

## La intel·ligència artificial: des dels orígens fins al futur

Aquest article defensa la tesi que és possible construir màquines intel·ligents. Aquesta afirmació es fonamenta en una anàlisi històrica dels esforços realitzats per la humanitat per tal de fer realitat aquest antic somni. L'aparició de l'ordinador va impulsar

el naixement de l'anomenada intel·ligència artificial que es perfila com l'esforç definitiu per construir aquestes màquines intel·ligents. Els més optimistes pensen que a finals d'aquest segle veurem culminar aquest esforç; d'altres diuen que n'hi ha per més

temps. Pels especialistes en informàtica el problema no és doncs si pot aconseguir-se sinó *quan* s'aconseguirà. Aquest fet comportarà un estil de vida tan diferent que sembla difícil per no dir impossible d'imaginar.

### Introducció

Què passaria si els ordinadors fossin tan intel·ligents com per exemple l'ordinador HAL de la pel·lícula de Stanley Kubrick *2001: A Space Odyssey*? Imaginem un ordinador capaç de pensar, raonar, fer deduccions lògiques, recordar experiències, resoldre problemes, comprendre el llenguatge parlat i escrit, etc.. Difícilment podríem refusar de qualificar-lo d'intel·ligent i hauríem d'acceptar que "Pensa, per tant existeix" i aquesta frase de Descartes deixaria de ser monopoli dels éssers humans. Hauríem d'aprendre a conviure amb una altra espècie intel·ligent, cosa que ens obligaria a replantejar la nostra pròpia existència, amb tots els avantatges i inconvenients que això comportaria. Tot això sembla utòpic i poc seriós i, tanmateix, la creació d'intel·ligència és un dels somnis més antics de la humanitat. L'aparició de l'ordinador va constituir una ajuda molt important en els esforços dirigits a fer realitat aquest somni, i, en particular, dins de la informàtica, aquests esforços configuren una de les àrees de recerca més importants: la INTEL·LIGÈNCIA ARTIFICIAL. Els governs dels països més avançats des del punt de vista científic i tecnològic estan "abocant el sac" per subvencionar projectes relacionats amb la intel·ligència artificial; cal destacar, per exemple, el gegantí projecte japonès anomenat "5th Generation Computer System" coordinat pel Ministeri d'Indústria i Comerç exterior d'aquell país i en el qual participen pràcticament totes les universitats tècniques i científiques, així com les indústries informàtiques més importants del Japó. Aquest projecte s'ha fixat la fita de construir l'ordinador intel·ligent a principis dels anys 90. La iniciativa japonesa ha tingut un gran ressò arreu del món i països com

Estats Units, França, Alemanya i Anglaterra ja estan elaborant projectes similars.

L'opinió generalitzada entre els especialistes en informàtica és que la qüestió no és saber si projectes com el japonès tenen probabilitats de succeir, (d'això ja en sabem que la resposta), sinó *quan*; i és que la història de la ciència ens ensenya que allò que és especulatiu per una generació esdevé probable en una altra de posterior i un fet en la següent. Una prova recent la tenim en els ràpids avenços deguts a l'electrònica en tots els seus aspectes (el telèfon, la televisió el mateix ordinador...). Afortunadament, els informàtics no estem sols en la realització de la tasca de construir màquines intel·ligents, ja que una important quantitat de filòsofs, matemàtics, lingüistes i psicòlegs hi col·laboren d'una manera determinant. El reconeixement d'aquest esforç comú ha quedat consolidat amb l'aparició de l'anomenada *ciència cognitiva* que agrupa les ciències esmentades. Cal remarcar que a la Universitat Politècnica de Barcelona hem constituït un grup de treball sobre ciència cognitiva i que tenim estrets contactes amb grups de treball similars a França, Alemanya i als Estats Units.

En aquest article voldria donar una visió de l'estat de l'art de la intel·ligència artificial a través d'un repàs històric des de l'origen fins avui dia, posant l'èmfasi en els últims, els quals, al cap i a la fi, seran els que determinaran el futur.

### Des dels orígens a la cibernètica

La intel·ligència artificial té la fortuna de tenir una de les més riques i divertides històries en la ciència; possiblement la raó sigui que investiga una qualitat tan profunda en l'esperit humà com és la intel·ligència.

D'altra banda, comprendre el desenvolupament històric d'una àrea de la ciència és sovint una pre-condició per poder aportar contribucions creatives a aquesta àrea. El període clau d'aquest desenvolupament se situa de principis a mitjan anys cinquanta, que va ser quan l'art es va transformar en ciència. El catalitzador d'aquesta transformació va ser el reconeixement que l'ordinador era el mitjà més prometedor per la realitzar un antic somni de la humanitat: *la creació d'intel·ligència*.

El somni és certament antic ja que algunes de les seves primeres manifestacions apareixen en la mitologia grega quan Homer, als voltants de 850 anys aC, ens parlava d'autòmats que executaven aquelles tasques que resultaven pesades per als déus. Vint segles més tard, el místic català Ramon Llull va aprendre àrab i se'n va anar a terres llunyanes amb la idea de convertir els musulmans al cristianisme. Es desconeix si el seu intent va tenir o no efecte entre els musulmans; però el que és segur és l'efecte que ells li van provocar ja que va quedar tan impressionat per la grandiositat d'una màquina de pensar àrab anomenada *Zairja* que va tornar immediatament al món cristià amb la idea de construir una màquina pensant anomenada *Ars Magna* la finalitat de la qual era arribar a la veritat mitjançant un raonament mecanitzat i, per tant, sense la molèstia de pensar. Sigui com sigui, l'esquema de Llull és remarcable, no per la finalitat que pretenia sinó perquè presuposava, sense vacil·lacions, que el pensament humà podia ser mecanitzat. Max von Bohlen en la seva obra *Puppets and Automata* ens parla d'estàtues parlants de bronze com, per exemple, les que posseïen Albert el Magne i el papa Silvestre II. A partir del segle XVI, els autòmats van començar a proliferar, un dels més coneguts és l'"ànc" de Jacques de Vau-



per Ramon López de Màntaras

Ramon López de Màntaras i Badia (Sant Vicenç de Castellet, Bages, 1952) és doctor en automàtica (robòtica) per la Universitat Paul Sabatier de Tolosa de Llenguadoc, Master of Science en enginyeria (informàtica) per la Universitat de Califòrnia a Berkeley i doctor en informàtica per la Facultat d'Informàtica de la Universitat Politècnica de Barcelona. Ha sigut professor ajudant a

la Universitat Paul Sabatier de Tolosa, Research Engineer a l'Electronics Research Laboratory de la Universitat de Berkeley i investigador convidat al Max Plack Institute de Munic amb una beca de viatge de la CIRIT. Ha publicat més de vint treballs relacionats amb la intel·ligència artificial. És editor associat de dues revistes internacionals relacionades amb aquesta qüestió i

membre de la junta directiva de l'Associació Internacional pel Reconeixement de Formes. Actualment és professor adjunt a la Facultat d'Informàtica de la UPB i coordinador del grup de treball sobre ciència cognitiva d'aquesta universitat. Premi Ciutat de Barcelona d'investigació l'any 1982.

ciència 31

octubre 1983 / Volum 3 / 601 17

canson que simulava el procés digestiu dels ànecs. Entre els autòmats construïts realment i els purament imaginaris van aparèixer també molts fraus, el primer i un dels més famosos és la màquina de jugar escacs de Von Kempelen, que data del 1770. Von Kempelen era un il·lusionista que va sorprendre Europa amb un androide capaç de jugar als escacs de manera excel·lent. Evidentment, era un truc ja que hi havia un ésser humà de talla reduïda amagat dins la màquina.

De tot l'exposat, sembla desprendre's que fins molt recentment tot el relacionat amb la intel·ligència artificial pertanyia a la llegenda i a la fantasia.

Si considerem l'especulació científica com a fantasia, l'anterior conclusió és probablement certa, però si per especulació científica entenem no solament aquells somnis que pretenem realitzar sinó també aquells possibles mitjans per realitzar-los, llavors estem davant a un plantejament molt diferent. Un exemple clar d'aquest plantejament ens el dona Charles Babbage i la seva col·lega la comtessa Ada Lovelace. El 1843, Lady Lovelace va publicar una llarga i detallada descripció del calculador analític de Babbage i, contràriament a les seves afirmacions anteriors, en les quals deia que una màquina tan sols pot fer el que expressament se l'ha programat fer, va afegir que la qüestió sobre si una màquina podria o no pensar romandria incontestada fins que es construís i la provessin. De totes maneres, Babbage i Lady Lovelace van considerar construir una màquina per jugar a escacs per poder finançar la construcció del calculador analític. Molt més tard, el 1915, l'inventor espanyol Leonardo Torres Quevedo va construir dues màquines capaces de jugar als escacs simplificats com, per exemple, rei i torre contra rei. Torres Quevedo sempre es va negar a afirmar que els seus autòmats pensessin

realment, encara que va suggerir que seria millor matisar la definició del procés pensant ja que els seus autòmats podien certament portar a terme funcions que eren considerades popularment com intel·ligents. En realitat, els autòmats de Torres Quevedo eren servomecanismes i van influir de manera important en la futura cibernètica de Norbert Wiener. La idea de Babbage de construir una màquina de calcular no era una idea nova: a part de l'existència de l'Abacus del 1642, Blaise Pascal va construir el primer calculador digital (aquí digital significa una màquina amb un conjunt finit d'estats). Leibnitz també va inventar una màquina de calcular i va escriure: *it is unworthy of excellent men to loose hours like slaves in the labor of calculation which could safely be relegated to anyone else if machines were used.*

Leibnitz mai no va atribuir capacitats pensants a la seva màquina; de fet, creia en el dualisme cartesí que considerava que cos i esperit són dues coses distintes i que, per tant, els actes humans es divideixen en dues classes: mecànics i racionals. Obviament, els mecànics podien imitar-se i els racionals no. Malgrat tot, els seus viatges freqüents per Europa i les converses amb científics sobre la seva màquina de calcular van fer que Leibnitz sentís la necessitat d'un llenguatge comú entre científics per tal de facilitar els intercanvis d'idees i va tenir la idea de reduir el raonament a una àlgebra del pensament anomenada *calculus ratiocinator*. Aquesta idea va tenir la seva eclosió en el segle XIX amb George Boole, gràcies al seu importantíssim treball, que va posar les bases de la moderna lògica simbòlica, titulat *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities*.

El treball iniciat per Boole es va anar desenvolupant i millorant, i va culminar amb el *Principia Mathematica* de Whi-

tehead i Russell, que al meu entendre constitueix una fita no solament en la història de la matemàtica i de la filosofia sinó també del pensament humà en general.

### Alan Turing, pare de la intel·ligència artificial

Tanmateix el pare de la intel·ligència artificial va ser sens dubte el matemàtic Alan Turing. L'any 1937, va proposar la màquina de Turing i el 1947 va escriure un treball titulat *Intelligent Machinery*, publicat... el 1969!, en el qual comença donant una sèrie d'arguments contra les màquines intel·ligents (entre altres coses afirma que el teorema de Gödel demostra que les màquines són inherentment incapaces de resoldre problemes que no puguin ser resolts per éssers humans) per, finalment, refutar tots els arguments donats i acabar conjecturant sobre la possibilitat de construir màquines intel·ligents. L'any 1950, Turing va publicar el seu conegut treball *Computing Machinery and Intelligence* en el qual va proposar el conegut test que porta el seu nom.

Com és natural, Turing va tenir molts detractors; un dels que no compartien el seu optimisme sobre la possibilitat de fabricar màquines intel·ligents va ser l'extraordinari matemàtic John von Neumann.

Von Neumann estava, tanmateix, fascinat per la idea sobre la qual havia discutit llargament amb Turing, però no veia possible la connexió entre el pensament humà i la màquina. En un article publicat el 1951 va enumerar les diferències entre el sistema nerviós i l'ordinador, i, tanmateix, ell mateix, sis anys abans havia proposat termes com: *memòria* i *òrgans de control* per designar algunes de les





Norbert Wiener (1894-1964), el principal introductor del pensament cibernètic.



funcions d'un nou ordinador que estava dissenyant.

Segons Von Neuman, la raó més important que impedia fabricar ordinadors intel·ligents era la falta d'una teoria lògica d'autòmats. La falta d'aquesta teoria impedia construir màquines d'elevada complexitat i afirmava que un elevat nivell de complexitat era essencial per produir comportaments intel·ligents. En aquest mateix article, menciona el treball de McCulloch, un neurofisiòleg, i Pitts, un matemàtic, que el 1943 van publicar: *A Logical Calculus of the Ideas Immanent in Nervous Activity*, treball que clarament va impressionar Von Neumann, ja que el va utilitzar per professar sobre la "teoria de màquines de calcular", ignorant la increïble complexitat del funcionament de les neurones i demostrant que una noció simplificada d'aquest funcionament podria ser imitada mitjançant relès o tubs de buit.

El 1943 es va publicar un altre treball important d'Artur Rosenblueth, Norbert Wiener i Julian Bigelow titulat *Behaviour, Purpose and Teleology* que, el mateix que el de McCulloch i Pitts, proposava un model per al tractament de la informació que proporcionava una major flexibilitat per fer hipòtesis sobre el raonament. La primera frase d'aquest article reflecteix les idees de Leibnitz, de Whitehead i Russell, i de Rudolf Carnap: *Because of the 'all-or-none' character of nervous activity, neural events and the relations among them can be treated by means of propositional logic*". A continuació, l'article descriu un càlcul lògic així com els principis per construir una classe de màquines que permetrien incorporar qualsevol teoria del pensament o del comportament, sempre que fossin satisfets alguns principis molt generals de finitud i causalitat. En opinió meua, els articles de Rosenblueth-Wiener, Bigelow i McCulloch-Pitts, introdueixen d'una manera tan clara l'estructura del

pensament que la seva publicació podria considerar-se com el naixement de la cibernètica, malgrat que el mateix Wiener no va introduir aquest terme fins al 1947, és a dir, quatre anys més tard.

### La cibernètica: de l'energia a la informació

La cibernètica, el sant patró de la qual, segons paraules de Wiener, seria Leibnitz, va significar reemplaçar el model basat en el concepte central de la mecànica newtoniana, l'energia, per un altre model basat en el concepte d'informació. Les idees i conceptes aportats per Shannon, en la teoria de la informació, com, per exemple, codificació, emmagatzematge, etc., proporcionaven una explicació millor per a moltes situacions que abastaven des del comportament de circuits electrònics fins al comportament de la divisió cel·lular. La principal raó era que l'antiga mecànica newtoniana s'aplicava a sistemes tancats i conservatius, mentre que el model de la teoria de la informació es podia aplicar a sistemes oberts, és a dir, sistemes acoblats al món exterior tant per actuar com a receptors d'informació com per executar accions. S'ha d'assenyalar que el principal corrent d'investigació en psicologia d'aquella època, encara molt arrelat en el model newtonià, va oposar molta resistència davant el canvi tan dràstic de paradigmes.

La cibernètica, gràcies a homes com Wiener i McCulloch, ha exercit una gran influència en la moderna intel·ligència artificial. McCulloch, en el seu treball amb Pitts citat anteriorment, va proporcionar una nova definició del concepte de màquina, afirmant que tot allò que pot ser descrit completament, pot ser realitzat mitjançant una xarxa de neurones, el que permetia considerar el

cervell constituït com una màquina però d'una manera més precisa que els models anteriors. En certa manera, el seu model de xarxa neuronal podia comparar-se amb una màquina de Turing.

Els treballs de McCulloch-Pitts i de Rosenblueth-Wiener-Bigelow impliquen que les lleis que governen l'activitat mental es troben més aviat entre les lleis que governen la informació que entre les que governen l'energia o la matèria.

Va ser precisament llavors que va aparèixer el primer ordinador, que, malgrat els americans i anglesos, va ser construït per un enginyer alemany anomenat Konrad Zuse, segons s'ha pogut comprovar mitjançant els documents corresponents a les sol·licituds de patents per part de Zuse, ja que els ordinadors que va construir van ser destruïts durant la segona guerra mundial. S'ha d'assenyalar que Konrad Zuse també es va deixar fascinar per la idea de les màquines intel·ligents i que, ja el 1943, es preguntava si el seu ordinador podria jugar als escacs. Per si fos poc, el 1945 ja havia desenvolupat un llenguatge de programació anomenat "Plankalkül" que, segons ell, podia ser utilitzat no solament per efectuar càlculs numèrics sinó també per resoldre problemes de tipus simbòlic, com, per exemple, moviments en els escacs. Malgrat tot, ell creia que la intel·ligència artificial estava encara a una o dues generacions vista. Aïllat per la derrota del seu país i per les prohibicions de desenvolupar l'electrònica durant la postguerra, Zuse no podia treure fruit del seu treball pioner i, quan a mitjan anys 50, va tenir la notícia dels desenvolupaments aconseguits als Estats Units en el camp de la intel·ligència artificial, es va quedar horroritzat per la seva falta de preocupació per les conseqüències del que estava fent. "They were like children playing with matches" va dir el 1976 i va afegir: "I was shocked. I see no reason why machines cannot think, and





Els robots d'Asimov es lligaven amb una imatgeria típicament industrial.

*it wouldn't surprise me if they someday out-think humans. But...*"

Konrad Zuse no va ser, evidentment, l'únic a adonar-se de les noves possibilitats que oferia l'ordinador. No és, doncs, sorprenent que comencés a anomenar-se màquina pensant. En la llarga història d'autòmats i màquines, tant reals com producte de la fantasia, construïts amb el propòsit que fossin capaces de pensar, cap, no havia demostrat ser més prometedor que l'ordinador per imitar, i potser ajudar a comprendre millor, la ment humana.

Malgrat la seva importància, el model neuronal de McCulloch i Pitts era enganyosament simple. Ells mateixos afirmaven que el coneixement és complex, les neurones del cervell relativament simples (encara que no tant com deixava entreveure la neurofisiologia d'aquells temps) i les interaccions entre les dues coses era el que quedava per dilucidar. Els neurofisiòlegs d'avui dia dirien que els processos del coneixement són complexos, que la neurona sembla molt més complicada del que es creia i que comprendre les connexions entre les dues coses podria molt bé ser un problema impossible de resoldre.

De fet, la investigació neurofisiològica ha demostrat que "la neurona és un sistema complex de mecanismes elèctrics i químics amb interaccions no lineals i que combinen processos variables, tant discrets com continus". De tota manera, l'excessiva simplificació de la neurona es va deure també a la manca d'eines matemàtiques capaces de tractar models més complexos. Ningú no era més conscient d'aquestes limitacions que el mateix McCulloch, que va passar la resta de la seva vida intentant de trobar una teoria fisiològica del coneixement, una cosa així com la descripció que feia Marvin Minsky del cervell com una màquina de carn (*meat machine*).

El canvi de paradigma de l'energia a la

informació també va tenir lloc, de forma independent, a Anglaterra el 1949, on un grup de físics, enginyers i fisiòlegs començaren a treballar junts sobre l'estudi del cervell i del tractament de la informació en animals i màquines (cibernètica, en una paraula). Un dels científics més coneguts d'aquest grup és el físic Donald Mackay; segons ell, la lògica tot-res era nefasta ja que no podia modelitzar els senyals de tipus gradual que, segons semblava, transportaven els productes químics del cervell. Per tant, va intentar imaginar una màquina amb els avantatges dels ordinadors digitals (precisió) i l'avantatge de les màquines analògiques de poder representar gradacions. Mackay afirmava que les dades utilitzades en el pensament humà no eren exactes i que, per tant, no es podien reduir a proposicions del tipus sí/no i a voler introduir la idea de veritat parcial. Aquesta idea apareixia més tard, en els anys 60, i per exactament les mateixes raons a què al·ludia Mackay, en els treballs de Lotfi Zadeh sobre l'anomenada lògica difusa.

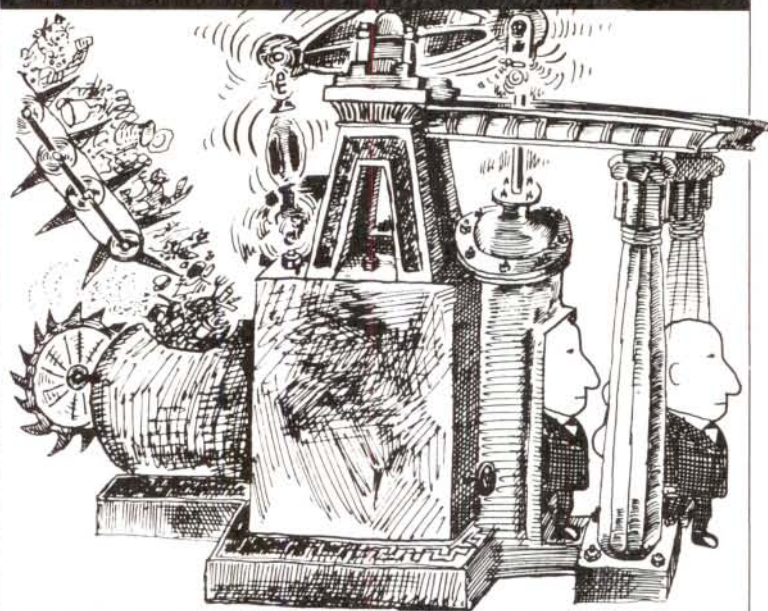
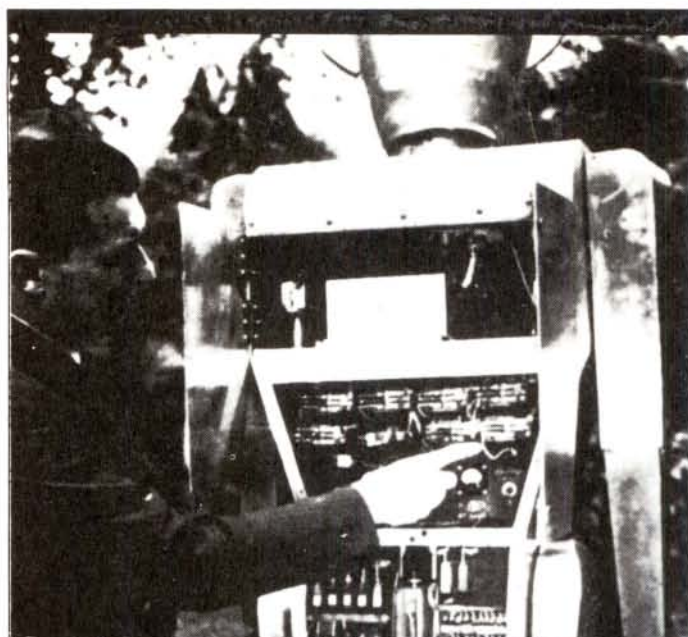
El 1949, Mackay va fer circular un document en el qual plantejava el problema de combinar tècniques digitals i analògiques en una màquina completament autònoma, és a dir, l'existència de la qual seria independent de tota interacció humana, essent capaç d'ajustar el seu comportament en funció de la informació que aniria rebent, amb la finalitat d'arribar a un equilibri, és a dir, que seria capaç d'aprendre.

El 1948, un altre membre d'aquest grup, el psiquiatre W. Ross Ashby, va publicar un article que més tard es va convertir en un llibre titulat *Design for a Brain* (1952), en el qual proposava uns mitjans per imitar l'habilitat del cervell per produir comportament adaptatiu, dit d'una altra manera, capaç d'aprendre. El model d'Ashby recorda els principis cibernètics enunciats per Wiener: "*The free*

*living organism and its environment, taken together, form an absolute system... The two parts act and re-act on one another*" (Ashby, 1952). Malgrat tot, Ashby va fins i tot més lluny que Wiener en introduir conceptes com el de l'estabilitat (equilibri per Mackay), és a dir, una espècie de supervivència en l'organisme. Un pasatge clau del seu llibre aclareix aquesta idea: *A determinate 'machine' changes from a form that produces chaotic, unadapted behaviour to a form in which the parts are so coordinated that the whole is stable, acting to maintain certain variables within certain limits-how can this happen?*. La resposta és que la màquina és un sistema autoorganitzatiu, que respon a estímuls, modificant el seu comportament, i fins i tot la seva forma, amb la finalitat d'arribar a l'estabilitat o el que Ashby va preferir anomenar ultraestabilitat. Un exemple clar té lloc quan queda danyada una part del cervell humà, ja que altres parts d'aquest cervell assumeixen responsabilitats per realitzar les funcions de la part danyada, particularment quan aquest dany té lloc en els primers anys de la vida d'una persona. En altres paraules, sembla que hi hagi un cert grau de redundància en les connexions entre neurones; a més a més, aquestes connexions sembla que estiguin fetes d'una forma més o menys aleatòria, de tal manera que els impulsos elèctrics poden desviar-se per altres camins quan un d'aquests camins està bloquejat o danyat. Aquest mètode va representar una enorme temptació per als investigadors, que van intentar construir sistemes *self-organizing*, que s'autoorganitzessin aleatòriament per tal d'aprendre i adaptar-se de la mateixa manera que el cervell sembla fer-ho.

Un dels esforços més importants en aquesta direcció va ser un sistema anomenat *Perceptron*, desenvolupat per un grup de la Universitat de Cornell dirigit per Frank Rosenblatt. El Perceptron





contenia tres nivells. El primer consistia en una reixeta de fotocèl·lules, que pretenien imitar la retina de l'ull, i que reaccionaven davant d'estímuls lluminosos. En el següent nivell hi havia unes unitats col·lectores dels impulsos procedents de les fotocèl·lules; aquestes unitats havien sigut connectades aleatòriament i, a la vegada, transmetien senyals a un tercer nivell format per unitats de resposta. Ja que tant els animals com els éssers humans neixen coneixent algunes coses, Roseblatt va modificar el Perceptron per tal que el connexionat no fos completament aleatori, la qual cosa millorava el seu funcionament en experiments de reconeixement i "aprenentatge" de caràcters impresos. Dit pel mateix Roseblatt, el Perceptron era capaç de coses extraordinàries i potser sigués cert, però ell mai no va poder demostrar-ho. Aquest fet va irritar molts dels seus col·legues, entre ells Marvin Minsky que, malgrat tot, va seguir treballant amb la idea del Perceptron, desenvolupant tota una teoria junt amb el seu col·lega Seymour Papert. Aquest esforç va quedar recollit en el seu llibre *Perceptrons* (Minsky-Papert, 1968).

L'esperança de construir una màquina pensant, mitjançant una imitació literal del cervell a nivell cel·lular, va acabar per esvaïr-se cap a mitjan anys 50. La idea no va funcionar tal com Von Neumann ho havia profetitzat.

En aquells temps, no es coneixia suficientment, ni tampoc es coneix ara, el cervell a nivell cel·lular; en qualsevol cas, s'ha començat a especular que cap màquina que tracti la informació en sèrie, no seria capaç d'imitar el que sembla ser un conjunt de processos treballant en paral·lel.

Però els investigadors en intel·ligència artificial tenien la certesa que existien altres formes més efectives de modelitzar o simular el pensament, i van ser aquest models alternatius els que arriba-

rien a donar lloc al canvi decisiu en la història de la intel·ligència artificial.

### El canvi decisiu: El model del tractament de la informació

L'ordinador digital anava a ser finalment la principal eina per simular els processos cognitius humans; però l'enfocament decisiu seria bastant diferent de l'utilitzat per la cibernètica. Aquest model decisiu s'anomena *tractament de la informació*, i malgrat que conté la paraula informació, és bastant diferent al model de la teoria de la informació, ja que la idea central és la manipulació de símbols i no la simple realimentació (*feedback*).

El tractament de la informació és un nivell intermedi de modelització que admet utilitzar tant expressions matemàtiques com expressions no matemàtiques, i permet formular simbòlicament moltes més hipòtesis sobre el comportament del cervell que els anteriors models i, al mateix temps, fa innecessària la comprensió del seu funcionament a nivell cel·lular. El mètode està admirablement adaptat (o fins i tot inspirat) per a l'experimentació amb l'ordinador, ja que considera l'ordinador com un processador d'informació i manipulador de símbols, un punt de vista aplicable també als éssers humans.

El model de tractament de la informació ha dominat tota la investigació en intel·ligència artificial des de mitjan anys 50 fins avui dia, i s'han obtingut resultats espectaculars i esperançadors.

L'èxit d'aquest model ha fet que hagi arribat a dominar altres àrees del saber, com, per exemple, la psicologia cognitiva, la lingüística i les ciències socials en general. El conjunt de totes aquestes ciències basades en el model del tractament de la informació constitueix el que

avui dia es coneix amb el nom de ciència cognitiva.

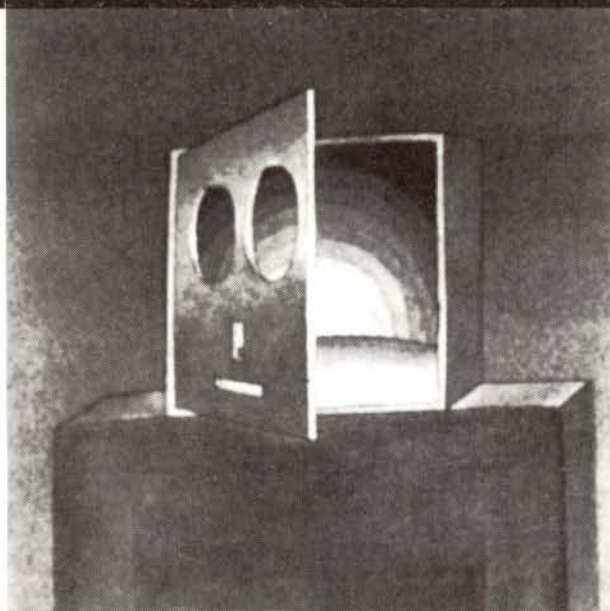
En el model del tractament de la informació es poden distingir tres períodes diferents en relació amb les teories utilitzades per representar coneixements. El primer d'aquests períodes podríem situar-lo des de mitjan anys 50 fins a principis dels 60 i jo l'anomenaria *període formatiu*.

### Període formatiu

Durant aquest període, es va pretendre representar el coneixement de manera molt tradicional, mitjançant proposicions del tipus subjecte-predicat, utilitzant el llenguatge que ens proporciona la lògica dels predicats. Els processos de raonament es simulaven mitjançant tècniques uniformes tals com la recerca heurística i demostració de teoremes. Aquestes tècniques van jugar un paper tan important que, per a molts informàtics amb excessiva tendència a la teoria, la intel·ligència artificial ha quedat identificada com la ciència de la recerca heurística, és a dir, l'estudi de tècniques per explorar espais d'estats massa amplis per ser explicats de manera exhaustiva.

El "Logic Theorist" (LT) de Newell i Simon (1956), juntament amb un programa per jugar a les dames (1955-1959) van ser els primers programes heurístics realitzats en un ordinador. El LT, que era un programa de resolució general de problemes ben definits (Torres d'Hanoi, Missioners i caníbals, etc...), també va significar un esdeveniment de gran importància en el camp de la psicologia, essent considerat la base de l'anomenada "Theory of Human Problem-Solving". De manera similar, un altre programa de Newell i Simon anomenat General Problem Solver (GPS,





Les màquines que cada cop saben jugar més a escacs...

1957-1959) va ser, durant anys, el tema central de tota la investigació en resolució de problemes d'I.A. El GPS va ser considerat el model més complex i detallat dels processos mentals humans que mai havia estat construït i provat.

A part del LT i del GPS, en una impressionant demostració de creativitat, Newell, Shaw i Simon van concebre els nous llenguatges de programació IPL1, IPL2, IPL3 i IPL4, millor adaptats a les exigències de sistemes processadors de símbols, com per exemple el LT, introduint el concepte de tractament de llistes que acabaria sent la base dels futurs llenguatges de programació per a intel·ligència artificial que van culminar en el famós LISP de John McCarthy (LISP es caracteritza, igual com la revista (ciència), per la presència de molts parèntesis).

Entre els estudiants de Newell i Simon que també van realitzar importants contribucions, podem citar Feigenbaum i el seu programa EPAM, que era un model d'aprenentatge verbal (1959), Felman, que va desenvolupar un model per a la formació d'hipòtesis en forma de decisions binàries, i Lindsay, que va desenvolupar el primer programa de comprensió del llenguatge natural anomenat SAD-SAM.

En aquell període, la investigació en intel·ligència artificial ja s'estenia fora de la universitat, tenint lloc també a la indústria. El focus industrial més important d'aquell temps era IBM, on Gelemter i el seu grup estaven desenvolupant un programa de demostració de teoremes de geometria i, paral·lelament, van crear un llenguatge de programació amb capacitat per al tractament de llistes basat en el llenguatge FORTRAN, anomenat FLPL (Fortran Processing Language, 1959). Per altra part, Samuel va continuar desenvolupant la seva màquina per jugar a les dames, afegint-li possibilitats d'aprenentatge, de tal ma-

nera que automillorava el seu joc ajustant els valors dels paràmetres d'una funció d'avaluació. Un altre científic d'IBM (i mestre en escacs) anomenat Bernstein va desenvolupar el que seria un programa influent per jugar a escacs ja que gairebé vint anys més tard Greenblatt, al MIT, va desenvolupar un programa basat en el de Bernstein que va ser classificat en la classe C, és a dir, una categoria superior a la dels principiants.

Més tard, el 1976 un programa desenvolupat per David Slate i Larry Atkin de la Northwestern University, anomenat Chess 4.5, va rebre una avaluació de 2070 punts per part de la Federació d'Escacs dels EUA, el que va suposar pertànyer a la categoria d'expert. Una versió millorada de Chess 4.5, anomenada Chess 4.7, és avui dia a tocar la categoria dels mestres. El paper que han jugat els escacs en la història de la intel·ligència artificial ha sigut impressionant, des de fraus com el de Von Kempelen, realitzacions pioneres com les de Torres Quevedo, importants estudis teòrics com els de Shannan el 1950, fins a programes com el Chess 4.7, que, en l'últim dels campionats mundials per a programes que se celebren cada tres anys, ja s'ha convertit en una versió millorada anomenada 4.9.

Una altra activitat molt important en aquella època, que ha marcat el naixement oficial de la intel·ligència artificial, va ser la Conferència de Dartmouth l'estiu del 1956, que va reunir deu homes durant dos mesos per estudiar les seves possibilitats. En aquesta conferència, es va utilitzar per primera vegada de manera oficial el terme *intel·ligència artificial*. Els deu homes eren: John McCarthy, Marvin Minsky, Claude Shannon, Allen Newell, Herbert Simon, Oliver Selfridge, Ray Solominoff, Nathaniel Rochester, Trenchard More i Arthur Samuel. Aquesta reunió d'experts va

tenir una gran ressonància publicitària; això junt amb el fet que Rochester, More i Samuel treballaven per a IBM va provocar una anècdota curiosa i és que IBM va témer que la idea de dotar d'intel·ligència els ordinadors era tan amenaçadora que els seus clients potser deixarien de comprar ordinadors!

També en aquella època (aproximadament el 1955) i en els laboratoris Lincoln, Selfridge i Dineen va desenvolupar el primer programa capaç de reconèixer caràcters impresos. Més tard el mateix Selfridge i el seu grup van desenvolupar un sistema de reconeixement denominat Pandemonium que exerciria una gran influència en el futur del reconeixement de formes.

El 1962, Slagle va estendre les idees de la recerca heurística amb el seu programa d'integració simbòlica anomenat SAINT.

Minsky va recopilar i va organitzar les idees desenvolupades durant els últims anys en el seu important treball *Steps toward Artificial Intelligence* i McCarthy, reflectint la inquietud de la comunitat científica que treballa en intel·ligència artificial sobre l'eficiència per desenvolupar programes complexos, va suggerir la idea que l'ordinador pot funcionar en temps compartits, és a dir, que pot ser utilitzat per diversos usuaris simultàniament.

El 1963, Feigenbaum i Feldman van editar una col·lecció d'importants treballs en un llibre titulat *Computers and Thought* que s'ha convertit en la referència clàssica que reflecteix l'activitat intel·lectual d'aquest període formatiu.

### El període modern

Podríem dir que el període modern s'estén des de mitjan anys 60 fins a principis dels 70. El període va comen-



→  
El robot mòbil *Shakey* de l'Stanford Research Institute.

22 ( 606 / Volum 3 / octubre 1983

çar amb un èmfasi en la utilització de la semàntica per representar coneixements (Quillian, 1968). El treball de Quillian va ser utilitzat i ampliat per Hendrix en els seus treballs sobre el llenguatge natural. Altres realitzacions que caracteritzen aquest període són el programa ANALOGY (Evans 1968), que resolva tests d'intel·ligència del tipus de buscar analogies entre figures geomètriques; el programa de Winston, per aprendre conceptes mitjançant exemples i contraexemples (Winston, 1970); i alguns sistemes de comprensió del llenguatge natural com, per exemple, SIR (Raphael, 1964); STUDENT (Bobrow, 1964) basat en el concepte "Question-Answering"; i ELIZA (Weizenbaum).

Al final d'aquest període es va començar a utilitzar la tècnica dels procediments (és a dir, programes) per representar el coneixement de tal manera que, quan es necessita algun coneixement en particular, el corresponent programa és executat. Mitjançant aquesta tècnica d'inclusió procedimental del coneixement, Winograd va desenvolupar el sistema SHRDLU (1971), que era capaç de conversar en anglès sobre els objectes geomètrics. SHRDLU va donar una empena definitiva a la investigació sobre la comprensió del llenguatge natural. Un altre tema d'investigació, que, en certa mesura, va veure la llum durant aquest període, va ser la robòtica, amb la construcció en el SRI (Stanford Research Institute) del robot mòbil SHAKEY. La investigació en robòtica és important ja que comporta diversos aspectes de la intel·ligència tals com reconeixement de formes, resolució de problemes, representació de la informació i comprensió del llenguatge natural i, per tant, representa un domini d'experimentació immillorable d'aquelles tècniques que es desenvolupin en tots i cada un dels aspectes mencionats.

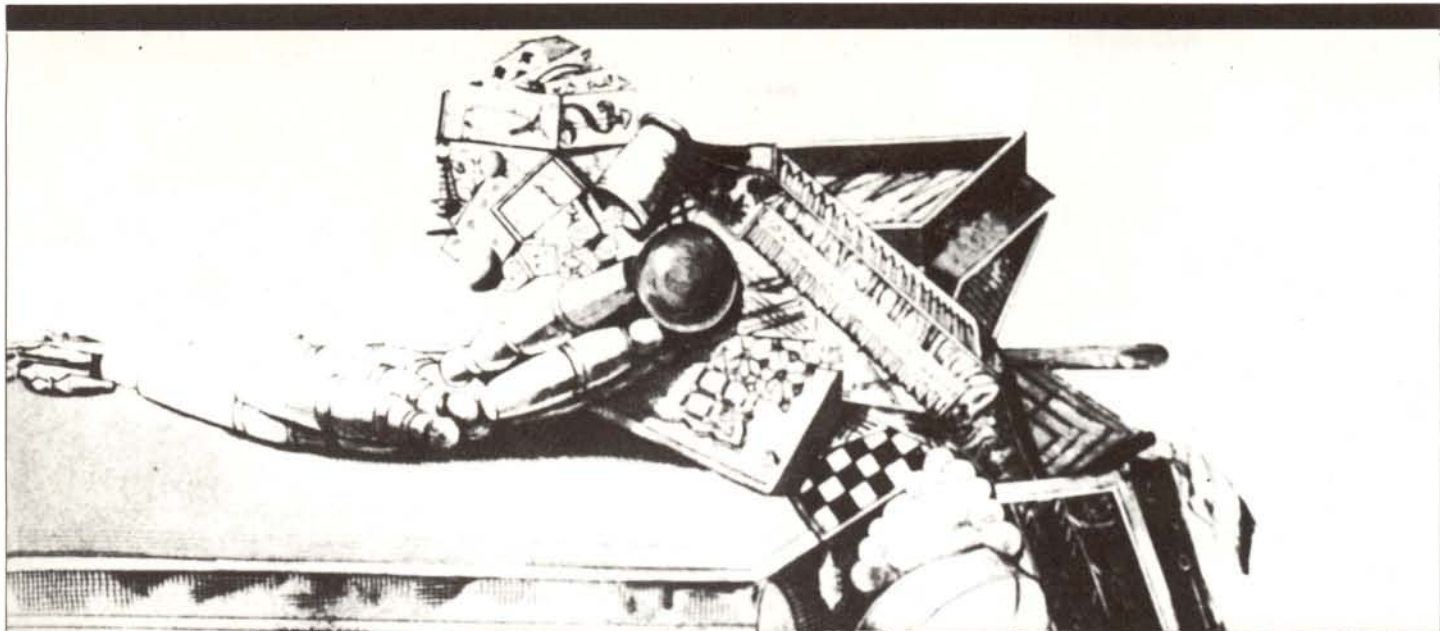
A partir de mitjan anys setanta, la in-



vestigació en robòtica general va disminuir per donar pas a la robòtica industrial, menys lligada a problemes d'intel·ligència artificial. Recentment, però, ha tornat a créixer l'interès per la robòtica intel·ligent, amb la construcció d'un robot mòbil en el Laboratori d'Automàtica i d'Anàlisi de Sistemes de Tolosa (França), anomenat HILARE, que ha significat un avenç important en robòtica general amb efectes secundaris també importants tant en intel·ligència artificial com en robòtica industrial. També en aquesta àrea, comptem a Barcelona amb un grup de treball sobre robòtica a l'Institut de Cibernètica i a la Facultat d'Informàtica. Juntament amb la robòtica, un altre fet important que caracteritza aquest període ha sigut l'aparició dels anomenats *Sistemes Experts*, basats en un conjunt de regles de producció per representar judicis d'experts sobre àrees del coneixement tals com medicina, química, teoria de circuits, matemàtiques, etc. Els primers sistemes van ser MACSYMA (Martin i Fateman, 1971), expert en matemàti-

ques; PROSPECTOR (Duda i Hart, 1978), expert en prospeccions minerals; i DENDRAL (Buchanan, 1969, 1971), expert en química (espectrometria de masses). DENDRAL és tan competent com un doctor en química a l'hora d'analitzar l'espectrograma d'una massa complicada i molècules orgàniques, amb l'avantatge de realitzar la feina en pocs minuts en lloc d'hores o dies. Cal destacar que DENDRAL està essent utilitzat quotidianament per químics de la Universitat de Stanford i ha donat lloc a més de 20 articles científics en revistes especialitzades en química. Aquest èmfasi cap a sistemes experts va portar a una presa de consciència de la importància que tenen els processos de recerca heurística i, sobretot, de les tècniques de representació del coneixement, arribant a la conclusió que el problema fonamental en intel·ligència artificial no era desenvolupar poderoses tècniques heurístiques sinó com representar grans quantitats de coneixements, de tal manera que sigui possible utilitzar-los eficaçment i de manera interactiva.





Aquest canvi d'enfocament es basava en una dècada d'experiències amb programes basats en tècniques heurístiques i lògiques que van resultar ser irreversiblement ineficaces davant de problemes que portaven amb ells mateixos un ampli espai de coneixements. El punt de vista adoptat actualment és que un sistema de resolució de problemes (ja sigui humà o artificial) ha de saber explícitament com utilitzar els seus coneixements mitjançant tècniques generals recolzades pel "Know-how" específic relatiu al problema que s'intenta resoldre.

La importància del paradigma basat en el *coneixement* pot comprendre's clarament des de la perspectiva evolucionista. Evidentment, la intel·ligència és impossible sense l'existència d'una certa capacitat bàsica a nivell de *hardware* (natural o artificial). En canvi, i d'acord amb Piaget, penso que aquesta capacitat ja està present a una edat molt primerenca, i el desenvolupament posterior de la intel·ligència no és degut a l'aparició de noves capacitats biològiques en l'individu sinó, més aviat, a l'adquisició de coneixements. Aquests coneixements poden estendre's des de fets específics en un determinat camp d'activitat fins a esquemes de resolució general de problemes o fins i tot pot tractar-se de coneixements sobre com aprendre nous coneixements. En alguns casos, el resultat és simplement una millora local de les capacitats de l'individu mentre que, en altres, els efectes poden ser globals, donant lloc a una etapa de transició, segons la terminologia de Piaget.

Des del punt de vista de l'ordinador, es va veure la necessitat de millorar la seva velocitat i capacitat de memòria, especialment per a aplicacions en visió i en reconeixement de la paraula; però les dificultats fonamentals a les quals ens enfrontem els investigadors en intel·ligència artificial no són degudes a les limitacions del *hardware* sinó a com re-

presentar grans quantitats de coneixement sense prejudicis per a l'ús eficaç de fets individuals.

### El període contemporani

**E**n aquest període s'ha consolidat el paradigma de coneixement i s'ha vist una tendència cap a la representació mitjançant unitats de coneixements més àmplies i més ben estructurades, com, per exemple, els "grafos conceptuals" de Schanck (1974) i sobretot mitjançant els "frames" de Minsky (1975).

Els frames constitueixen avui dia un dels millors mitjans de representar coneixements ja que, a més de permetre un tipus de representació declaratiu, han estat estesos per poder ser també procedimentals.

Una altra característica del període actual ha sigut el desenvolupament de la programació interactiva en línia, com, per exemple, UCI-LISP (1975), MACLISP (Moon, 1974) i INTERLISP (Teitelman, 1974), que ha estat la base dels llenguatges avançats de segona generació, com, per exemple, PLANNER, QAY, CONNIVER, SAIL, QLISP, KRL i fuzzy. Aquest darrer permet, a més, la representació d'informació imprecisa. Crec que la representació d'informació imprecisa és, precisament, una de les característiques més importants que hauran de posseir els sistemes intel·ligents del futur, ja que la major part de la informació que manipula l'ésser humà és de naturalesa imprecisa.

Un fet molt important d'aquest període ha estat la proliferació d'aplicacions gràcies als sistemes experts. En aquest moment (1983) es poden catalogar més de 40 sistemes experts en àrees tan diverses com les matemàtiques, la química, la medicina, l'ensenyament de l'electrònica, la geologia, la informàtica, l'engi-

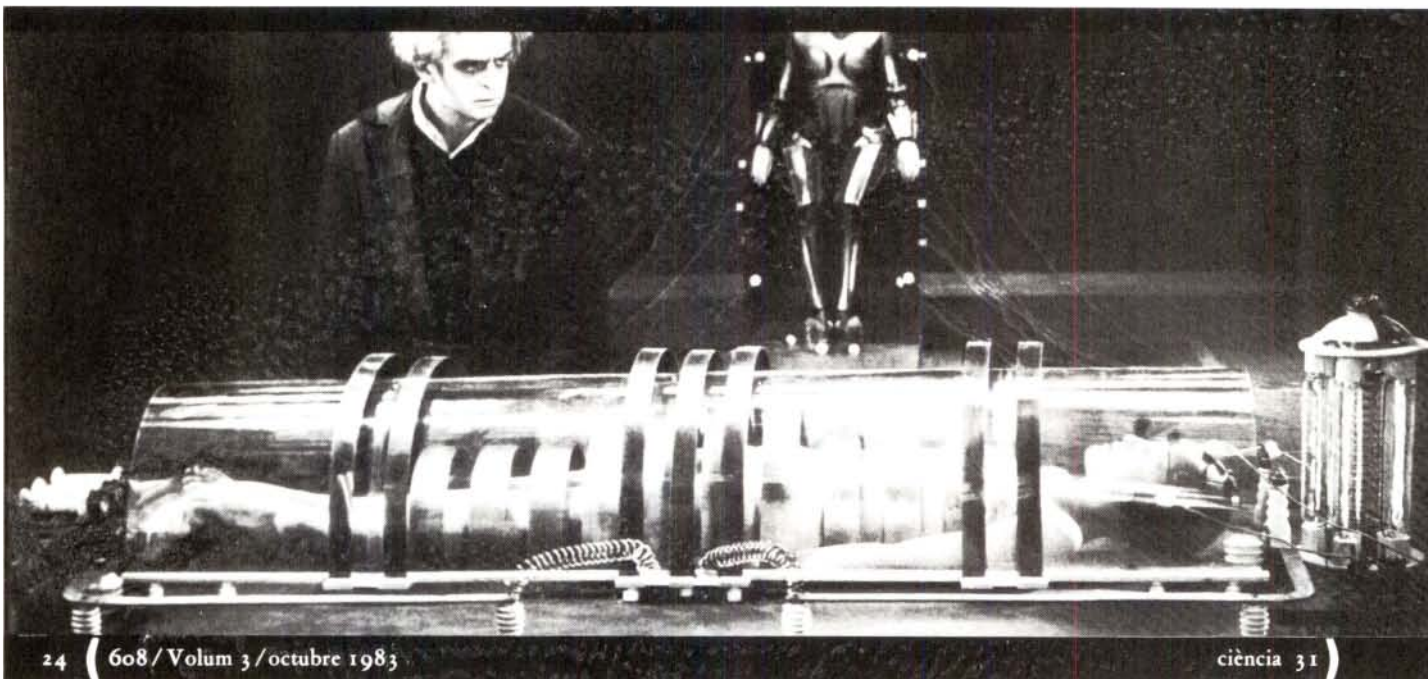
nyeria civil, etc.

La tendència actual és dotar els sistemes experts de la capacitat d'aprenentatge, és a dir, d'adquisició de coneixements.

El progrés en sistemes experts és una de les peces clau en el desenvolupament de la intel·ligència artificial, a més de les que ja hem mencionat, i s'ha de destacar el sistema anomenat INTERNIST desenvolupat a la Universitat de Pittsburg el 1977, expert en medicina interna. El sistema interroga el metge sobre els símptomes, historial mèdic, resultats d'anàlisi, etc., del pacient; el mateix que el metge, INTERNIST descarta alguns diagnòstics i n'explora altres més a fons, fins que, finalment, formula el seu diagnòstic explicant el raonament que ha utilitzat fins a arribar-hi. INTERNIST posseeix coneixements del nivell d'un expert sobre gairebé totes les malalties a les quals s'enfronta un especialista en medicina interna i és utilitzat cada cop més pels metges als Estats Units.

La comprensió del llenguatge (escrit i parlat) constitueix també un dels camps més actius de la recerca en intel·ligència artificial. A la Universitat de Yale han desenvolupat un programa anomenat FRUMP que llegeix històries curtes i és capaç de resumir-les en una frase gramaticalment perfecta. Tanmateix, i per allò que "és humà equivocar-se", els programes com el FRUMP cometen errors. Una mostra d'ells va tenir lloc quan FRUMP va haver de resumir una notícia sobre l'assassinat de l'alcalde de San Francisco el 1978, en la qual es deia, entre altres coses: "la mort que va fer trontollar San Francisco" i l'ordinador va arribar a la conclusió que a San Francisco hi havia hagut un terratrèmol. En una altra ocasió, la història explicava un segrest d'un missioner per terroristes filipins i, curiosament, FRUMP va resumir la història dues vegades, el que va portar-li a la conclusió que els terroristes tenen tendència a segrestar missioners!





### Conclusions

Donada la impressionant cadència en els desenvolupaments científics del nostre segle, és perfectament probable que en un futur la humanitat hagi "creat" màquines intel·ligents; el que no sabem és quan serà. Sembla evident, però, que en aquest futur la vida serà molt diferent de la d'ara. L'ordinador gestionarà la nostra pròpia informació filtrant aquelles dades que sabrà que no ens interessen i mantenint-nos alerta sobre el que ens interessa. Aquest ordinador podrà intervenir en les nostres decisions buscant en gegantines bases de coneixements aquells fets que ens ajudin a formar decisions, suggerint-nos línies d'acció i fent-nos veure les possibles conseqüències que es poden derivar d'elles. Utilitzarà coneixements de medicina, dret, etc., per actuar com a ajudant i conseller, controlarà robots que realitzaran treballs desagradables, peri-

llosos, etc. per a les persones, com, per exemple, recollir escombraries, treballar en mines, en atmosferes contaminades, etc. Tanmateix, les promeses de la intel·ligència artificial van molt més enllà, en el sentit que l'ordinador servirà per elevar l'intel·lecte i creativitat humans fins a nivells insospitats i actuarà, en definitiva, com un amplificador del poder mental tal com la màquina de vapor va constituir un "amplificador del poder físic". Però això no ho és tot; fins aquí he enumerat algunes conclusions optimistes i positives que constitueixen una cara de la moneda. A l'altra cara hi ha els perills, que si bé són difícils d'evitar —això seria impossible sense destruir la "moneda"—, hem de provar, almenys, de reduir-los en tot el possible. Pensem, per exemple, que es podria programar ordinadors capaços de comprendre la paraula escrita i parlada per escoltar totes i cada una de les converses humanes (recordem HAL, capaç de llegir els

llavis dels protagonistes de 2001) i decidir per ells mateixos si el contingut d'una conversa pot ser motiu de preocupació per part de les autoritats (l'any 1984 de Georges Orwell és a tocar). Un altre perill és que la humanitat depengui excessivament de les màquines i que siguem incapaços de comprendre el perquè de les decisions i accions que prenguin. Finalment, hi ha el temor de la següent pregunta: ¿podrà la màquina posseir, en alt grau, totes i cadascuna de les capacitats humanes: coneixements, intel·ligència, judicis de valor, capacitat de decisió, estètica i..., sobretot, emocions? Si la resposta és sí, significaria que, creant aquesta forma immortal d'intel·ligència, estem, sense voler, sembrant la llavor de la nostra pròpia extinció.

Ramon López de Mántaras

A *Metropolis*, Fritz Lang concebia una transferència d'intel·ligència d'una persona a un robot.

### Material de lectura

A. Barr i E.A. Feigenbaum (editors): *The Handbook of Artificial Intelligence*. Los Altos, Califòrnia, William Kaufmann Inc., 1981, 1982 (3 volums).

Obra de gran valor ja que dona una visió completa de l'estat actual de la intel·ligència artificial. És composta per 200 articles.

M. Boden: *Artificial Intelligence and Natural Man*. Nova York, Basic Books, 1977.

Un tractament del tema des del punt de vista de la filosofia i la sociologia. L'èmfasi se centra a remarcar la importància de la informàtica per ajudar-nos a comprendre millor la naturalesa humana.

H. Dreyfus: *What Computers Can't Do; A Critique of Artificial Reason*. Nova York, Harper & Row, 1972.

Una crítica ferotge de la intel·ligència artificial per part d'un filòsof que compara els esforços en aquest camp amb les follies dels alquimistes. Com era d'esperar, hi varen haver reaccions no menys ferotges per part dels grans noms de la

intel·ligència artificial.

D. Hofstadter: *Gödel, Escher, Bach: an External Golden Braid*. Nova York, Basic Books, 1979.

Discuteix el desenvolupament d'idees relacionades amb la formalització del raonament a través d'una interessantíssima anàlisi comparant els treballs de Gödel, els dibuixos d'Escher i la música de Bach.

R. López de Mántaras: *Resultados y tendencias en Inteligencia Artificial*. "Novática", Maig-juny, 1979

Una visió general del que és la intel·ligència artificial.

B. Raphael: *The Thinking Computer*. San Francisco, W.H. Freeman, 1976.

Una introducció a la intel·ligència artificial fàcil de llegir encara que una mica obsoleta.

D. Waltz: *Artificial Intelligence*. "Scientific American", octubre 1982.

Article molt interessant que va donar una bona idea del que és la intel·ligència artificial.

J. Weizenbaum: *Computer Power and Human Reason*. San Francisco, W.H. Freeman, 1976.

Una crítica a la intel·ligència artificial més rao-

nada i constructiva que la de Dreyfus. Weizenbaum era un dels pioners en aquest camp i va desenvolupar un programa (ELIZA) que simulava un psiquiatre, quan alguns dels seus col·laboradors varen començar a estendre la idea que aquest tipus de programa podria ser utilitzat com a teràpia tot i substituir els psiquiatres. La reacció de Weizenbaum va ser molt forta i el va conduir a escriure aquest llibre en el qual analitza els aspectes ètics i morals derivats de la substitució de l'home per la màquina.

Principals publicacions periòdiques sobre intel·ligència artificial:

"Artificial Intelligence", "Cognitive Science", "Computational Linguistics", "Sigart Newsletter", "AISB Newsletter", "AI Magazine", "Cognition", "International Journal of Man-Machine Studies", "Behavioral and Brain Sciences", "Transactions on Pattern Analysis and Machine Intelligence", "Transactions on Systems", "Man and Cybernetics", "Pattern Recognition", "Communications of ACM", "Cognition and Brain Theory", "Robotics Age", "Robotics Research" i "Fuzzy Sets and Systems".