S.P. basado en la capacidad de manejar fácilmente el lenguaje mediante su continuo aprendizaje; si por ejemplo se ha aprendido el *paradigma* "Juan pesa 60 kg, ¿cuánto pesa Juan?, respuesta: 60 kg"; ante la nueva entrada "Carlos mide 1,70 m, ¿cuánto mide Carlos?", el sistema reconoce las analogías formales entre ambos y responde: "1,70 m". Por otra parte, el programa se plantea a sí mismo las respuestas, pudiéndose obtener así comportamientos más complejos, como por ejemplo si se aprendieron: "¿Quién mordió a Juan?, respuesta: el perro" y "¿Por quién fue mordido Pedro?, respuesta: ¿Quién mordió a Pedro?"; entonces, ante la pregunta "¿Por quién fue mordido Juan?", el sistema plantea la analogía con el segundo paradigma, contestando con la misma pregunta en voz activa; al plantearse ahora esta última, responde utilizando el primer paradigma.

Desde un punto de vista más formal, las analogías fueron estudiadas bajo dos aspectos; en uno, al establecer una analogía entre dos problemas, si ésta cumple ciertas condiciones se pueden deducir de las estrategias que hayan resultado satisfactorias en uno de ellos, las que sirven para resolver el otro. El segundo aspecto es el que utiliza la analogía para decidir cuáles aspectos son relevantes en la solución de un problema, dados los correspondientes a un problema "similar".

El primero de los enfoques, el *cambio de representación* de un problema, fue estudiado para ciertos juegos y sus estrategias por R. B. Banerji y G. Ernst, demostrando en particular la "equivalencia" entre varios juegos y los posicionales, para los cuales existe una estrategia ganadora (3.1.1). El segundo enfoque (*Zorba*, de R. E. Kling) se aplica a la lógica de predicados, permitiendo que un algoritmo de resolución (en este caso QA3) demuestre teoremas basándose sólo en axiomas "relevantes" determinados por la analogía con un teorema similar; todos estos conceptos fueron formalizados, y un programa del tipo de *Zorba* podría llegar a ser esencial en los futuros *demostradores de teoremas*.

Los métodos de almacenamiento podrían llegar a cambiar sustancialmente debido a los resultados en memorias asociativas sobre la base de la *holografía*; en que mediante dispositivos técnicos, luego de fotografiar (holográficamente) una escena, si se ilumina sólo una parte del dispositivo con parte de dicha escena, ésta aparece en su integridad; en la Conferencia se presentó un modelo usando estos principios.

En cuanto al procesamiento del lenguaje natural, siguen en desarrollo un nuevo tipo de gramáticas ("redes de transición

recursivas”), aparentemente muy prácticas; habiéndose implementado en E.E.UU. un sistema de recuperación de información sobre piedras lunares que lo utiliza. Señalemos también la presentación de un estudio sobre programación evolutiva (34) por K. R. Fialkowski, de la Universidad Técnica de Varsovia, Polonia.

Por último se pueden advertir algunos adelantos interesantes en robótica. Estudios realizados en Italia conducen al R.C. táctil como complemento del visual; cuando alguna parte de la imagen no es clara para el robot, éste podría acercarse y aclarar las dudas tocando (con sensores especiales) los objetos a reconocer. En el mismo sentido se realizan estudios en Japón, habiéndose presentado los resultados de las experiencias con una mano artificial.

Para decidir el *planeamiento* de las acciones del robot del S.R.I. y S.P. en general, R. E. Filkes y N. J. Nilsson presentaron el sistema STRIPS (“Stanford Research Institute Problem Solver”). Se trata de una nueva concepción que combina el método de *reducción de problemas* con las técnicas de demostración de teoremas. Esta combinación permite separar las últimas del problema de la búsqueda de la solución, para lo cual son generalmente ineficientes, como ya lo hemos señalado. Aquí, el papel de estos algoritmos se limita a determinar si la *descripción de un estado* (dada formalmente) se corresponde con la de otro; y en caso de no ser así, a identificar las *diferencias* y determinar los operadores relevantes para reducirlas. Las descripciones de los estados y los operadores, junto con las *condiciones de aplicabilidad* de estos últimos, se almacenan en formato de lógica de predicados, lo que permite el uso de las técnicas mencionadas. Los resultados ya son realmente interesantes en cuanto a la compleja capacidad de razonamiento demostrada, y se espera poder aumentarla con la implementación de un aprendizaje de los problemas ya resueltos (por ejemplo, parametrizándolos de algún modo), y otros agregados en estudio.

En particular, pueden ser fundamentales los estudios en *lógicas multivaluadas*, en las que no sólo se admiten dos valores de verdad (“falso” o “verdadero”), sino toda una gama de valores, indicando el grado de incertidumbre. Es altamente probable que para realizar un razonamiento con “sentido común” sobre temas “de la vida diaria” este tipo de sistemas lógicos constituya una necesidad ineludible; inclusive en los trabajos científicos es común trabajar con datos aproximados, a los que puede asignarse un cierto “grado de confianza”. Este tipo de razonamiento podría ser usado por los robots en el proceso de planear y controlar su interacción con el incierto medio ambiente.

Constantemente se realizan, además de las mencionadas,

conferencias y simposios sobre temas particulares de la I.A. (demostración de teoremas, R.C. visual, etc.), y el intercambio de ideas es fundamental para impulsar y acelerar el desarrollo en el área.

4.2. Conclusiones

Luego de haber leído los capítulos precedentes, esperamos que el lector se haya formado una imagen bastante completa sobre el campo de la inteligencia artificial, sus problemas y posibles soluciones.

Hemos tratado de no dar definiciones estrictas y terminantes sobre los diversos términos, con el fin de no caer en un dogmatismo que sólo serviría para entorpecer los estudios. Es por ello que en todo momento preferimos comparar los diversos puntos de vista, dando en ocasiones una opinión, pero dejando abiertas las posibilidades a las modificaciones que pudiesen ser necesarias.

El tema de la definición de *inteligencia* puede, sin embargo, requerir algo más de discusión. En la introducción al Capítulo 3 hemos citado al famoso "Test de Turing", que sirvió de guía durante muchos años a los estudiosos en I.A. Si bien se puede reconocer varias ventajas a su uso como definición, no resulta del todo adecuado debido a que deja sin aclarar cuáles son los atributos de la inteligencia; además, el depender la definición del diálogo con un interlocutor humano, el conocimiento por el programador de la psicología de aquél puede llevarlo a diseñar el sistema con el fin de engañarlo, aprovechando tal conocimiento, pero sin necesariamente dar una "inteligencia" a la máquina.

No queremos dar (en realidad no conocemos) una definición de *inteligencia* amplia y sin ambigüedades. Podemos, sin embargo, asociarla a los *procesos de decisión* que se encuentran subyacentes a un comportamiento *finalista*. Estos procesos de decisión pueden apoyarse sobre una gran cantidad de medios que le permitan desarrollarse con facilidad (algoritmos de demostración de teoremas, memorias enciclopédicas, manejo del lenguaje, etc.), pero no deben confundirse con estos medios en sí. El poseer una gran cantidad de recursos ayuda a que se desarrollen dichos procesos inteligentes, pero no constituyen su esencia. ¿Cómo serán en sí esos órganos de decisión? La respuesta quizá deba esperar mucho tiempo, pero es probable que la investigación en simulación de procesos intelectuales nos señale en un futuro el camino; la exploración heurística de posibilidades y el reconocimiento de configuraciones no estarán se-

guramente ausentes. En este esquema, el desarrollo de más y mejores técnicas para efectuar las diversas tareas asociadas a la inteligencia puede considerarse como tendiente a aumentar la eficiencia del sistema cuando se completan las inteligencias artificiales. Esto no significa que las investigaciones actuales rendirán frutos sólo en un futuro lejano, ya que lo anterior se refiere a una "completa" inteligencia. Los estudios en curso tienen mientras tanto aplicaciones extremadamente importantes: la formación automática de conceptos y teorías en las ciencias, la traducción automática, el manejo de gigantescos bancos de datos, ejecución de tareas difíciles o rutinarias por parte de robots (investigación espacial o del fondo oceánico, por ejemplo), la investigación matemática y, en fin, prácticamente cualquier actividad para la que pueda requerirse su ayuda.

Al respecto, el famoso académico soviético V. M. Glushkov comentó recientemente sobre las aplicaciones de la computación: "... el gran interés, sin embargo, está en la automatización de la demostración de teoremas, el área de las teorías deductivas, y en la construcción de esquemas teóricos que generalicen los resultados experimentales. Hasta ahora sólo se han obtenido resultados iniciales tentativos; pero la perspectiva que sugieren es realmente inmensa. Usando computadoras para automatizar al menos parte de la resolución de problemas en las investigaciones, se podrían acortar algunos pasos en el desarrollo de teorías que de otro modo podrían insumir tanto como una década. Esto aumentaría poderosamente la potencia intelectual de la humanidad. Quizá el resultado más importante de este tipo de automatización sería el desarrollo de teorías complejas, un hecho ahora imposible. Los resultados prácticos de tales teorías incrementarían el poder del hombre sobre la naturaleza ..."

Si bien la experimentación seguirá en el campo de la programación, es posible que se desarrollen nuevos principios de almacenamiento (como las memorias holográficas, por ejemplo), que la computación híbrida llegue a desempeñar un papel fundamental (para la resolución de una ecuación diferencial, por ejemplo, la unidad de control podría encargarse de esa tarea a un dispositivo analógico específico), o que se desarrolle un nuevo tipo de estructuración lógica (*computación paralela*, en la cual una máquina "central" ordena la realización de los diversos procesos de una tarea a otros centros "esclavos").

La ciencia-ficción se ha dedicado mucho al tema de las máquinas que dominen al hombre; un ser, en principio, mucho más limitado. La posibilidad de máquinas que *sientan* no es quizás tan remota si se aceptan ciertos "sentimientos" como simples sensaciones físicas y se reconoce la posibilidad de una ver-

dadera "percepción" por parte de las máquinas. Se plantea entonces la posibilidad de una "clase social" compuesta por máquinas pensantes que podría llegar a oponerse al hombre. Sin embargo, no creemos en tal alternativa, fundamentalmente porque habrá sido el mismo *hombre* el que les dio la "vida", y es claro que en tal caso habrá tomado las debidas precauciones para programar sus sentimientos de modo tal de convertirlas en sumisos esclavos. Los secretos laboratorios donde un maquiavélico hombre prepara un cerebro que le ayudará a dominar al resto de la humanidad, pertenecen exclusivamente al dominio de la fantasía; en la realidad tales proyectos necesitan de equipos gigantescos de científicos; que si bien pueden estar al servicio de un Estado antes que de la Humanidad, es de esperar que la conciencia pacifista que aparentemente caracterizará al hombre del futuro, canalice el uso de tan poderosos mecanismos hacia fines compatibles con los de la Humanidad.

En síntesis, una profunda confianza en los valores racionales y éticos del hombre puede inspirarnos tranquilidad y profunda fe en el futuro.

Este libro se terminó de imprimir en
"ARTES GRÁFICAS PROLAM"
Miró 1540, Buenos Aires, Argentina, a
los 15 días del mes de junio de 1972