

# Intel·ligència artificial

Miquel Barceló García

PID\_00153131



Universitat Oberta  
de Catalunya

[www.uoc.edu](http://www.uoc.edu)



# Índex

<b>Introducció</b> .....	7
<b>Objectius</b> .....	8
<b>1. Definició, objectius i altres denominacions</b> .....	9
1.1. El test de Turing .....	9
1.2. Objectius centrals de la investigació en intel·ligència artificial .....	10
1.3. Altres definicions .....	10
1.4. Dos enfocaments oposats .....	12
1.5. Altres denominacions .....	12
<b>2. Aplicacions de la IA</b> .....	14
2.1. Sistemes experts .....	14
2.2. Sistemes de visió per ordinador .....	14
2.3. Sistemes de llenguatge natural .....	15
2.4. Resolució de problemes .....	15
2.5. Robòtica .....	15
2.6. Aprenentatge .....	16
2.7. Jocs i puzles .....	16
<b>3. Tècniques de la IA</b> .....	17
3.1. Tècniques de base de la IA .....	17
3.2. Inferència simbòlica .....	17
3.3. Heurística .....	18
<b>4. La representació del coneixement</b> .....	19
4.1. Components i elements que s'han de representar .....	19
4.2. Etapes en l'ús del coneixement .....	20
4.3. Propietats d'un bon formalisme de representació del coneixement .....	20
4.4. Els espais d'estats .....	21
4.5. Esquemes moderns de representació del coneixement .....	21
4.5.1. Els esquemes lògics .....	21
4.5.2. Les xarxes semàntiques .....	22
4.5.3. Els esquemes de representació de procediments .....	23
4.5.4. Els <i>frames</i> .....	23
4.5.5. Els <i>scripts</i> .....	24
<b>5. Sistemes experts</b> .....	25
5.1. Definició i característiques .....	25

5.1.1.	Atributs d'un expert humà .....	25
5.1.2.	Definició i característiques d'un sistema expert .....	26
5.2.	Estructura bàsica d'un sistema expert .....	26
5.2.1.	La base de coneixements .....	27
5.2.2.	La base de fets .....	27
5.2.3.	El motor d'inferència .....	28
5.2.4.	Les interfícies .....	28
5.3.	Bases de coneixements: tipus de regles .....	29
5.3.1.	Regles antecedent-conseqüència .....	30
5.3.2.	Regles condició-acció .....	30
5.3.3.	Regles inexactes .....	30
5.4.	Motor d'inferència: funcionament i mecanismes de raonament .....	31
5.4.1.	Raonament per encadenament cap endavant .....	31
5.4.2.	Raonament per encadenament cap enrere .....	32
5.4.3.	Raonament inexacte .....	33
5.5.	Enginyeria del coneixement .....	33
5.5.1.	Funcions de l'enginyer del coneixement .....	34
5.5.2.	<i>Shells</i> i conjunts d'eines ( <i>toolkits</i> ) .....	35
5.6.	Viabilitat i beneficis d'un sistema expert .....	35
5.6.1.	Requisits de viabilitat d'un sistema expert .....	36
5.6.2.	Beneficis que s'han d'obtenir d'un sistema expert .....	36
5.7.	Problemes en l'ús de sistemes experts .....	37
5.8.	Alguns sistemes experts famosos .....	38
5.8.1.	MYCIN .....	38
5.8.2.	MACSYMA .....	38
5.8.3.	PROSPECTOR .....	38
5.8.4.	DENDRAL .....	39
5.8.5.	XCON .....	39
5.8.6.	Altres aplicacions .....	39
<b>6.</b>	<b>Xarxes neuronals</b> .....	40
6.1.	Descripció .....	40
6.1.1.	Models de xarxes neuronals artificials .....	40
6.1.2.	Algunes característiques de les xarxes neuronals .....	41
6.1.3.	Altres denominacions .....	41
6.1.4.	Aplicacions i funcionament .....	42
6.1.5.	Neuroordinadors i <i>netware</i> .....	42
6.1.6.	Característiques particulars de les xarxes neuronals .....	43
6.2.	Arquitectura i topologia de les xarxes neuronals .....	44
6.2.1.	Estructura topològica .....	44
6.2.2.	Estructura en capes d'una xarxa neuronal .....	45
6.2.3.	Connexionat de neurones .....	47
6.3.	Dinàmica i aprenentatge d'una xarxa neuronal .....	48
6.3.1.	Fase d'execució .....	48
6.3.2.	Fase d'aprenentatge .....	50
6.4.	Història de les xarxes neuronals .....	51

6.4.1.	El <i>perceptron</i> de F. Rosenblatt .....	51
6.4.2.	El predictor meteorològic de Widrow .....	52
6.4.3.	El nou interès per les xarxes neuronals .....	52
6.5.	Les xarxes neuronals i el cervell humà .....	53
<b>7.</b>	<b>Llenguatges específics per a la IA.....</b>	<b>55</b>
7.1.	Lisp .....	55
7.1.1.	Origen .....	55
7.1.2.	Objectius i característiques principals .....	55
7.1.3.	Dialectes .....	56
7.2.	Prolog .....	56
7.2.1.	Orígens .....	56
7.2.2.	Objectius i característiques principals .....	57
7.2.3.	Dialectes .....	57
7.3.	Smalltalk .....	57
	<b>Activitats.....</b>	<b>59</b>
	<b>Exercicis d'autoavaluació.....</b>	<b>59</b>
	<b>Solucionari.....</b>	<b>60</b>
	<b>Bibliografia.....</b>	<b>61</b>



## Introducció

La intel·ligència artificial és, amb tota seguretat, l'intent més ambiciós dels que aborda la informàtica. Superat l'estadi de construir màquines més fortes i ràpides que el seu propi constructor, a l'ésser humà encara li queda el repte de construir màquines que almenys semblin intel·ligents, com se suposa que ho som nosaltres.

Malgrat que pugui semblar un camp emparentat amb la ciència ficció, la realitat és que avui en dia disposem de programes que superen els humans en alguns aspectes del que considerem activitat en certa manera intel·ligent.

La intel·ligència és difícil de definir, i per això ningú no pretén construir programes intel·ligents sinó, simplement, programes amb un comportament que qualificaríem d'intel·ligent en un ésser humà. Una mica semblant a aconseguir moure's per les aigües marines com fa, per exemple, un submarí, però que pocs s'atrevirien a etiquetar de "nedar" el que fa un submarí...

Aquest mòdul presenta els **problemes centrals** de la intel·ligència artificial i les seves **tècniques particulars**, i explica amb més detall alguns casos concrets com ara els resultats obtinguts amb els sistemes experts i el renaixement de la tecnologia de les xarxes neuronals.

### Intents de la intel·ligència artificial

Després d'intents una mica esotèrics com els *homuncles* de Paracels o el *golem* que obeeix el rabí de Praga per l'efecte d'una paraula màgica, la intel·ligència artificial intenta portar a la realitat tecnològica la construcció de màquines i robots que es comportin intel·ligentment.

## Objectius

Els objectius que l'estudiant pot assolir són els següents:

- 1.** Definir i delimitar l'ampli camp de la IA, indicant les seves principals aplicacions.
- 2.** Introduir-se en l'estudi de les principals tècniques que utilitza la IA.
- 3.** Assumir la problemàtica de com es representa el coneixement i les diferents propostes en curs.
- 4.** Conèixer amb un cert detall les possibilitats dels sistemes experts, les tècniques específiques i els exemples més característics.
- 5.** Comentar el renaixement de la tècnica de les xarxes neuronals i exposar-ne les principals característiques.
- 6.** Saber quins són els llenguatges específics utilitzats en la IA.



## 1. Definició, objectius i altres denominacions

És força difícil definir exactament el contingut i l'abast de la intel·ligència artificial (IA), malgrat els llargs anys transcorreguts des que John McCarthy n'encunyés el nom l'any 1956 a la Darmouth Conference a Hannover (New Hampshire, EUA).

Potser una de les caracteritzacions més breus i senzilles és aquella que, parafraçant Marvin Minsky, un dels experts i investigadors més famosos de la IA, assigna a la intel·ligència artificial la "realització de sistemes informàtics amb un comportament que en l'ésser humà qualificaríem d'intel·ligent".

### 1.1. El test de Turing

A aquest efecte, és famós l'anomenat **test de Turing**, establert per Alan Turing en un article que data de 1950 ("Computing machinery and intelligence") per proposar la realització d'un experiment que permeti discernir el caràcter intel·ligent o no del comportament d'una màquina.

El test esmentat parteix d'un joc en què un interrogador ha d'esbrinar el sexe de dos interlocutors, A i B, situats en una habitació separada i que, encara que ambdós diuen que són dones, en realitat són un home i una dona. En la proposta original de Turing es tracta de substituir la dona per un ordinador, i això s'ha generalitzat després en la forma següent: l'interrogador ha d'esbrinar qui és la màquina a partir de la conversa amb els dos interlocutors, una persona i un ordinador, encara que ambdós diuen ser persones.

Aquest objectiu s'ha d'aconseguir tot i saber que ambdós interlocutors no estan obligats a dir la veritat i que, per exemple, la màquina pot decidir donar un resultat erroni a una multiplicació i, fins i tot, comunicar-lo després de força segons o minuts d'haver-lo obtingut per enganyar l'interrogador.

En la hipòtesi optimista del mateix Turing, cap a l'any 2000 es disposaria d'ordinadors prou potents "per a fer-los jugar tan bé l'esmentat joc que un interrogador normal no tindrà més del 70% de possibilitats d'efectuar la identificació correcta al cap de cinc minuts d'haver plantejat les preguntes".

Si fos així, seríem davant d'una màquina veritablement "intel·ligent" o, si més no, que se sap fer passar per tal.

No cal dir que la predicció de Turing peca d'un optimisme exagerat, cosa que, com veurem, és un error molt freqüent en els inicis de la IA.

En realitat el problema no afecta tan sols la potència de l'ordinador, sinó les potencialitats de la seva programació per a poder mantenir una conducta intel·ligent.

## **1.2. Objectius centrals de la investigació en intel·ligència artificial**

Es pot dir que, actualment, quan s'investiga en el camp de la intel·ligència artificial es poden perseguir dos objectius complementaris que posen l'èmfasi respectivament en l'aspecte teòric o tecnològic de la IA.

Un primer objectiu és l'**estudi dels processos cognoscitius en general**, la qual cosa justificaria la definició de Patrick J. Hayes: *L'estudi de la intel·ligència com a computació* (1973), que s'orienta en la consideració de la intel·ligència artificial com a estudi de la conducta humana intel·ligent.

En aquest estudi és fonamental l'ajut obtingut de la informàtica com a eina de gran potència en la representació de qualsevol sistema de símbols com el que fa servir l'activitat racional. D'aquí els múltiples punts de contacte de la intel·ligència artificial amb altres ciències com ara la neurofisiologia, la psicologia, la lògica formal, la lingüística, etc.

Un altre objectiu central de la IA és l'intent d'obtenir sistemes automàtics capaços de dur a terme tasques reservades, fins i tot ara, als éssers humans.

Amb aquest enfocament, la IA apareix com una disciplina eminentment tecnològica que persegueix la construcció de màquines i programes capaços de realitzar tasques complexes amb una habilitat i eficiència iguals o superiors a les que aconsegueix l'ésser humà.

## **1.3. Altres definicions**

Els dos objectius se succeeixen en les diverses definicions que els especialistes més importants de la IA han anat donant als continguts d'aquesta ciència.

Per a Marvin Minsky, professor del MIT (Massachusetts Institute of Technology) i un dels investigadors i experts més famosos en intel·ligència artificial, la intel·ligència artificial és:

"la ciència de construir màquines que fan coses que realitzades per l'home requereixen l'ús de la intel·ligència".

Un altre punt de vista l'ofereixen B.G. Buchanan i E.A. Feigenbaum, de la Universitat de Stanford, quan diu que:

"la investigació sobre IA és la part de la ciència dels ordinadors que investiga processos simbòlics, raonaments no algorísmics i representacions simbòliques del coneixement usats en màquines intel·ligents".

És important recordar que es tracta d'un ús dels ordinadors ("màquines intel·ligents") que no respecta l'habitual caracterització d'un programa com la conjunció d'uns algorismes actuant sobre unes estructures de dades. En realitat els programes de la IA són també, i tal vegada essencialment, no algorísmics i així ho constatarem més endavant.

També, en una posició intermèdia, P.H. Winston, director del laboratori d'intel·ligència artificial del MIT, diu que:

"l'objectiu de la IA es pot concretar en com aconseguir fer ordinadors més útils per a comprendre els principis que fan possible la intel·ligència".

En altres definicions s'intenta analitzar el contingut semàntic dels termes "intel·ligència" i "artificial" que componen la denominació, però aquest enfocament també sembla plantejar greus problemes en vista de les dificultats de definir els mateixos conceptes d'intel·ligència o d'artificialitat.

Al costat del contingut ja pràctic de la definició de Minsky, altres autors no dubten a proposar continguts encara més pragmàtics i diu que:

"la intel·ligència artificial és l'estudi de les facultats mentals mitjançant l'ús de models computacionals"

Charniak i McDermott, 1985.

Aquesta definició posa èmfasi en l'interès de les tècniques de la IA a duplicar (reproduir) les facultats mentals de les persones: visió, llenguatge natural, raonament i comprensió, per posar exemples concrets ja tractats per les tècniques de la IA.

### **Una altra definició**

Una altra definició complementària i en la mateixa línia, també procedent d'un llibre de text, és la que dona Elaine Rich el 1983 en dir que "la intel·ligència artificial és l'estudi de com es pot fer que els ordinadors facin coses que, de moment, les persones fan millor".

Potser aquestes definicions il·lustren més exactament el contingut actual de la intel·ligència artificial que l'ambició objectiu de superar el test de Turing amb una màquina o un programa d'ordinador.

#### 1.4. Dos enfocaments oposats

Un element important en la realitat actual i històrica de la intel·ligència artificial és que hi ha dos enfocaments oposats sobre com s'ha d'aconseguir aquest estudi i duplicació tecnològica de la intel·ligència i els processos cognoscitius habituals en l'ésser humà.

Per una part, **John McCarthy**, inventor de la denominació "intel·ligència artificial", creador del llenguatge Lisp i un dels investigadors més famosos de la IA, recomana una línia d'investigació que persegueix l'obtenció de programes d'ordinador que raonin seguint bàsicament les exigències dels dictats de la **lògica matemàtica**.

A l'extrem oposat, **Marvin Minsky**, també famós expert de la IA, proposa imitar fonamentalment el mètode de **raonament de la ment humana** que, segons ell, no té per què ser necessàriament el mateix de la lògica matemàtica.

Potser l'enfocament de McCarthy presenta més facilitats per a arribar, si més no, a certs resultats. Però l'observació de Minsky estableix les bases d'un dubte raonable sobre si els resultats obtinguts per l'enfocament proposat per McCarthy tenen realment alguna cosa a veure amb la intel·ligència humana o tan sols són meres aproximacions que poden tancar el camí a la veritable imitació de la intel·ligència humana, si aquest n'és l'objectiu principal, la qual cosa, com ja s'ha dit, és posada en dubte per altres investigadors més pragmàtics.

D'aquí que, en conjunt, les tècniques de la IA, en el seu intent d'imitar les característiques de la ment humana, tractin, en qualsevol cas, l'estudi dels mètodes basats en la representació simbòlica del coneixement o en la inferència simbòlica, i es preocupin per problemàtiques com la comprensió dels processos cognoscitius, el raonament i la presa de decisions, la comprensió i la utilització del llenguatge natural -en oposició a llenguatges més formals i més "fàcils" per als sistemes informàtics-, la percepció visual, els processos d'aprenentatge, etc.

#### 1.5. Altres denominacions

El nou apogeu de les tècniques de la intel·ligència artificial en els últims anys va provocar, a la dècada dels vuitanta, que alguns intentessin utilitzar les mateixes inicials (IA) com a indicadors també d'*informàtica avançada*, en un intent d'expressar que, malgrat els seus més de trenta anys d'antiguitat, l'èxit de les tècniques de la IA la configurava com un dels aspectes més prometedors de la informàtica i un dels més importants amb vista al futur immediat de la tecnologia informàtica. Aquesta denominació, encara que ha estat freqüentment utilitzada en el món comercial i empresarial, resulta força menyspreada en el

món acadèmic, preocupat bàsicament pels aspectes teòrics de la IA i que, per tant, recomana preferentment denominacions com ara **sistemes basats en el coneixement** o **ciència dels processos cognoscitius**.

També cal esmentar l'observació de John Haugeland, que proposa el terme **intel·ligència sintètica** per a destacar, al costat de l'artificialitat de la intel·ligència aconseguida per la IA, el fet que aquesta resulta originada en l'activitat humana que "sintetitza" així una nova forma de la seva pròpia intel·ligència.

## 2. Aplicacions de la IA

Avui dia les tècniques de la intel·ligència artificial s'apliquen a molts camps. A continuació en destacarem alguns:

- Sistemes experts
- Sistemes de visió per ordinador
- Sistemes de llenguatge natural
- Resolució de problemes
- Robòtica
- Aprenentatge
- Jocs i puzles

### 2.1. Sistemes experts

Els sistemes experts constitueixen l'eix central d'una nova activitat informàtica coneguda com **enginyeria del coneixement**. Amb això es persegueix la construcció de programes que reproduïxin de forma correcta el comportament d'un expert humà en el seu domini concret de competència i perícia.

Els sistemes experts també reben el nom de *sistemes basats en el coneixement* i, amb intencions més teòriques, *sistemes de producció*.

### 2.2. Sistemes de visió per ordinador

Després de l'objectiu central de l'antic *perceptron* de Rosenblat, els sistemes de visió per ordinador persegueixen el **reconeixement de formes** amb l'objectiu d'aconseguir la **visió per ordinador**, tant amb la finalitat d'aplicar-la en l'àmbit de la robòtica com en la seva aplicació en medicina, per exemple per a ajudar les persones que pateixen de ceguesa.

#### Perceptron de Rosenblat

Aquest tema es tracta més endavant en l'apartat "El perceptron de Rosenblat"

### 2.3. Sistemes de llenguatge natural

El llenguatge que utilitzem els éssers humans s'ha etiquetat de "llenguatge natural" en contraposició als llenguatges "artificials", que s'han dissenyat per als ordinadors. La **comprensió del llenguatge natural** és molt difícil, ja que hi ha termes o expressions amb més d'un significat.

#### Temes d'estudi

Entre els temes d'estudi del llenguatge natural s'inclouen la comprensió i l'anàlisi de textos, la traducció automàtica i la generació d'informes o textos. També s'hi inclouen el tractament de la paraula parlada i altres tècniques auxiliars com el reconeixement vocal i el tractament i la comprensió de l'expressió oral.

El tractament del llenguatge natural és un dels temes més interessants que tracta la IA, tant per l'objectiu de l'estudi en si mateix com per la possibilitat de poder accedir en un futur als ordinadors fent servir el mateix llenguatge natural que utilitzem els éssers humans, propòsit central del projecte japonès de l'anomenada "cinquena generació d'ordinadors".

### 2.4. Resolució de problemes

Un dels temes centrals de la IA és l'ús de la intel·ligència per a la resolució de tot tipus de problemes.

Els objectius centrals han estat trobar els mecanismes de deducció, cercar solucions, i també aconseguir l'explicació de la solució que s'ha obtingut i la seva justificació.

També s'estudia la planificació intel·ligent de conductes i procediments d'acció. Encara que en aquest tema els sistemes experts han avançat molt més en les seves aplicacions puntuals, l'estudi de la resolució de problemes continua essent la base teòrica d'un plantejament més general.

### 2.5. Robòtica

Com a resultat de l'evolució directa de l'enginyeria mecànica i l'automatització, la IA també intenta aconseguir l'**augment de l'habilitat** i l'**autonomia** d'aquests mecanismes que manipulen instruments i que anomenem *robots*.

La robòtica necessita tant la percepció (visual o d'una altra forma) de l'entorn on treballa el robot com la planificació de l'acció que s'ha d'emprendre en cada cas. Encara que amb gran èxit i una aplicació creixent en el món industrial, els robots que s'ha aconseguit crear encara són molt lluny dels de ficció, que són els que configuren la imatge i les expectatives de la robòtica en la majoria del públic.

#### Resolutor General de Problemas

Després de l'intent del Resolutor General de Problemas (GPS, *General Problem Solver*), l'any 1972 Newell i Simon van publicar un influent llibre sobre aquest tema: *Human Problem Solving*.

## 2.6. Aprenentatge

És possible que l'essència de la intel·ligència sigui la capacitat d'aprendre, és a dir, d'obtenir coneixements nous a partir de l'experiència i els coneixements ja adquirits. L'objectiu de fer que els programes "aprenquin" és central en les tècniques de la IA. Un dels exemples més coneguts són els programes que juguen a certs jocs de contingut estratègic (escacs, dames, bridge, etc.) en què, freqüentment, també es persegueix que, a partir dels resultats obtinguts, millorin el seu comportament en futures partides.

Al costat de l'interès que els programes d'ordinador puguin "aprendre", també es dóna molta importància al fet que la informàtica es pugui utilitzar per a afavorir el procés d'aprenentatge dels éssers humans. Es parla així d'**ensenyament assistit per ordinador** (EAO), en el qual destaca, entre altres utilitzacions, la creació de les anomenades *situacions de descobriment* gràcies a l'experimentació directa amb certes eines informàtiques.

Un element central d'aquesta utilització pedagògica de la IA és el llenguatge Logo, desenvolupat al MIT per Seymour Papert adaptant idees nascudes al Lisp. L'ampli univers pedagògic que ha obert el Logo i els seus "micromons" és un dels escassos exemples en què la utilització d'una nova tecnologia en l'educació també ha vingut acompanyada d'un mètode pedagògic *ad hoc*.

## 2.7. Jocs i puzles

Una de les aplicacions més immediates de la investigació en IA ha estat l'elaboració de **programes de jocs** que, al mateix temps, puguin ser exemples de l'aplicació de les tècniques de resolució de problemes i d'aprenentatge, quan s'apliquen a universos o dominis restringits com són, en definitiva, els dominis que descriuen les regles dels jocs. En aquest aspecte destaquen l'aplicació de tècniques d'IA en programes d'ordinador que juguen a dames, escacs o bridge i altres jocs estratègics de sala per a la pràctica dels quals es presumeix una certa necessitat d'intel·ligència.

S'han estudiat els jocs i els puzles perquè permeten tractar de forma restringida les dificultats generals de la resolució de problemes. Bàsicament s'intenta trobar una representació adequada de l'espai de casos propi del problema, per a configurar el que es denomina "espais d'estats".

La reducció de possibilitats dels universos explorats als jocs -respecte del món real-, ofereix la possibilitat de tractar de forma particular els greus problemes derivats de l'explosió combinatòria del nombre de casos possibles i de les solucions que cal provar.

### Exemples de jocs

Exemples utilitzats àmpliament en la investigació són els jocs senzills com el marro, els puzles simples, o disposar vuit reines en un tauler d'escacs sense que es "mengin" entre si.



### 3. Tècniques de la IA

Quan s'aborda la resolució d'un problema amb tècniques d'IA, es tracta tant d'obtenir-ne la solució com de poder exposar el procés de raonament que s'ha utilitzat. Òbviament s'hauran d'obtenir diverses solucions si n'existeix més d'una, i justificar l'interès de cada una d'elles. També és important que el sistema d'IA sàpiga millorar el raonament utilitzat a mesura que augmenta l'experiència, fet que constitueix la seva forma concreta d'aprenentatge.

#### 3.1. Tècniques de base de la IA

De la mateixa manera que la informàtica tradicional es compon d'algorismes que operen sobre determinades estructures de dades, la tècnica pròpia de la IA és obtenir una adequada representació simbòlica del coneixement.

També és possible aplicar tant els mecanismes de la inferència simbòlica, propis del raonament deductiu, com els mecanismes de tipus empíric que formen la base de la recerca heurística de solucions.

#### 3.2. Inferència simbòlica

Els mecanismes de deducció típics de la inferència simbòlica inclouen la **deducció** -també anomenada de vegades *inferència lògicament correcta*- emparada en les diverses regles de la inferència lògica:

a) El **modus ponens**, amb regles del tipus: "si  $p$  llavors  $q$ ", és a dir, del fet o premissa  $p$  es pot afirmar el fet o premissa  $q$ .

b) La **instanciació universal**, que ens indica que si alguna cosa és certa de tots els elements d'un conjunt, també és cert per a cada cas particular, com recorda el conegut raonament: "tots els homes són mortals, Sòcrates és un home, per tant, Sòcrates és mortal".

Però la IA també utilitza per als seus raonaments altres tipus de regles d'inferència lògica, no tan immediates ni sempre segures.

c) Un exemple és el mecanisme "raonador" de la **inducció** pel qual, predicada una propietat d'una àmplia sèrie d'individus, s'"indueix" que la propietat pot ser predicada de tots els individus de la mateixa espècie.

d) Un altre exemple és una versió particular de l'anomenat *modus tollens* de la lògica en l'anomenada abducció, que a partir de la regla "si  $a$  llavors  $b$ " i el fet o premissa  $b$  s'atreveix a postular la possibilitat de  $a$ . Encara que l'abducció pugui semblar una manera de procedir estranya i arriscada, té la seva utilitat en la temptativa de recerca de solucions.

### 3.3. Heurística

Però els mecanismes de la inferència simbòlica no són suficients, i per això els sistemes d'IA utilitzen molt sovint les tècniques de l'heurística.

L'heurística es podria definir com el conjunt dels criteris, mètodes o principis que s'utilitzen per a trobar, d'entre diversos camins possibles, quin o quins són els més efectius per a obtenir un objectiu determinat.

Evidentment, l'heurística també té molt a veure amb els mecanismes experimentals i empírics, la síntesi dels quals acaba elaborant aquestes regles de l'experiència que s'empren per a seleccionar un camí d'acció davant l'explosió combinatòria dels molts casos possibles.

Investigacions d'IA han demostrat que un jugador d'escacs expert, no analitza *totes* les jugades possibles en un moment determinat del joc, sinó tan sols aquelles que, heurísticament, reconeix, gràcies a la seva experiència, com a més prometedores o interessants.

#### La diagnosi mèdica

Un exemple concret és la diagnosi mèdica, que ve a ser, en realitat, un ús prudent de l'abducció ja que, a partir de diversos símptomes, el metge pot fer una abducció de l'existència d'una causa possible.

## 4. La representació del coneixement

Un dels temes centrals en la intel·ligència artificial moderna és el dels diferents sistemes que la IA utilitza per a la representació del coneixement i, implícitament, les possibilitats que l'esmentada representació ofereix per a la seva utilització.

El que es persegueix no és tan sols unes estructures de dades que proporcionin un sistema de representació efectiu i eficient dels coneixements, sinó també quin coneixement s'ha de representar en cada utilització particular.

Els problemes annexos i importants són:

- a) L'extracció del coneixement de qui el posseeix.
- b) La seva formalització en un determinat sistema de representació.
- c) La possibilitat de crear el coneixement esmentat, si escau, o **modificar el coneixement existent** gràcies a la interacció del sistema d'IA amb l'entorn en un món canviant.

Òbviament, cada sistema de representació del coneixement pot anar associat a una determinada forma de raonar amb el coneixement emmagatzemat, de manera que aquest coneixement s'utilitzi adequadament per a obtenir la solució desitjada. Un problema potser posterior, però no menys important, és la implementació informàtica de la representació del coneixement escollida de manera que sigui eficient.

### 4.1. Components i elements que s'han de representar

En qualsevol sistema de representació del coneixement són components essencials tant les estructures de dades que s'utilitzen com els seus procediments d'interpretació i maneig.

Els diversos elements que s'han de representar són:

- a) El coneixement sobre els **objectes**.
- b) El coneixement sobre els **processos**.
- c) El coneixement sobre l'**entorn** en què existeixen objectes i processos.

A més, també s'ha d'implementar la representació de diversos objectius, motivacions, elements de causalitat, temporalitat, etc. Un dels elements més difícils d'incorporar als sistemes d'IA és l'aparentment senzill **sentit comú**, que sempre comporta un coneixement difícil de determinar i precisar i que, sovint, s'acaba convertint en regles de tipus heurístic.

## 4.2. Etapes en l'ús del coneixement

En l'ús del coneixement es reconeixen diverses etapes:

- a) L'**adquisició**. Es tracta d'acumular nou coneixement i relacionar-lo amb el que ja es tenia.
- b) La **recuperació**. Es tracta d'obtenir, d'una àmplia base de coneixements, precisament aquell que sigui veritablement rellevant per a resoldre el problema en curs.
- c) El **raonament**. Es tracta d'inferir alguna cosa, a partir del coneixement ja seleccionat, per a obtenir i verificar fets nous diferents dels ja coneguts.

## 4.3. Propietats d'un bon formalisme de representació del coneixement

A un sistema de representació del coneixement se li solen exigir diverses propietats com ara:

- a) L'**adequació representacional** o possibilitat de representar totes les classes de coneixement que siguin necessàries en un domini donat.
- b) L'**adequació inferencial** o possibilitat de manipulació de les estructures de dades que serveixen per a representar el coneixement, de manera que sigui possible generar noves estructures que corresponguin a nous coneixements inferits dels anteriors.
- c) L'**eficiència inferencial** o capacitat d'incorporar també un nou coneixement addicional (*metaconeixement*) que pot emprar-se per a millorar l'ús dels mecanismes d'inferència i optimitzar el funcionament del sistema d'IA.
- d) L'**eficiència en l'adquisició del coneixement** o facilitat d'adquirir informació nova incloent-hi el fet que, en el cas ideal, el propi sistema d'IA ha de ser capaç de controlar l'adquisició de nou coneixement o metaconeixement.

#### 4.4. Els espais d'estats

Un dels primers formalismes utilitzats per a la representació del coneixement va ser l'anomenat **espai d'estats**, que descriu en els seus **estats** tots els moments i passos tant del problema com de la solució. Aquesta solució s'obté amb uns operadors o **transicions** que actuen transformant la situació del problema d'un estat a un altre dins de l'espai d'estats.

Aquest sistema resulta adequat quan no es produeix una explosió combinatòria del nombre de casos possibles, la qual cosa pot ser certa quan s'aborden universos reduïts com passa amb alguns jocs o puzles. Però en altres casos, com en el conegut joc dels escacs, l'explosió combinatòria d'estats és tal que resulta imprescindible recórrer a regles de tipus heurístic per a simplificar la multitud d'estats possibles.

#### 4.5. Esquemes moderns de representació del coneixement

L'esgotament de la capacitat de representació i manipulació dels espais d'estats va fer aparèixer altres esquemes de representació del coneixement.

Se sol distingir entre els **esquemes de tipus declaratiu**, com els esquemes lògics i les xarxes semàntiques, i els **esquemes de representació de procediments**, com els sistemes de representació procedural o els sistemes de producció, també anomenats *sistemes experts*. Destaca també la tendència integradora representada pels *frames* i els *scripts*.

##### 4.5.1. Els esquemes lògics

En els esquemes lògics (un exemple senzill d'esquemes de tipus declaratiu) els fets i les regles són fórmules expressades en algun sistema basat en la lògica com passa amb el llenguatge Prolog, que utilitza la lògica de primer ordre del càlcul de predicats. També hi ha l'ús de lògiques més complexes com la lògica multivaluada, la temporal o la difusa, per esmentar exemples ja utilitzats a la IA.

Una **fórmula lògica** és, en definitiva, una combinació de predicats, variables, connectors lògics, quantificadors i funcions.

L'avantatge d'aquest sistema és que es tracta d'un procediment precís, flexible, modular i bastant natural, encara que sol estar mancat de principis d'organització i selecció entre els fets (deficiències del metaconeixement). També acostuma a tractar objectes massa simples i resulta difícil de manipular si hi ha moltes dades.

#### La programació lògica

L'exemple més clàssic de l'ús d'esquemes lògics és la programació lògica implementada amb el llenguatge Prolog i el seu ús dels mecanismes de resolució, unificació i resolució per refutació.

### 4.5.2. Les xarxes semàntiques

Les xarxes semàntiques són uns sistemes de representació, inicialment gràfics, que resulten molt adequats per a establir taxonomies. Al graf que constitueix una xarxa semàntica, els **nodes** incorporen objectes o conceptes i els **arcs** són relacions entre els nodes.

Als nodes es descriuen tant **conceptes** –constants o paràmetres de la realitat que especifiquen objectes o abstraccions– com **esdeveniments** –accions que ocorren en la realitat que s'està modelant– i **característiques** –que particularitzen estats, modifiquen conceptes, esdeveniments i també altres característiques.

Als arcs se solen distingir els que serveixen a efectes de classificació, que relacionen un objecte amb altres de la seva classe com passa amb els casos "membre-de" o "instància-de", d'**agregació**, que relacionen un objecte amb els seus components amb mecanismes com "part-de", i de **generalització**, que relacionen un objecte amb una classe més general amb exemples del tipus "és-un" o "és-subconjunt-de".

#### Variants evolucionades

Algunes variants evolucionades a partir de les xarxes semàntiques són les xarxes particionades (com proposa Hendrix), els esquemes de propagació de *markers* (proposta de Fahlman), les jerarquies de tòpics (del tipus "és-un" o "part-de"), les xarxes proposicionals (com suggereix Shapiro), etc.

La característica fonamental de les xarxes semàntiques és la seva facilitat per a implementar l'**herència de propietats**, que permet deduir fets nous derivats del fet que els nodes més alts en la jerarquia són també assercions sobre els nodes més baixos.

En les figures que podeu veure al marge es mostren dos exemples concrets de les moltes formes que pot adoptar una xarxa semàntica.

#### Formes que pot adoptar una xarxa semàntica

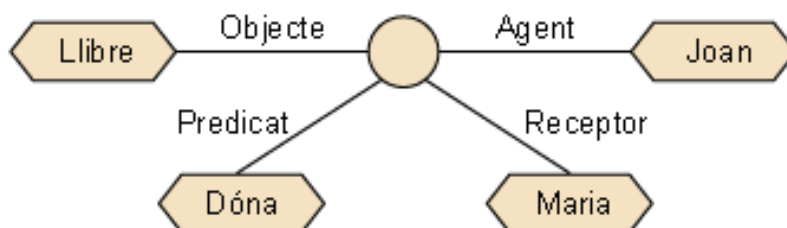


Figura 1. En Joan dóna un llibre a la Maria

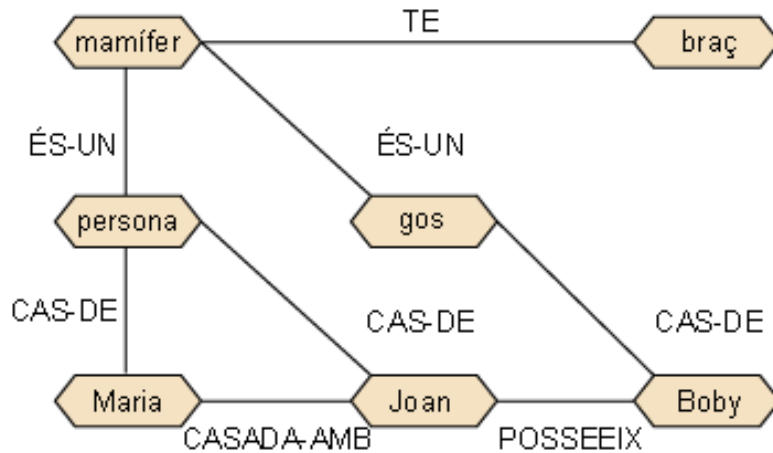


Figura 2. Una jerarquia "ÉS-UN"

#### 4.5.3. Els esquemes de representació de procediments

Els sistemes de representació procedural es basen en la idea d'emmagatzemar el coneixement en forma de procediments (*procedures*) en lloc de fer-ho en forma de proposicions.

Però sembla que s'ha obtingut més bon resultat amb el que els teòrics anomenen sistemes de producció i que, més freqüentment, reben el nom de sistemes experts, que estudiarem amb més detall a l'apartat 5.

En aquest cas es tracta de:

- Un conjunt de regles de producció que constitueix la base de coneixements.
- Una base de fets que descriu el context d'utilització del sistema expert.
- Un intèrpret o motor d'inferència.

El motor d'inferència obté un nou coneixement per un procediment anomenat d'unificació –selecció de la regla que s'ha d'utilitzar en cada cas–, una resolució dels possibles conflictes entre regles, i una acció que decideix regles o fets nous que s'han d'emmagatzemar a la base de fets.

#### 4.5.4. Els frames

El *frame*, que els experts espanyols han decidit NO traduir per "quadre", es pot assimilar a una xarxa semàntica complexa d'estructures que s'utilitzen per a descriure la col·lecció d'atributs que té un determinat objecte o una situació complexa. Compartint les propietats d'una xarxa es pot establir una jerarquia de frames, també apta per a establir taxonomies tant de pertinença a una classe (*member link*, del tipus "En Joan ÉS-UN home"), com d'adscripció a una altra (*subclass link*, del tipus "l'home ÉS-UN mamífer").

#### Informació en els frames

Proposats per Minsky en un article avui clàssic, "A Framework for representing Knowledge" (1975), els frames inclouen en les seves estructures de dades tant informació procedural com declarativa.

Els *frames* disposen de **ranures d'expansió** (*slots*) com a llocs físics per a inserir una part de coneixement o propietat dins de l'estructura. Hi ha les ranures d'expansió **membre**, que descriuen els atributs de cada membre d'una classe i s'utilitzen només als *frames* que representen classes, i les ranures d'expansió **pròpies**, que descriuen els atributs de la classe en si.

L'avantatge dels *frames* sobre les xarxes semàntiques és que també contenen informació procedural. Aquesta informació és als **mètodes**, que vénen a ser procediments escrits, per exemple en llenguatges com el Lisp, i que s'emmagatzemen com a valors de les ranures d'expansió.

També utilitzen els **valors actius** (*demons*) com a procediments o regles de producció lligats a una ranura d'expansió i que es fan servir quan canvia la informació que conté, tant a l'hora d'afegir-ne de nova com en eliminar o modificar la que ja conté.

A més, els *frames* incorporen mètodes d'inferència com l'herència de propietats, pròpia de les xarxes semàntiques, i també el raonament mitjançant *facets*, que són restriccions sobre el nombre de valors possibles que pot tenir un atribut i sobre les classes a què pot pertànyer cada valor.

#### 4.5.5. Els *scripts*

Un últim exemple de sistema de representació mixta és el dels *scripts* o "guions". Es tracta d'estructures especialitzades que descriuen seqüències d'esdeveniments en un context particular. Els components principals d'un *script* són:

- a) Les **condicions d'entrada** que s'han de satisfer abans que els esdeveniments descrits a l'*script* es puguin iniciar.
- b) El **resultat** o les condicions que siguin certes quan els esdeveniments descrits s'hagin complert.
- c) Les **propietats** (*prop*), que són ranures d'expansió que representen els objectes involucrats en els esdeveniments que descriu l'*script*.
- d) Els **personatges** (*rols*), que són les diferents entitats o persones que intervenen en els esdeveniments descrits.
- e) La **instància** (*track*), que és una variació específica o un cas particular d'un model més general que s'ha representat mitjançant aquest *script*.
- f) Les **escenes**, que són les diferents seqüències d'esdeveniments que tenen lloc a l'*script*.



## 5. Sistemes experts

### 5.1. Definició i característiques

Els sistemes experts, també anomenats sistemes basats en el coneixement, representen un dels èxits més importants de la IA, almenys pel que fa referència a l'èxit de les seves aplicacions en l'àmbit dels sistemes d'informació.

La finalitat principal dels sistemes experts és la reproducció correcta del comportament d'un expert humà en el seu domini de competència.

#### Els sistemes experts

Són un dels exemples més clars de l'evolució de la IA i del seu èxit quan aquesta ha prescindit dels objectius exageradament ambiciosos dels primers anys, i ha delimitat el domini i les característiques dels problemes que s'han de resoldre.

Algun especialista ha volgut etiquetar els sistemes experts com "la IA amb èxit", cosa que és, evidentment, una exageració, però indica la importància dels sistemes experts en la superació dels "anys difícils" de la IA, a l'inici dels anys seixanta, amb diversos projectes en curs i un nombre excessiu de fracassos.

#### 5.1.1. Atributs d'un expert humà

Un expert humà és acceptat com a tal per una sèrie d'atributs que caracteritzen la seva tasca, en la qual se li reconeix una vertadera competència, habilitat i experiència. En general, l'expert humà que aquí ens interessa considerar resol problemes complexos, actualitza els seus coneixements, és capaç de justificar les seves decisions i respostes, a les quals també pot atribuir graus de credibilitat.

El bon expert humà també és capaç de destacar els aspectes més rellevants d'un problema, a l'hora que detecta i reconeix la desaparició de la seva competència o perícia quan el problema en qüestió s'apropa a les fronteres del domini de la seva experiència, és a dir, és conscient dels límits de la seva competència.

També és capaç d'aprendre a raonar millor, reestructurant els seus coneixements, aplicant procediments o regles per als casos d'excepció i elaborant noves regles o procediments per enfrontar-se a nous problemes. Per si no n'hi ha prou amb tot això, a més disposa de sentit comú per a situar la seva manera de procedir dins d'un àmbit més general.

### 5.1.2. Definició i característiques d'un sistema expert

Un sistema expert és un sistema elaborat amb tècniques d'IA que, de la mateixa manera que l'expert humà a qui intenta emular, resol problemes complexos i difícils que se circumscriuen en un domini específic i delimitat.

El sistema expert utilitza processos que imiten el raonament humà (deducció, inducció, estratègies de recerca de solucions, etc.) a l'hora de resoldre els problemes. I tot això ho aconsegueix utilitzant els coneixements de base, subministrats en origen per un expert humà, als quals també incorpora els coneixements que el sistema expert informàtic "aprèn" durant la seva activitat com a "expert".

A més, el sistema expert ha de ser capaç de justificar les decisions i els resultats obtinguts i atribuir-los graus de credibilitat. També, com fan de vegades els experts humans, ha de ser capaç de raonar a partir de dades incertes.

### 5.2. Estructura bàsica d'un sistema expert

A la figura següent es mostra l'estructura bàsica d'un sistema expert en la qual es percep clarament un aspecte fonamental: la separació de les dades que formen el coneixement, de les estratègies de resolució o procediment amb què s'elaboren les solucions.

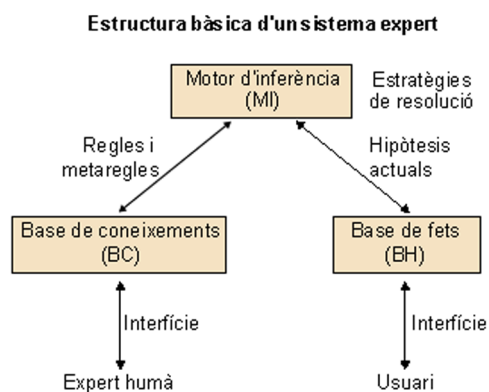


Figura 3

Els coneixements s'emmagatzemen a la **base de coneixements** i a la **base de fets**, i els procediments capaços de raonar s'implementen al **motor d'inferència**.

L'esquema també mostra la necessitat d'interfícies que permetin l'accés al sistema, tant de l'expert humà que alimenta la base de coneixements com de l'usuari del sistema expert que proporciona els fets que determinen una utilització concreta del sistema expert. L'usuari també obté les respostes que proporciona el sistema.

### 5.2.1. La base de coneixements

La base de coneixements ve a ser la **memòria a llarg termini** del sistema expert i conté les **regles** i les **metaregles** que resumeixen el coneixement de l'expert humà sobre el domini del problema. Les metaregles indiquen l'ordre o la preferència amb què s'han d'usar les regles que descriuen el coneixement de l'expert en cada cas.

Una de les formes habituals (i també més simples) és la utilització de sistemes senzills de representació del coneixement, com són les **regles lògiques** del tipus "Si... llavors...".

Les regles són actualitzables (i utilitzables) d'una en una, i independents del seu ús posterior.

#### Exemple 1

SI (la persona-X ÉS-UN home)

I (la persona-Y ÉS-FILL de la persona-X)

I (la persona-Z ÉS-FILL de la persona-Y)

LLAVORS (la persona X ÉS-AVI de la persona-Z).

#### Exemple 2

SI (el país -X i el país-Y SÓN-VEÏNS)

LLAVORS (el color del país-X ÉS-DIFERENT del color del país-Y).

### 5.2.2. La base de fets

La base de fets, també anomenada de vegades **memòria a curt termini** o **memòria de treball**, emmagatzema tant els fets proporcionats per l'usuari, i que descriuen la situació concreta analitzada, com els fets nous que el propi sistema expert va obtenint.

#### Exemple 1

En Joan ÉS-UN home

I ÉS-FILL d'en Carles.

#### Exemple 2

El groc ÉS-UN color.

### 5.2.3. El motor d'inferència

El motor d'inferència conté el **mecanisme de raonament** que segueix el sistema expert: deducció, inducció, estratègies de recerca de solucions, etc.

Amb això, el motor d'inferència, en aplicar les regles contingudes a la base de coneixements sobre la base de fets, obté nou coneixement que s'incorpora també a la base de fets i és objecte, al seu torn, de l'aplicació de les regles recollides a la base de coneixements.

Un dels problemes centrals en el funcionament del motor d'inferència quan les bases de coneixements i de fets són voluminoses, és la selecció adequada de les regles i els fets que s'han de considerar a cada pas del "raonament". En aquest aspecte és fonamental l'existència de **metaregles** específiques que orientin precisament la selecció esmentada.

### 5.2.4. Les interfícies

Un element essencial en la facilitat d'ús del sistema expert són les interfícies d'entrada de dades que permeten la interacció, tant de l'expert com de l'usuari, amb el sistema expert. També hi ha la possibilitat, no recollida a l'esquema de la figura 4 que repetim al marge, d'incorporar una interfície, realitzada amb els sistemes de la informàtica tradicional, que permeti l'extracció de dades de la base de dades de gestió per a alimentar part de la base de fets. En les interfícies amb els usuaris humans també és habitual buscar les formes més amigables del diàleg persona-màquina.

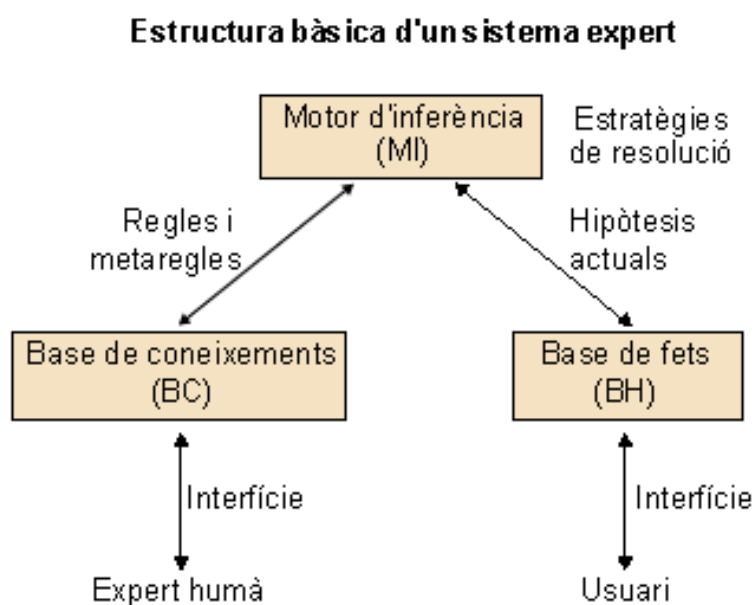


Figura 4

La figura 5 al marge mostra l'entorn de desenvolupament en què l'expert humà, potser ajudat per un enginyer del coneixement, dialoga amb el sistema per a crear la base de coneixements.

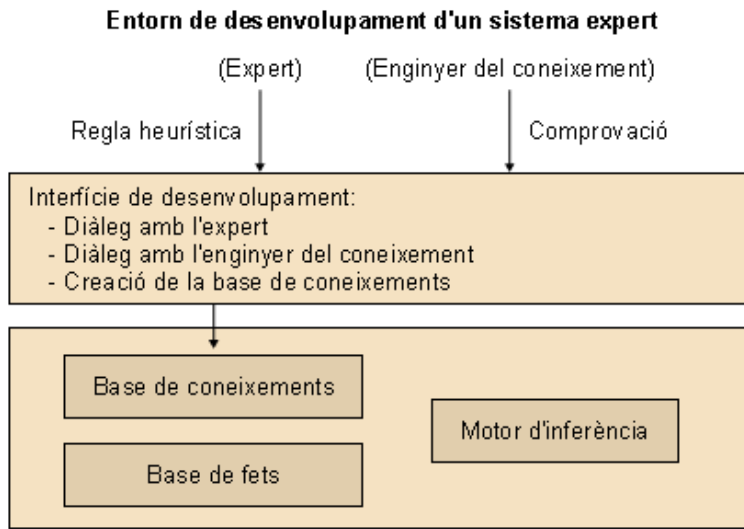


Figura 5

Així mateix, la figura 6 situada també al marge, mostra un possible entorn d'utilització del sistema expert, en el qual un usuari dialoga amb el sistema per a crear la base de fets i interrogar el sistema expert per a obtenir la justificació de les recomanacions o decisions d'aquest (proporcionades per mitjà de l'activitat "raonadora" i "explicadora" del motor d'inferència).

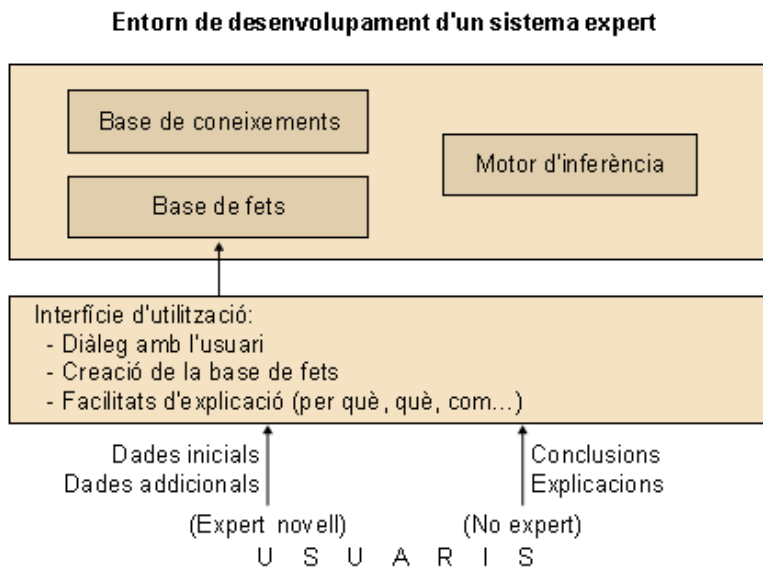


Figura 6

### 5.3. Bases de coneixements: tipus de regles

Encara en el cas senzill de coneixements emmagatzemats en forma de regles simples de la lògica elemental, cal considerar casos com els següents.

### 5.3.1. Regles antecedent-conseqüència

Són del tipus *si tots els antecedents són certs, llavors totes les conseqüències són certes*, i responen a l'esquema general:

```

SI          <antecedent-1>
           <antecedent-2>
           ...
LLAVORS    <conseqüència-1>
           <conseqüència-2>

```

del qual, en un sistema expert de diagnòstic mèdic, s'obtindria l'exemple següent:

```

SI          (pacient TÉ símptoma-1)
           I (pacient TÉ símptoma-5)
           I (pacient TÉ símptoma-7)
LLAVORS    (pacient TÉ malaltia-A)

```

### 5.3.2. Regles condició-acció

Són del tipus *si totes les condicions són certes, llavors s'han d'executar totes les accions*, i responen a l'esquema general:

```

SI          <condició-1>
           <condició-2>
           ...
LLAVORS    <acció-1>
           <acció-2>
           ...

```

del qual, en un sistema expert per a jugar als escacs, s'obtindria l'exemple següent:

```

SI          (torre-negra A B2)
           I (reina-blanca A D2)
           I (C2 ÉS-LLIUERE)
           I (MOUEN-NEGRES)
LLAVORS    ESBORRAR: (torre-negra A B2)
           ESBORRAR: (reina-blanca A D2)
           ESBORRAR: (MOUEN-NEGRES)
           INSERIR: (torre-negra A D2)
           INSERIR: (MOUEN-BLANQUES)

```

### 5.3.3. Regles inexactes

Són del tipus: *si tots els antecedents són certs, llavors la conseqüència és certa amb un determinat grau de certesa*, i responen a l'esquema general:

```

SI          <antecedent-1>
           <antecedent-2>
           ...
LLAVORS    <conseqüència-1>
AMB UNA CERTESA DEL: <factor-de-certesa>

```

del qual, en un sistema expert de diagnòstic mèdic, s'obtidria l'exemple següent:

```

SI          (pacient TÉ símptoma-2)
           I (pacient TÉ símptoma-5)
           I (pacient TÉ símptoma-8)
LLAVORS    (pacient TÉ malaltia-B)
AMB UNA CERTESA DEL: 0,7

```

#### 5.4. Motor d'inferència: funcionament i mecanismes de raonament

El motor d'inferència combina els fets i les regles per a obtenir nous fets.

El funcionament típic d'un motor d'inferència s'il·lustra en l'exemple següent:

```

FET:       (pacient TÉ símptoma-3)
FET:       (pacient TÉ símptoma-7)
REGLA:     SI (pacient TÉ símptoma-3)
           I (pacient TÉ símptoma-7)
           LLAVORS (pacient TÉ malaltia-C)

```

del qual el motor d'inferència pot obtenir la conclusió següent:

```
FET INFERIT: (pacient TÉ malaltia-C)
```

##### 5.4.1. Raonament per encadenament cap endavant

Una de les estratègies generals a disposició del motor d'inferència és l'anomenat encadenament cap endavant (*forward chaining*).

Es tracta d'un procediment governat per les dades (*data driven*), en què es parteix dels fets coneguts i, amb l'ús de les regles, es van deduint fets nous.

##### Raonament per encadenament cap endavant (*forward chaining*)

En l'exemple de la figura 7, els fets inicials (C i F) permeten obtenir el nou fet E en aplicar la regla 1 (*si C llavors E*). Un cop es disposa del fet E, la seva combinació amb un dels fets inicials (F) permet que la regla 3 (*si E i F llavors D*) obtingui un nou fet D a la base de fets.

Ara l'aplicació de la regla 2 (*si D llavors B*) permet obtenir un nou fet B. I finalment l'aplicació de la regla 4 (*si B i C llavors A*) permet completar la base de fets amb el nou descobriment A.

La figura 7 mostra a la seva part dreta inferior l'evolució de la memòria de treball (base de fets) i les conclusions finals obtingudes.

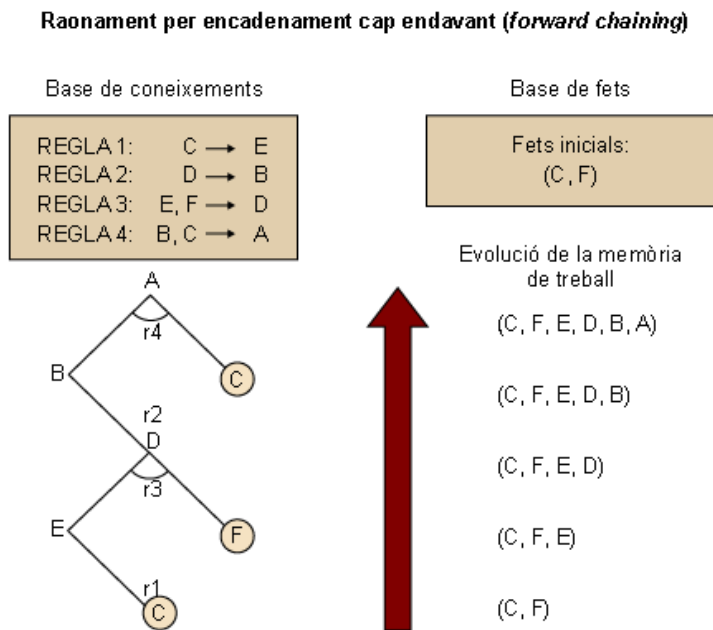


Figura 7

#### 5.4.2. Raonament per encadenament cap enrere

Una altra estratègia habitual és l'encadenament cap enrere (*backward chaining*).

En el cas de l'encadenament cap enrere (*backward chaining*) es tracta d'un procediment de recerca dirigida per l'intent d'afirmar un objectiu concret (*goal directed*) que s'intenta provar d'acord amb els fets i les regles disponibles.

##### Raonament per encadenament cap enrere (*backward chaining*)

En l'exemple de la figura 8, en intentar deduir si l'objectiu A és cert, la regla 4 (si B i C llavors A) permet substituir l'objectiu inicial A pel conjunt B i C.

Com que C ja existeix en la base de fets, l'objectiu A es complirà si es compleix l'objectiu B. Amb això B es converteix en el nou objectiu que s'ha de verificar.

Per a verificar B, la regla 2 (si D llavors B) permet substituir l'objectiu B per un nou objectiu D. Ara la regla 3 (si E i F llavors D) diu que D es compleix si ho fan E i F, i com que F ja forma part de les dades inicials de la base de fets, permet substituir l'objectiu D per un nou objectiu E.

Finalment, la regla 1 (si C llavors E) fa veure que E es compleix si es compleix C. Com que C forma part de la base de fets inicial es compleix i, consegüentment, també es compleixen tots els altres objectius parcials (D i B) fins a arribar a l'objectiu inicial A, que queda així comprovat.



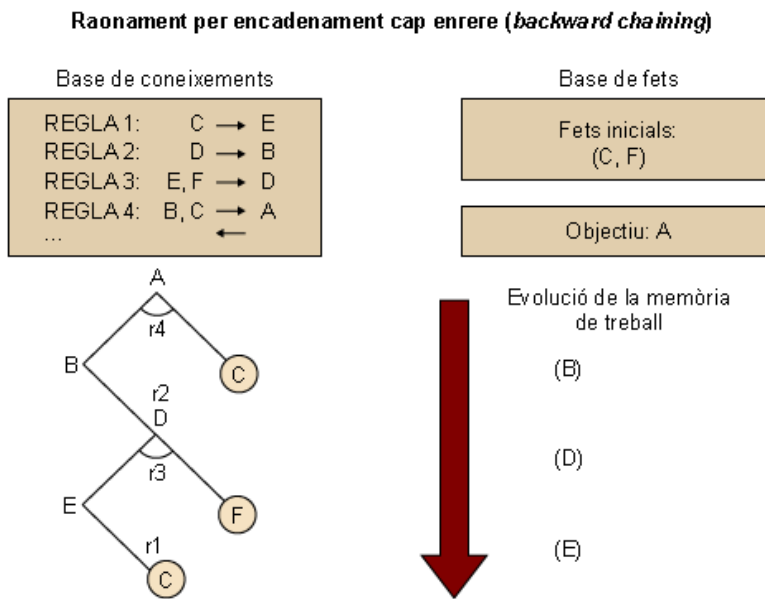


Figura 8

### 5.4.3. Raonament inexacte

Quan les regles poden ser inexactes, el motor d'inferència ha de saber tractar amb els factors de certesa (FC) per a compondre el factor de certesa de la solució proposada. Això es realitza utilitzant regles derivades de la lògica difusa (*fuzzy logic*) que combinen tant les lleis de la lògica com les de la probabilitat.

#### Forma de composició de factors de certesa

En l'exemple següent es mostra una possible forma de composició (producte) de factors de certesa (FC-entrada · FC-regla):

FET: A	[FC = 0,8 ]
FET: B	[FC = 0,7 ]
REGLA: Si A i B, llavors C	[FC = 0,5 ]
Certesa de l'entrada: mínim 0,8 · 0,7 = 0,7	
Certesa de la sortida: 0,7 · 0,5 = 0,35	

que permet establir que C és cert amb un 35% de probabilitat, a partir del 80% de probabilitat assignada a la presència del fet A, i el 70% de probabilitat assignada al fet B. L'entrada té, per tant, una certesa del 70%, mentre que la regla aplicada només és certa en el 50% dels casos. Per això la certesa final queda reduïda al 35%.

## 5.5. Enginyeria del coneixement

Els sistemes experts han portat a l'aparició d'una nova especialitat en la informàtica que se sol etiquetar d'"enginyeria del coneixement". De la mateixa manera que un administrador de bases de dades defineix la base de dades amb un llenguatge de definició de dades (DDL), l'enginyer del coneixement ajuda a definir la base de coneixements amb un sistema adequat de representació del coneixement. El paral·lelisme amb el món de les bases de dades es completa amb la figura de l'usuari del sistema expert, que actua en certa forma com l'usuari d'un sistema de base de dades que pot obtenir, modificar i afegir noves

dades puntuals amb un llenguatge de manipulació de dades (DML). Igualment l'usuari d'un sistema expert afegeix nous fets a la base de fets i obté nous fets deduïts dels que ja existeixen a la base.

L'activitat de l'enginyer del coneixement comporta la interacció amb l'expert humà, que és qui proporciona realment el coneixement que l'enginyer del coneixement estructurarà i ajudarà a introduir a la base de coneixements. En alguns casos, si la interfície de desenvolupament és prou amigable (*user friendly*), una vegada decidit el formalisme de representació del coneixement, l'expert humà és qui alimenta directament la base de coneixements, encara que generalment ho fa amb l'ajuda i supervisió de l'enginyer del coneixement.

### 5.5.1. Funcions de l'enginyer del coneixement

L'activitat de l'enginyer del coneixement el porta a:

- 1) Descompondre un problema en subproblemes discrets.
- 2) Dissenyar un esquema de representació del coneixement per als fets, les regles i el domini que s'han de considerar.
- 3) Caracteritzar les situacions i les conclusions del problema en termes de l'esquema de representació del coneixement escollit.
- 4) Definir les regles que permeten assolir solucions quan una situació particular arriba a ser reconeguda.
- 5) Seleccionar els estils de raonament que s'han d'emprar (encadenament cap endavant o cap enrere, raonament inexacte, etc.).
- 6) Especificar la interfície d'utilització que es farà servir durant les actualitzacions i consultes que realitzi l'usuari del sistema expert.

Per a això, l'enginyer del coneixement ha de disposar de competències pròpies de l'**enginyeria de programari**, és a dir, l'arquitectura del programari dels sistemes experts, els llenguatges que utilitza la IA (Lisp, Prolog, etc.) i les eines disponibles per a la construcció del sistema expert.

L'enginyer del coneixement també ha de ser competent en el que s'anomena **enginyeria del coneixement**, que comporta bàsicament conèixer els sistemes de representació del coneixement i també les tècniques adequades, fins i tot de tipus psicològic, per a extreure l'esmentat coneixement dels experts humans.

### 5.5.2. *Shells* i conjunts d'eines (*toolkits*)

Per al desenvolupament efectiu de sistemes experts és possible utilitzar un **paquet adaptable**, és a dir, una aplicació escrita prèviament en la qual realitzar uns ajustos mínims per a adaptar-la a unes necessitats concretes. A l'extrem oposat és possible un **desenvolupament específic**, utilitzant els llenguatges propis de la IA (Lisp i Prolog generalment) per a construir un sistema amb tots els seus elements i detalls: motor d'inferència, base de coneixements i interfícies.

La dificultat i el cost del desenvolupament específic es poden evitar recorrent als paquets generadors de sistemes experts que es coneixen habitualment com a *shells* o conjunts d'eines (*toolkits*).

En aquest cas es tracta d'un subsistema que inclou un motor d'inferència ja elaborat i, sovint, un únic esquema de representació del coneixement, també prefixat. La construcció del sistema expert es redueix a obtenir el coneixement de l'expert humà –tasca que no és fàcil ni banal– i introduir-lo en el sistema gràcies a interfícies de desenvolupament també preexistents i proporcionades per la *shell* o conjunt d'eines.

En la selecció d'una *shell* o conjunt d'eines és important conèixer l'esquema de representació del coneixement que utilitza (regles, *frames*, etc.), el mecanisme de raonament que incorpora el seu motor d'inferència (encadenament cap endavant o cap enrere), el maquinari en què la *shell* està implementada (ordinadors personals, màquines Lisp, miniordinadors, grans sistemes, etc.) i el caràcter obert o tancat de la seva arquitectura.

Altres elements d'interès són les facilitats d'explicació de les decisions obtingudes pel sistema expert, la possibilitat o no de raonament inexacte, el caràcter d'interpret o compilador de la *shell*, principalment en el moment de l'adquisició del coneixement, etc.

### 5.6. Viabilitat i beneficis d'un sistema expert

Els sistemes experts intenten emular i reemplaçar els experts humans i per això són adequats per a la seva implantació en les organitzacions i situacions en què hi hagi escassetat d'experts o aquests resultin sotmesos a les consultes d'un ampli conjunt d'usuaris molt dispersos geogràficament. També poden suplir l'expert humà en els casos en què la formació d'un expert nou sigui molt costosa i l'ús de la seva experiència no sigui prou freqüent per a justificar-la.

Altres causes que aconsellen recórrer a la implementació d'un sistema expert solen ser el risc de la pèrdua de l'expert humà i l'existència d'un domini en evolució constant.

Pel que fa referència al problema abordat, haurà de ser un problema que no tingui una solució algorísmica directa –si la tingués, es tractaria amb les tècniques de la informàtica tradicional– i que requereixi tasques deductives i inductives al costat de mecanismes heurístics, en un domini sobre el qual els coneixements siguin de tipus predominantment qualitatiu i simbòlic, sovint incomplets i imprecisos i, fins i tot, de vegades, aparentment contradictoris.

### **5.6.1. Requisits de viabilitat d'un sistema expert**

Perquè el sistema expert sigui viable i pugui oferir resultats també és necessari que:

- a) El domini en què actua sigui restringit i perfectament definit.
- b) Els casos (fets, regles i solucions) siguin clarament formalitzables.
- c) Les solucions amb un cert grau d'incertesa siguin acceptables.
- d) El problema sigui d'una complexitat "raonable" i la seva resolució per l'expert no sigui ni massa curta (cas banal) ni massa llarga (cas massa complex), la qual cosa se sol concretar, per exemple, en un nombre de regles inferior al miler i una profunditat d'inferència entre 2 i 50.
- e) Existeixi disponibilitat d'un expert humà, si és possible motivat, i amb una certa capacitat de formalització, ja que, en definitiva, és ell qui proporciona el coneixement.

Aquesta disponibilitat d'un expert humà motivat és un requisit primordial. Es pot dir que no hi pot haver un sistema expert sense un bon expert humà a qui emular.

### **5.6.2. Beneficis que s'han d'obtenir d'un sistema expert**

Un bon sistema expert pot proporcionar molts beneficis tant per a l'expert, al qual teòricament substitueix, com per als usuaris i les organitzacions que el tenen a la seva disposició.

Per a l'**expert humà**, un sistema expert pot ser una eina de gran utilitat pel que fa a formalització, estructuració i validació dels seus propis coneixements, ja que pot veure com aquests coneixements s'apliquen a casos menys freqüents i, potser, amb resultats insospitats. Gràcies al sistema expert, l'expert humà pot quedar reservat per a les consultes menys rutinàries.

Per a l'**usuari**, un sistema expert comporta la disponibilitat immediata de la competència de l'expert; una interactivitat possiblement més fàcil si la interfície d'utilització està ben dissenyada i es pot disposar d'una justificació adequada, tant de les respostes que obté del sistema expert com de les preguntes que el mateix sistema pot anar efectuant.

Per a les **organitzacions** o empreses que utilitzen un sistema expert, hi ha diverses millores evidents: la possible disponibilitat d'un expert en cada punt de consulta, la perpetuació de l'habilitat i competència de l'expert i un augment del nivell general de competència de l'organització.

També hi ha més facilitat per a la formació de nous experts humans, que es poden formar utilitzant el sistema expert i comparant-ne les recomanacions amb les que ells mateixos haurien realitzat.

### **5.7. Problemes en l'ús de sistemes experts**

Actualment la realització de sistemes experts encara presenta serioses limitacions i problemes entre els quals destaquen:

- a) La dificultat en l'adquisició dels coneixements obtinguts de l'expert humà, no sempre prou col·laborador i motivat.
- b) Una aplicació efectiva que continua essent efectiva només en dominis encara molt restringits.
- c) Un comportament molt fràgil a prop de les fronteres del domini, i una gran dificultat perquè el sistema expert detecti el desbordament del seu àmbit de competència.
- d) L'obtenció d'explicacions no sempre rellevants.
- e) Les limitacions serioses per raons d'eficiència que obliguen a restringir el nombre de regles.
- f) Les limitacions implícites a la potencialitat de les tècniques actuals de representació del coneixement.
- g) L'escassa capacitat dels sistemes experts actuals per a l'aprenentatge, la generalització i el raonament per analogia.

També és un problema la dificultat addicional de la integració dels sistemes experts, i en general dels sistemes derivats de la IA, amb la informàtica tradicional, fins ara gestora i dipositària dels "fets".

## 5.8. Alguns sistemes experts famosos

En aquest apartat esmentarem alguns exemples concrets de sistemes experts famosos:

- MYCIN
- MACSYMA
- PROSPECTOR
- DENDRAL
- XCON

### 5.8.1. MYCIN

El MYCIN va ser un dels primers sistemes experts aparegut a mitjan anys setanta.

El va desenvolupar Shortliffe a la Universitat de Stanford i s'ocupa del **diagnòstic d'infeccions de la sang i la seva teràpia**. Fa servir regles de tipus lògic i el seu motor d'inferència és per encadenament cap enrere (*backward chaining*), i és capaç d'emprar factors de certesa.

Posteriorment es va completar amb **TEIRESIAS**, que és una versió "explicativa" del MYCIN que justifica el diagnòstic i la teràpia recomanada.

### 5.8.2. MACSYMA

Derivat tardà del SAINT de Slage, desenvolupat al MIT a mitjan anys cinquanta, el MACSYMA, documentat el 1977, tracta el **càlcul diferencial i integral** amb un motor d'inferència compost d'una sèrie de funcions implementades directament en Lisp.

### 5.8.3. PROSPECTOR

Documentat per Hart (1978) i Duda (1979), s'ocupa de la prospecció i avaluació de jaciments de minerals, particularment coure i urani. Utilitza la inferència probabilística.

#### **5.8.4. DENDRAL**

Desenvolupat a la Universitat de Stanford sota la direcció de Feigenbaum, és capaç d'analitzar l'estructura molecular d'un compost a partir de l'espectograma de masses i altres dades. El seu motor d'inferència usa el mecanisme de *generate and test*, que consisteix a generar casos possibles i provar si compleixen totes les condicions.

#### **5.8.5. XCON**

El XCON és una versió operativa d'un sistema anterior denominat *R1* i, possiblement, el primer sistema expert utilitzat en l'àmbit comercial. Es va utilitzar, en la firma DEC, per a la configuració dels sistemes DEC/VAX i el va desenvolupar J. McDermott de la Universitat de Carnegie-Mellon. Utilitza un motor d'inferència amb encadenament cap endavant (*forward chaining*) i és operatiu des de 1981.

#### **5.8.6. Altres aplicacions**

El camp d'aplicació possible dels sistemes experts és realment ampli. A tall d'exemples no exhaustius cal esmentar-ne la possible utilització en banca (selecció de productes per a clients, anàlisi de riscos i balanços, regles d'auditoria i inspecció, gestió de carteres, etc.), disseny (configuració de sistemes, disseny de circuits etc.), diagnòstic de situació (diagnòstic mèdic, d'avaries de màquines, del funcionament d'una organització, etc.), ensenyament, etc.

## 6. Xarxes neuronals

En la segona meitat de la dècada dels vuitanta es va produir la reactivació d'una tecnologia d'intel·ligència artificial ja estudiada en els anys cinquanta i seixanta, que havia estat pràcticament abandonada a finals dels anys seixanta.

Com a efecte de l'atenció creixent al procés en paral·lel i del recent descobriment d'algorismes d'aprenentatge nous per a les simulacions de sistemes de neurones, es torna a confiar en un dels primers camps d'estudi de la intel·ligència artificial, al qual es fa referència amb l'expressió *xarxes neuronals*.

Avui dia les **xarxes neuronals** constitueixen una tecnologia "renascuda", encara en desenvolupament, en la qual es dipositen grans esperances a l'hora de tractar amb èxit alguns dels problemes clàssics de la intel·ligència artificial, en particular el del reconeixement de formes i de la paraula parlada.

### 6.1. Descripció

A continuació descriurem alguns aspectes més destacats de les xarxes neuronals.

Explicarem quins són els models de xarxes neuronals, quines són les seves característiques, quines aplicacions tenen, i també algunes particularitats.

#### 6.1.1. Models de xarxes neuronals artificials

Una **xarxa neuronal** està constituïda per un nombre variable de processadors interconnectats entre si i que realitzen una transferència mútua de valors anomenats *activacions*.

Cada processador rep una sèrie d'activacions (**activacions d'entrada**) i, a partir d'elles, genera un valor de sortida (**activació de sortida**) que, al seu torn, transfereix a un altre grup de processadors que té connectats.

Per analogia amb les xarxes biològiques, es dóna el nom de sinapsi a les interconnexions, i **neurones** o **unitats** als processadors que formen els nodes de la xarxa de processadors, que rep el nom de **xarxa neuronal**.

Aquesta descripció funcional s'aplica tant a les xarxes neuronals biològiques com a la seva modelització elèctrica, és a dir, les **xarxes neuronals artificials** (ANN - *Artificial Neural Networks*) amb què es pretén reproduir, mitjançant sistemes electrònics, l'estructura i el comportament de les xarxes neuronals biològiques.



El costum ha portat a ometre el qualificatiu "artificial" i, de forma general, es parla de xarxes neuronals fent referència a la modelització electrònica de les xarxes neuronals biològiques.

### 6.1.2. Algunes característiques de les xarxes neuronals

En el funcionament de les xarxes neuronals és possible distingir la fase d'execució i la fase d'aprenentatge.

En la **fase d'execució** cada neurona realitza una actualització de l'activació de sortida, d'acord amb les activacions d'entrada que rep en un moment donat. El funcionament durant aquesta fase se sintetitza en les equacions que regeixen la dinàmica del sistema i el comportament de cada neurona en un moment donat.

Abans d'usar o "executar" la xarxa neuronal, en la **fase d'aprenentatge**, es modifiquen i actualitzen els pesos que afecten les diverses activacions d'entrada amb l'objectiu d'"ensenyar" a la xarxa neuronal a comportar-se segons es desitja, és a dir, a produir una sortida determinada d'acord amb l'entrada rebuda.

És possible caracteritzar les xarxes neuronals per:

- a) **L'arquitectura de la xarxa**, determinada per la topologia i l'estructura que té i pel grau de connexió entre les neurones que la formen.
- b) **La dinàmica de la xarxa**, és a dir, les equacions que regeixen el comportament de cada neurona, i en conjunt de la xarxa, durant la fase d'execució, és a dir, quan la xarxa ja s'ha "ensenyat".
- c) **La regla, procediment o algorisme d'aprenentatge** amb què s'"ensenyava" a la xarxa la seva comesa.

### 6.1.3. Altres denominacions

Encara que el nom més estès és el de xarxes neuronals, els models com l'indicat també han rebut altres denominacions. A causa de la interconnexió de diversos processadors simples se'ls ha anomenat **models connexionistes** i, de fet, representen el suport d'aquest plantejament al camp de la intel·ligència artificial.

A causa del paral·lisme de l'actuació conjunta de diversos processadors, també es parla de **models de procés en paral·lel distribuït** com es va fer en un llibre famós de 1986, *Parallel Distributed Processing: Explorations in the microstructure of Cognition*, editat per D.E. Rumelhart i J.L. McClelland, que ha servit per a la reactivació de l'interès sobre les xarxes neuronals.

Per la similitud estructural amb el sistema nerviós biològic se'ls ha denominat també **sistemes neuromòrfics** (*neuromorphic systems*) i es parla també de **càlcul neuronal** (*neural computing*) per a referir-se a l'aspecte de la informàtica i de la intel·ligència artificial que tracta de les xarxes neuronals i les seves possibilitats.

#### 6.1.4. Aplicacions i funcionament

Les xarxes neuronals s'utilitzen principalment en els sistemes classificadors que han de donar resposta a un estímul concret per a reconèixer la seva pertinença o no a una classe determinada.

En concret, sembla que les xarxes neuronals són el mecanisme que mostra més possibilitats davant de problemes com **el reconeixement de formes i el reconeixement de la paraula parlada** i, en general, davant de tot tipus de problemes que requereixin una alta capacitat de tractament efectuat en paral·lel. En lloc d'executar una seqüència d'instruccions prèviament emmagatzemades, com fan els ordinadors que utilitzen l'arquitectura clàssica deguda a Von Neumann, els models amb xarxes neuronals exploren simultàniament diverses hipòtesis fent servir àmpliament el procés en paral·lel, per la qual cosa també han servit per a impulsar aquest aspecte de l'arquitectura dels ordinadors.

Se sol dir que la referència a les xarxes neuronals es distingeix de la informàtica clàssica perquè no es "programa" una xarxa neuronal, sinó que s'"ensenya" a una xarxa neuronal a comportar-se d'una manera determinada.

La manera habitual de fer-ho és mostrar a la xarxa neuronal diversos senyals d'entrada conjuntament amb les sortides esperades, i repetir el procés més d'un cop mentre la xarxa neuronal "aprèn" ajustant els pesos i les ponderacions que determinen la dinàmica de la xarxa i el comportament de les diverses neurones i sinapsis que la formen. El problema queda resolt quan s'ha trobat un conjunt de pesos que permeten a la xarxa compondre la sortida adequada per a cada entrada, és a dir, "reconèixer" l'entrada.

#### 6.1.5. Neuroordinadors i *netware*

La modelització d'una xarxa neuronal pot tenir lloc al maquinari utilitzant circuits de procés en paral·lel i, a tal efecte, s'utilitzen arquitectures especials per al procés en paral·lel com, per exemple, els ordinadors matricials. També es dissenyen xips VLSI que incorporen l'estructura d'una xarxa neuronal. Generalment s'assigna el nom de neuroordinadors (*neurocomputers*) a aquests sis-

temes, i se solen utilitzar com a coprocessadors especialitzats i auxiliars dels ordinadors digitals tradicionals. En aquest cas s'accedeix a la xarxa neuronal com si fos una subrutina del sistema.

En el cas dels *neurocomputers* no té sentit parlar de la seva potència fent referència als milions d'instruccions per segon (mips) que són característics dels ordinadors amb arquitectura Von Neumann. Als neuroordinadors es parla d'**interconnexions per segon** i, per exemple, el sistema Neuro-07 anunciat per NEC el 1988 ja incorporava un neuromotor (*neuro-engine*) de 210.000 interconnexions per segon. El sistema incloïa un ordinador personal convencional amb pantalla de color, un neuromotor amb el seu teclat particular i programari per a "ensenyar" a la xarxa neuronal, per un preu de poc més d'un milió de pessetes (uns 6000 euros).

El *netware* és un programari que simula les neurones i la seva interconnexió en un ordinador convencional amb arquitectura Von Neumann.

El mateix any a què fèiem referència, el 1988, es van vendre uns 10.000 paquets de *netware* neuronal als Estats Units.

Fins i tot amb estacions de treball, i també amb els ordinadors personals de la microinformàtica, és possible la investigació i utilització de petites xarxes neuronals (menys de cinquanta elements processadors o neurones en un PC compatible, segons alguns autors).

Tot i que s'han de tenir en compte les limitacions de procés de l'arquitectura Von Neumann sotmesa a les exigències de simular el procés en paral·lel de les xarxes neuronals, i que és important disposar de compiladors adequats per a la simulació del procés en paral·lel, els PC resulten ser adequats per a l'estudi i la investigació amb xarxes de dimensions reduïdes.

### 6.1.6. Característiques particulars de les xarxes neuronals

Gràcies a l'elevat grau de paral·lisme, les xarxes neuronals presenten una tolerància a errors elevada (*fault tolerance*) a causa de la redundància d'elements processadors i a les característiques de la seva dinàmica de funcionament. Per això també presenten un procés de degradació controlat i limitat (*graceful degradation*) quan es presenta l'error d'un dels processadors.

A més, cal recordar que les xarxes neuronals no són eficients per a realitzar les tasques habitualment encomanades als ordinadors de procés seqüencial (càlculs amb gran exactitud, manipulació de xifres i dades, etc.), però resulten millors per al maneig de conjunts incomplets de dades o d'informació difusa o contradictòria.

#### La dècada dels anys 80

A finals de la dècada dels anys 80, dades publicades l'agost de 1989 estimaven ja en unes 300 les empreses nordamericanes especialitzades en la tecnologia de les xarxes neuronals, i diverses d'elles es dediquen, precisament, a elaborar *netware* per a ordinadors personals i estacions de treball.

## 6.2. Arquitectura i topologia de les xarxes neuronals

A continuació explicarem quina és l'estructura topològica i l'arquitectura en les xarxes neuronals.

### 6.2.1. Estructura topològica

Pel que fa a l'arquitectura i topologia de la xarxa es pot parlar, per exemple, de xarxes amb alimentació cap endavant (*feed forward*) o de xarxes amb retroalimentació (*feedback*). En les primeres, el sentit de propagació de les activacions en la fase d'execució és únic (cap endavant), mentre que en les segones, les activacions de sortida d'unes neurones s'utilitzen també com a activacions d'entrada en un circuit de retroalimentació (vegeu la figura 9 al marge).

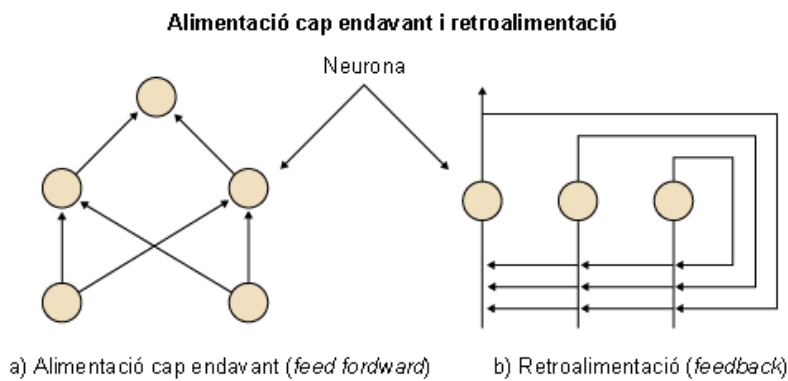


Figura 9

També es poden dissenyar xarxes en les quals es formen estructures jeràrquiques entre les neurones o utilitzar qualsevol altra topologia.

Un cas particular d'aquestes xarxes és el que s'ha anomenat xarxes d'estructures competitives, en el qual les neurones s'agrupen en grups de sectors (*clusters*), i dins de cada grup de sectors, tan sols una neurona pot estar activa en un moment determinat gràcies al caràcter inhibidor de les sinapsis entre elles (vegeu la figura 10 al marge).

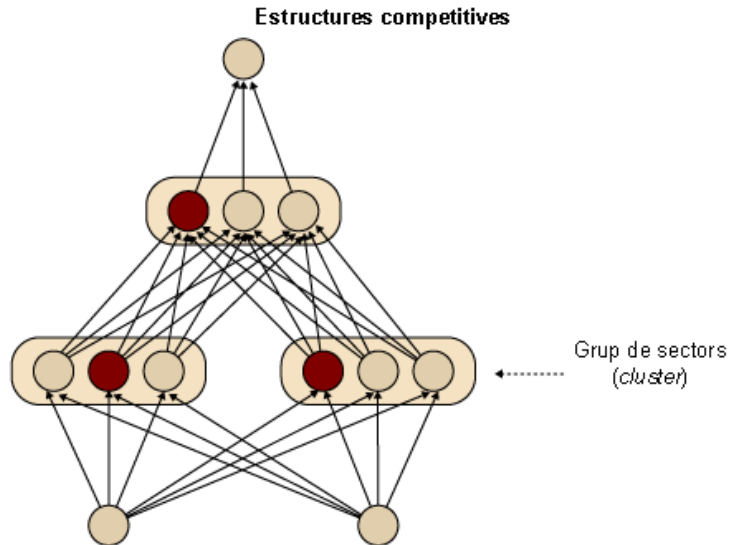


Figura 10

### 6.2.2. Estructura en capes d'una xarxa neuronal

En el cas més utilitzat, les neurones en una xarxa neuronal solen estructurar-se de manera que integrin una o diverses "capes" o nivells (*layers*). Segons els dissenys, hi ha xarxes d'una, dues i diverses capes.

La primera capa s'anomena **capa d'entrada** (*input layer*) i és la que rep els estímuls de l'exterior. Si hi ha més capes, l'última d'elles és la **capa de sortida** (*output layer*), de la qual s'obté el resultat de la fase d'execució de la xarxa neuronal. Si hi ha més de dues capes, entre la primera i l'última hi ha un conjunt de capes ocultes (*hidden layers*) formades per neurones que emmagatzemen la informació en una "representació interna", i per això també reben el nom d'unitats de representació interna (*internal representation units*) (vegeu la figura 11 al marge).

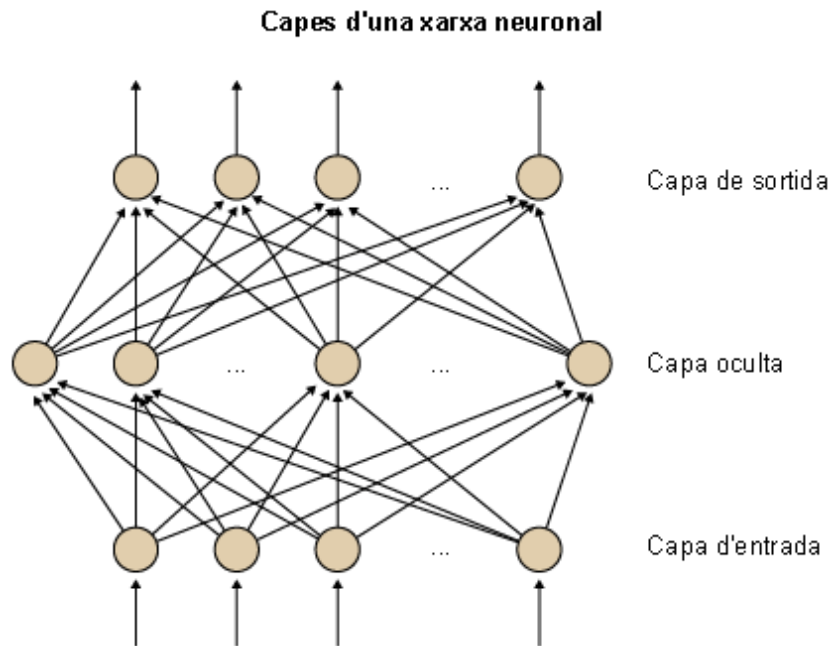


Figura 11

Les capes poden estar constituïdes per un nombre qualsevol de neurones, normalment imposat per les característiques del problema que s'ha de resoldre o per les prestacions que es desitgi obtenir de la xarxa neuronal. Així doncs, en el disseny d'una xarxa neuronal es donen dos primers graus de llibertat en escollir tant el nombre de capes com el nombre de neurones a cada capa.

### Les dues espirals

Un exemple possible és el que es mostra a la figura 12, amb dues neurones d'entrada, una de sortida i dues capes ocultes de cinc neurones cada una. K. Lang i M. Witbrock, de la Carnegie Mellon University, van utilitzar aquesta xarxa per a resoldre el problema conegut com "de les dues-espirals", que consisteix a reconèixer si un punt pertany o no a dos espirals entrelaçats. Gràcies a la simplicitat del model, ha estat possible estudiar-hi el comportament de les capes ocultes en una xarxa neuronal, i aprofundir en el coneixement de la representació interna de la informació en una xarxa amb més de dues capes, un dels problemes més interessants i suggestius en la investigació sobre xarxes neuronals.

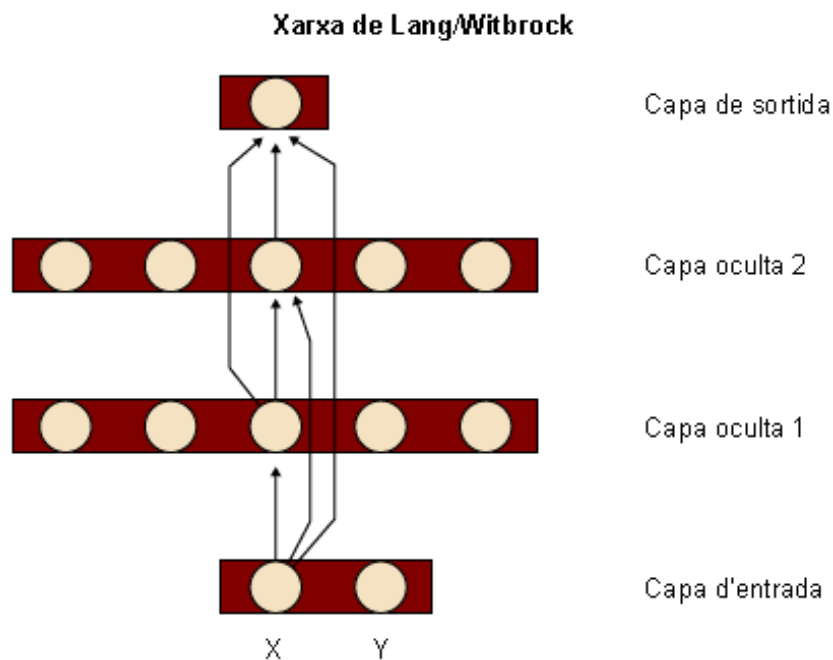


Figura 12

### 6.2.3. Connexionat de neurones

Les sinapsis o connexions entre les neurones que formen les diverses capes de la xarxa es poden realitzar de diverses maneres, la qual cosa influeix en la topologia i arquitectura de la xarxa i en el seu funcionament.

El sistema més habitual segueix el principi de la connectivitat total segons el qual cada neurona pertanyent a una certa capa està connectada amb totes i cada una de les neurones que formen la capa immediatament inferior i de totes elles rep activacions d'entrada.

A més, cada neurona d'una capa està connectada també amb totes i cada una de les neurones que integren la capa immediatament superior, a les quals transmet la seva activació de sortida. Així es mostrava al connexionat de la figura 11, que repetim al marge, amb una única capa oculta.

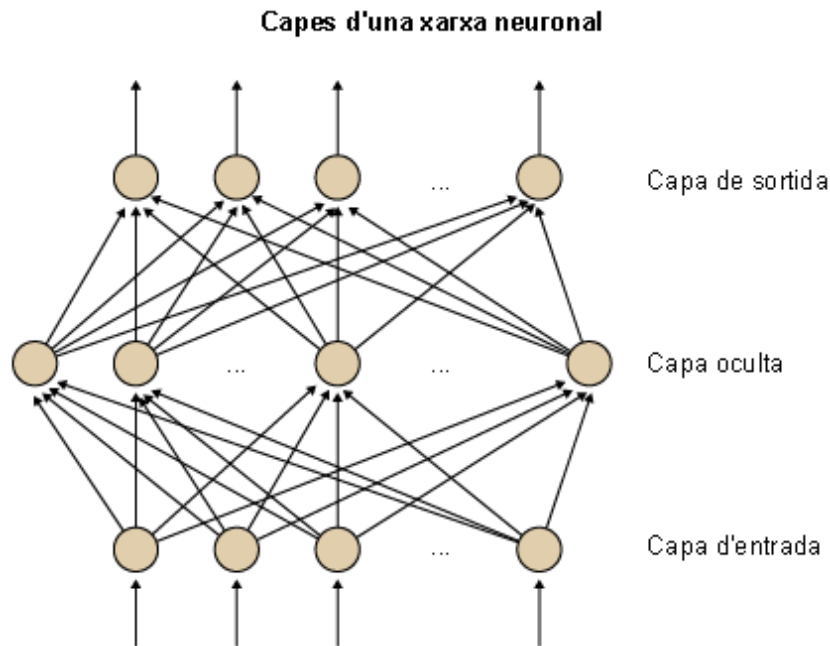


Figura 11

### 6.3. Dinàmica i aprenentatge d'una xarxa neuronal

A continuació explicarem la dinàmica i l'aprenentatge d'una xarxa neuronal en dues fases: fase d'execució i fase d'aprenentatge.

#### 6.3.1. Fase d'execució

En la fase d'execució cada interconnexió o sinapsi té un pes establert obtingut en la fase d'aprenentatge.

La dinàmica de funcionament de la xarxa neuronal estableix que el senyal d'entrada a cada neurona és afectat pel pes esmentat en una combinació lineal de totes les activacions d'entrada que rep la neurona. Un cop efectuada aquesta combinació lineal, s'afegeix al resultat una certa funció llindar característica de la neurona. Al resultat obtingut s'aplica una funció pròpia de la neurona que és la que determina finalment l'activació de sortida. En el cas, avui dia freqüent, de diverses capes de neurones connectades seguint el principi de la connectivitat total, la fórmula que determina el comportament dinàmic de la xarxa neuronal és:

$$a_k(c) = f\left(\sum [W_{ki}(c) * o_k(c-1)] + u_k(c)\right)$$

on:

$a_k(c)$  és l'activació de sortida de la neurona de la columna  $k$  a la capa  $c$ ;

$W_{ki}(c)$  és el pes assignat a la interconnexió entre la neurona de la columna  $k$  de la capa  $c$  i la neurona de la columna  $i$  de la capa  $c-1$ ;



$o_k(c-1)$  és l'activació d'entrada per aquesta interconnexió i és igual a  $i_k$  (el senyal d'entrada a la columna  $k$  de la capa d'entrada) quan la capa és la primera ( $c = 0$ ), i igual a  $a_k(c-1)$  a les altres capes;

$u_k(c)$  és la funció llindar per a la neurona de la columna  $k$  de la capa  $c$ ; i finalment,

$f()$  és la funció d'activació de la neurona.

Tot això permet esquematitzar gràficament una neurona amb les seves activacions en el diagrama de la figura 13 al marge.

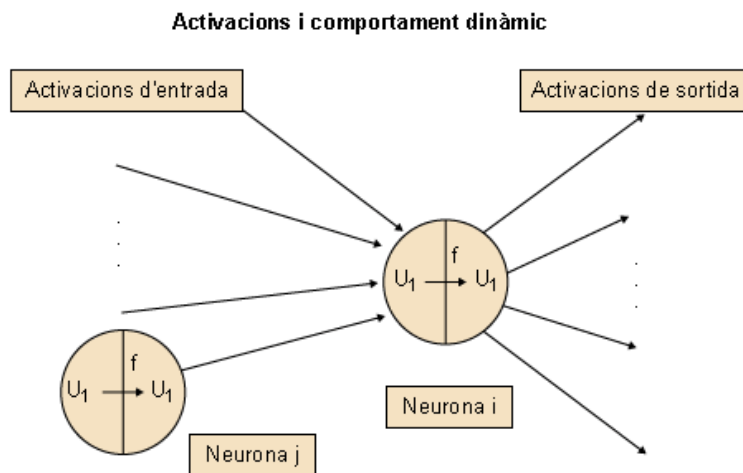


Figura 13

La funció  $f()$ , que defineix el comportament d'una neurona en una xarxa neuronal, pot ser **determinista** o **estocàstica** segons que el potencial postsinàptic o activació de sortida s'obtingui de forma determinada de l'entrada o d'acord amb una distribució probabilística.

Les funcions deterministes poden ser lineals com  $f(x) = x$ , però en la majoria dels casos són no lineals. Les funcions més utilitzades són:

- a) Funcions de tipus binari si la resposta és 0 o 1.
- b) Funcions isign o *hard delimiter* si la resposta és +1 o -1, és a dir, del tipus  $f(x) = \text{sign}(x)$ .
- c) Funcions sigmoïdes, generalment segons la tangent hiperbòlica en la forma  $f(x) = \text{th}(x)$ .

La fase d'execució pot ser paral·lelitzada fàcilment (fins i tot en una arquitectura Von Neumann) mitjançant una partició del procés de càlcul. Això s'aconsegueix utilitzant un processador (o la seva simulació) per a cada neurona, i el processador esmentat s'encarrega d'emmagatzemar en una memòria local els pesos de ponderació associats a la neurona que ha de simular. Per

a això, cada processador pren l'activació de la capa immediatament inferior i la multiplica pel pes corresponent. L'activació obtinguda es transfereix processador següent mentre es rep la nova activació procedent del processador anterior que, al seu torn, es multiplica (\* en la fórmula) pel pes corresponent i s'acumula (S en la fórmula) al resultat anterior. Aquest procés de multiplicació i acumulació es repeteix fins que s'obtenen tots els productes parcials indicats en l'equació.

### 6.3.2. Fase d'aprenentatge

En la fase d'aprenentatge s'"ensenya" a la xarxa neuronal el comportament desitjat.

Les regles d'aprenentatge poden ser supervisades quan existeix un "professor" que dirigeix l'entrenament de la xarxa i indica si la resposta a l'estímul d'entrada és correcta o no; o autoorganitzades quan no hi ha un "professor" extern a la xarxa. En aquest últim cas es parla, més freqüentment, d'una regla d'aprenentatge no "supervisada".

Hi ha diversos algorismes o regles per a procedir a l'actualització dels valors dels pesos de les interconnexions o sinapsis durant el procés d'entrenament de la xarxa. Cada algorisme dóna lloc a una certa estructura i comportament de la xarxa neuronal.

El problema central en l'eficàcia d'un algorisme d'aprenentatge és determinar quines són les interconnexions que tenen una modificació de pesos que comporta una millora ràpida de l'eficiència de la xarxa neuronal. El problema resulta especialment complex quan es tracta d'interconnexions en xarxes neuronals de diverses capes i hi ha "capes ocultes" que mantenen una "representació interna" desconeguda de la informació que circula per la xarxa neuronal. Precisament, la difusió al final dels anys vuitanta d'un d'aquests algorismes d'aprenentatge es considera una de les causes principals del renaixement de l'interès per les xarxes neuronals.

L'esmentat algorisme és conegut com el de la propagació d'errors cap enrere (*back propagation of errors*), però sol abreuja-se en la literatura especialitzada com a *back propagation* i, fins i tot, com a *back-prop*. És un algorisme supervisat que va proposar D.E. Rumelhart i es pot aplicar a xarxes de diverses capes, tot i que habitualment es fan servir xarxes de tres capes en la majoria d'estudis i investigacions. Aquest algorisme presenta, entre altres, l'avantatge de permetre un ampli rang d'aplicacions, principalment a causa del fet que també accepta entrades i sortides contínues.

#### La regla d'aprenentatge

En alguns casos es prescindeix del procés d'aprenentatge i es diu que la regla d'aprenentatge és "instantània". Però, en la majoria dels casos, la regla o algorisme d'aprenentatge és de gran importància.

En primer lloc s'estableixen a l'atzar els valors dels pesos i s'ofereix una entrada a la xarxa neuronal que produirà activitat en cada una de les capes de la xarxa. Un "professor" que coneix quina ha de ser la resposta de cada neurona de la capa de sortida, indica en cada una d'elles l'amplitud i el signe del seu error. Els esmentats senyals d'error s'utilitzen per a ajustar els pesos de totes les interconnexions entre la capa de sortida i la capa oculta immediatament anterior a ella, de manera que l'error obtingut a la capa de sortida es propaga cap enrere, cap a les capes ocultes que, al seu torn, utilitzen els valors d'error rebuts per a ajustar el valor de les seves sinapsis. Així doncs, en la fase d'aprenentatge el flux de dades té un sentit contrari al de la fase d'execució i es propaga "cap enrere".

De fet, l'algorisme aconsegueix realitzar petits ajustos en cada sinapsi de la xarxa, de manera que el conjunt de les modificacions redueix l'error total en el funcionament de la xarxa. Un cop s'ha aplicat el mateix procediment diverses vegades, s'arriba a un valor mínim de l'error total. Com tots els mètodes de gradient descendent, es podria tractar d'un mínim local, però això rarament passa a les xarxes neuronals, possiblement pel nombre de neurones involucrades i la gradualitat de la seva resposta.

#### **6.4. Història de les xarxes neuronals**

A continuació farem una breu introducció a la història de les xarxes neuronals, des dels primers estudis que es van fer sobre el tema fins al nou interès que ha despertat en els nostres temps.

##### **6.4.1. El *perceptron* de F. Rosenblatt**

Els primers estudis sobre xarxes neuronals daten de 1943, quan McCulloch i Pitts van iniciar la interpretació mecanicista de les xarxes neuronals. Però només a partir dels anys seixanta les xarxes neuronals van començar a ser considerades com una possible solució als problemes de realitzar classificacions automatitzades. Per a això va ser decisiu el treball de **Franz Rosenblatt**, un psicòleg de la Universitat Cornell, que per mitjà de l'anàlisi matemàtica, la simulació en ordinadors digitals i alguns experiments amb sistemes analògics de procés en paral·lel va arribar a mostrar com es podia ensenyar a actuar com a classificadors a les xarxes neuronals amb valors variables als pesos de les interconnexions.

El *perceptron*, desenvolupat per Rosenblatt a partir de 1958, era en realitat un classificador de categories aplicat a la simulació de la visió humana. Rosenblatt proposava una regla d'aprenentatge que, encara que era lenta, havia de convergir sempre que existís una solució. En realitat, el model de *perceptron* amb dues capes de neurones podia computar, per exemple, les funcions lògiques *or* i *and*, però no la funció *xor*. La introducció d'una tercera capa, una capa oculta, aportaria la potència suficient per a implementar aquesta funció. Però un article de Minsky i Papert de 1967 demostrant que el procediment

de convergència de Rosenblatt només podia resoldre problemes linealment separables, va fer que a la pràctica s'abandonés durant diversos anys aquesta via d'investigació.

#### **6.4.2. El predictor meteorològic de Widrow**

Malgrat el descrèdit que el treball de Minsky i Papert va aportar a la teoria dels perceptrons, un altre dels pioners que hi treballava des dels anys cinquanta, **Bernard Widrow**, de la Universitat de Stanford, va implementar el 1963 una xarxa neuronal que actuava com un **senzill sistema de predicció meteorològic** capaç de donar prediccions correctes en un 85% dels casos, davant l'estimat 65% que obtenia l'home del temps local. Widrow utilitzava una regla d'aprenentatge amb una funció de "cost" que havia de ser minimitzada. La funció de cost esmentada mesurava la diferència entre la sortida real de la xarxa neuronal a la seva capa de sortida i el valor esperat.

Però la realitat és que, després de l'influent article de Minsky i Papert, els perceptrons i les xarxes neuronals van ser pràcticament abandonats durant els anys setanta, i fins i tot es va arribar a dubtar de la possibilitat de continuar considerant-los com a objectes vàlids d'investigació.

#### **6.4.3. El nou interès per les xarxes neuronals**

Es pot dir que l'interès per les xarxes neuronals va renéixer l'any 1982 a partir de l'èxit dels treballs de **J.J. Hopfield** sobre xarxes neuronals. El model de Hopfield va introduir les xarxes neuronals que seguien el principi de la connectivitat total i utilitzava senyals binaris i amb una arquitectura realimentada. La xarxa neuronal de Hopfield, en les seves diverses versions, es pot utilitzar com una forma de memòria associativa i també com un sistema per a resoldre problemes d'optimització.

Un altre sistema amb gran èxit i influència va ser la xarxa **NETTalk**, presentada el 1986 per T.J. Sejnowski i C.R. Rosenberg, de la Universitat John Hopkins. Es tracta d'una xarxa neuronal de tres capes que té com a objectiu aprendre a pronunciar l'anglès. L'entrada és text que ha de pronunciar una màquina que pot produir sons. Encara que al principi la xarxa balbujeja i s'equivoca, gràcies a l'entrenament i la correcció dels pesos de les interconnexions, acaba pronunciant correctament, fins i tot en intentar-ho amb textos amb què no ha estat entrenada (només registra un 10% d'errors, possiblement perquè la pronunciació anglès de vegades depèn del context).

Amb aquest sistema també es va demostrar que el nombre de neurones en capes ocultes és de gran importància: poques neurones impedeixen que la xarxa neuronal compleixi la seva funció, però amb un excés de neurones a la capa

oculta la xarxa pot funcionar una mica millor, malgrat que no sap generalitzar (en el cas de la NETTalk, és incapaç, per exemple, de llegir textos que no hagi preparat prèviament).

També es considera de gran influència la difusió el 1986 de l'algorisme *back propagation* i el treball recapitulatiu sobre el procés paral·lel distribuït editat per Rumelhart i McClelland. Nous investigadors com Feldman ( propietats dels models connexionistes), Grossberg (el cervell adaptatiu) i molts altres han proporcionat un gran impuls a l'estudi de les xarxes neuronals.

Els fets que resumeixen el nou interès se centren en el desenvolupament noves topologies i algorismes, la possibilitat d'implementacions en circuits VLSI, els èxits de sistemes com el NETTalk de Sejnowski i Rosenberg i de mateixes xarxes d'Hopfield i, també, la creixent fascinació sobre el funcionament del cervell humà que les xarxes neuronals ajuden a estudiar. En qualsevol cas, la xarxes neuronals són, altre cop, una tecnologia en creixement de qual encara resulta prematur fer hipòtesis sobre el seu futur.

## 6.5. Les xarxes neuronals i el cervell humà

La similitud estructural entre les xarxes neuronals biològiques i les xarxes neuronals artificials encoratja a utilitzar aquestes últimes per a entendre millor el cervell humà i la dinàmica del seu funcionament.

En qualsevol cas, l'analogia no pot ser completa i cal tenir en compte molts factors de diferenciació com, per exemple, les evidents limitacions de mida de les xarxes artificials. En un cervell humà s'estima que hi ha al voltant de  $10^{11}$  neurones que presenten una connectivitat mitjana de l'ordre de 1.000; això situa el nombre d'interconnexions o sinapsis en  $10^{14}$ , xifres per ara inabastables amb els models artificials que es veuen reduïts a casos exageradament simples.

D'altra banda, l'algorisme de propagació cap enrere (*back propagation*), que tant d'èxit ha tingut en les xarxes neuronals artificials, no sembla realista en l'àmbit del cervell humà, en què l'activació de sortida d'una neurona (almenys al neocòrtex) és o bé excitadora o inhibidora, però no totes dues alhora.

Un altre problema important que sorgeix en assimilar el comportament de les xarxes neuronals artificials amb el del cervell humà és l'exigència d'una part diferenciada del cervell humà, la que faria de "professor" en el procés d'aprenentatge, ja que hauria de tractar-se de neurones amb propietats noves i específiques que sembla que no existeixen en el cervell humà.

Aquestes raons i moltes altres que no fa falta esmentar aquí fan pensar a alguns especialistes com F. Cricks que aquests suposats "models" neuronals no modelitzen res real en el cervell humà, i que un altre cop es tracta de sistemes que "manifesten" en alguns casos un comportament similar al cervell humà, sense que d'això es dedueixi que es tracti d'un model real.

De fet, encara que les xarxes neuronals biològiques no utilitzin, per exemple, un algorisme com el de propagació cap enrere, el cert és que el comportament final és semblant en alguns casos. Cal no oblidar que les xarxes neuronals resulten el sistema artificial més eficient per al reconeixement de formes i, curiosament, aquesta és una activitat cerebral molt senzilla que fins i tot els animals són capaços de realitzar fàcilment, encara que, per la gran potència de càlcul que exigeix, ha fracassat l'intent de resoldre-la amb sistemes artificials (sovint basats en l'abstracció matemàtica) que no siguin les xarxes neuronals.

## 7. Llenguatges específics per a la IA

Fem constar tot seguit els llenguatges específics per a la intel·ligència artificial més rellevants.

### 7.1. Lisp

És freqüent la consideració que el Lisp és el llenguatge propi de la intel·ligència artificial a causa de les seves característiques particulars, la seva existència des de 1962 i per haver estat àmpliament utilitzat pels investigadors en el camp de la IA.

#### 7.1.1. Origen

El Lisp es va desenvolupar al MIT arran d'un treball publicat per John McCarthy el 1960: *Recursive functions of symbolic expressions and their computation by machine*. Posteriorment, el 1962, es va publicar el primer manual del Lisp per a l'equip IBM 7090 i els seus autors van ser el mateix McCarthy juntament amb Abrahams, Edwards, Hart i Levin, tots ells del MIT. El llenguatge utilitza també les anomenades funcions "lambda" sorgides del *lambda calculus*, un llenguatge formal sobre funcions desenvolupat per A.Church.

#### 7.1.2. Objectius i característiques principals

El Lisp implementa el tractament mitjançant expressions simbòliques i recursivitat, i disposa d'una gran potència i simplicitat en la definició d'estructures. Al Lisp no hi ha diferència entre les estructures de dades i les estructures de programa; totes dues es redueixen a **expressions simbòliques** que tenen com a casos particulars els **àtoms** i les **l·listes**. D'aquí el nom de Lisp derivat de *LIST Processor* (processador de l·listes).

En realitat, el Lisp es pot tractar com un llenguatge matemàtic formal que, amb les extensions adequades, es converteix en un llenguatge de programació. Potser per això és un llenguatge summament flexible, i fins i tot el pot canviar el mateix programador.

Es tracta, en els seus inicis, d'un **llenguatge funcional** que no té per què disposar de la instrucció d'assignació, els programes de la qual són funcions que es defineixen per composició d'altres funcions més simples. Per a executar-lo, s'"aplica" a les dades d'entrada, que són els paràmetres de la funció, per a obtenir un resultat, que és el valor calculat de la funció.

Les variants del Lisp més utilitzades a la pràctica també incorporen molts elements no funcionals en nom de l'eficiència.

### 7.1.3. Dialectes

Un problema habitual per als usuaris del Lisp és que no existeix cap estandarització vàlida i hi ha diversos dialectes en ús. Es reconeix un nucli central, anomenat de vegades Pure Lisp, i variants com el Maclisp, utilitzat al MIT; el PSL, amb pretensió d'estandarització portable desenvolupat a la Universitat d'Utah; el Le-Lisp, versió utilitzada a França, etc.

Els intents d'estandarització es van iniciar sense gaire èxit amb l'Standard Lisp d'Hearn, el 1966, i van arribar a un consens amb el Common Lisp aparegut el 1984. També s'han de destacar els entorns de programació Lisp, d'entre els quals el Maclisp és el que es va convertir en un dels més coneguts en ampliar-se amb variants com el Franzlisp o el Zetalisp. També és possible esmentar l'Interlisp, de la Universitat de Cambridge.

El llenguatge **Logo** és, en el fons, una derivació simplificada del Lisp per a la seva utilització en l'ensenyament, en s'utilitza per a crear els **entorns de descobriment**.

## 7.2. Prolog

El Prolog és un altre dels llenguatges específics més rellevants que hi ha sobre intel·ligència artificial.

### 7.2.1. Orígens

A començament dels anys setanta, els partidaris de la programació lògica es van dedicar a implementar a l'ordinador un sistema que gestionés la lògica causal que és, en realitat, una restricció de la lògica de predicats de primer ordre (clàusules d'Horn). El llenguatge que implementava l'esmentada **PROgramació LÒGica** es va anomenar Prolog i va néixer a la Universitat de Marsella, sota la direcció d'Alan Colmerauer.

Algunes fites fonamentals en la història del Prolog són un primer programa per a la prova de teoremes escrit el 1973 per Colmerauer en FORTRAN; la formulació completa de la programació lògica de Robert A. Kowalski el 1974; la implementació d'un Prolog eficient en un PDP-10 realitzada el 1977 per David H.D. Warren i, finalment, l'estandarització de l'anomenada **sintaxi d'Edimburg** al llibre *Programming in Prolog*, de Cloksin i Mellish el 1981.

Un element addicional a favor de la difusió i la fama del Prolog va ser el seu ús anunciat en el projecte japonès de la "cinquena generació" d'ordinadors.



### 7.2.2. Objectius i característiques principals

La implementació de la lògica causal s'aconsegueix amb una representació del coneixement a base de fets i regles amb la forma:

FETS:  $P(X_1, \dots, X_n)$

REGLES:  $P(X_1, \dots, X_n) :- P_1(\dots), \dots, P_m(\dots)$

(on ":-" es llegeix com "si")

que constitueixen els elements emmagatzemats en la base de dades (coneixements en forma de regles i fets en forma de fets) del sistema Prolog.

Per tant, es tracta d'un sistema de producció basat en regles i és un **llenguatge declaratiu** en el qual s'indiquen els fets i les regles i no el procediment (algorisme) amb què es troben les solucions.

L'usuari proporciona els fets i les regles i el Prolog incorpora un motor d'inferència que implementa els mecanismes de recursivitat, unificació i raonament per encadenament cap enrere (*backtracking*). Això es fa patent en la forma d'utilitzar el Prolog una vegada implementada la base de fets i regles: preguntar si és cert o no un objectiu determinat (*goal directed*). Tot això fa que el Prolog sigui especialment adequat per a la implementació de sistemes experts basats en regles, i en els quals sigui correcte l'ús d'aquest mecanisme de raonament.

### 7.2.3. Dialectes

La problemàtica dels molts dialectes del Lisp no apareix al Prolog, en l'ús del qual és freqüent utilitzar la sintaxi d'Edimburg, establerta el 1981 al llibre ja esmentat de Cloksin i Mellish.

## 7.3. Smalltalk

Encara que el llenguatge Smalltalk es va crear inicialment per a altres objectius, també s'ha considerat amb gran interès en l'àmbit de la IA. Es tracta d'un llenguatge concebut com a suport de la programació orientada a objectes i que presenta l'avantatge de complir els requisits generals d'un llenguatge per a IA: incorpora tractament simbòlic i disposa d'una estructura d'*array* molt similar a les llistes del Lisp.

Gràcies a la seva adaptació a la programació orientada a objectes, l'Smalltalk disposa de mòduls que es transmeten **missatges** que vénen a ser ordres d'execució d'operacions. Cada **objecte** es defineix mitjançant el conjunt de **variables** que constitueixen el seu estat intern i dels **mètodes** associats a les operacions que serviran per a manipular les variables esmentades.

També és de gran importància la facilitat de l'Smalltalk per a utilitzar el concepte d'**herència**. La definició d'un objecte a l'Smalltalk comporta indicar la seva classe, ja que tot objecte és una *instància* d'una classe. Les classes s'organitzen jeràrquicament i tota subclasse "hereta" les variables i els mètodes de la classe superior.

Alguns autors aposten pel futur d'un sistema anomenat **Loops** que integra l'ús de l'Smalltalk amb el Lisp i el Prolog, encara que qualsevol decisió és prematura en aquest sentit.

## Activitats

Una activitat complementària a l'abast de molts és atrevir-se a jugar als escacs contra un programa informàtic. Sense necessitat d'arribar a la sofisticació del *Deep Blue*, el programa al qual s'ha enfrontat diverses vegades el campió humà Gari Kasparov, la realitat és que els jugadors no professionals solen perdre davant de programes d'escacs fins i tot senzills. I aquest no és més que un primer exemple de les tècniques de la intel·ligència artificial en funcionament.

Per a arribar a fer-se una idea de com podria ser un món amb programes i màquines intel·ligents, és recomanable la lectura de la bona narrativa de ciència ficció entorn dels robots.

La millor opció, evidentment, són els relats i les novel·les de robots del científic, divulgador i novel·lista nordamericà Isaac Asimov, creador de les *Tres lleis de la robòtica*, veritable codi ètic per als robots intel·ligents del futur. Una proposta de lectura divertida i suggeridora, que recomanem especialment, seria:

Relats:

Asimov, Isaac (1950). *Yo robot*. Barcelona: Plaza y Janés.

Asimov, Isaac (1982). *Los robots*. Barcelona: Martínez Roca.

Novel·les:

Asimov, Isaac (1953). *Bóvedas de acero*. Barcelona: Plaza y Janés.

Asimov, Isaac (1957). *El sol desnudo*. Barcelona: Plaza y Janés.

Una altra opció, menys optimista, és la lectura d'algunes crítiques molt serioses que han sorgit en contra del projecte de l'anomenada "intel·ligència artificial dura". En resum, aquestes crítiques plantegen la impossibilitat del projecte de la intel·ligència artificial.

En aquest sentit resulta recomanable la lectura, no tan senzilla però molt interessant i estimulant, del llibre:

Penrose, Roger (1989). *La nueva mente del emperador*. Barcelona: Mondadori, 1991.

## Exercicis d'autoavaluació

1. Quin és l'objectiu principal de la intel·ligència artificial?
2. Ja s'han construït programes que superen el test de Turing. Cert o fals?
3. Els programes d'intel·ligència artificial són capaços d'aprendre. Cert o fals?
4. Esmenteu un exemple de programes d'intel·ligència artificial a l'abast del gran públic.
5. La inducció és la "inferència lògicament correcta". Cert o fals?
6. Què és l'heurística?
7. Què és la base de coneixements d'un sistema expert?
8. Els enginyers del coneixement són els experts l'experiència dels quals emulen els sistemes experts. Cert o fals?
9. Esmenteu dos llenguatges específics de la intel·ligència artificial.
10. Una xarxa neuronal artificial sol tenir milers de neurones. Cert o fals?

## Solucionari

### Exercicis d'autoavaluació

1. L'objectiu principal de la intel·ligència artificial és obtenir sistemes informàtics amb un comportament que considerariem intel·ligent en un ésser humà.
2. Fals. Malgrat les expectatives del mateix Turing, que creia que més o menys cap a l'any 2000 se superaria el test que porta el seu nom.
3. Cert. L'aprenentatge és un dels objectius importants de la intel·ligència artificial.
4. Alguns programes informàtics que juguen als escacs utilitzen tècniques d'intel·ligència artificial.
5. Fals. La "inferència lògicament correcta" és la deducció (*modus ponens* o instanciació universal).
6. L'heurística és el conjunt de criteris, mètodes o principis que s'utilitzen per a trobar, entre diversos camins possibles, quin o quins són els més efectius per a obtenir un objectiu determinat. Ve a significar les "regles de l'experiència" per a seleccionar un camí d'acció.
7. La base de coneixements d'un sistema expert és el conjunt de regles que sintetitzen l'experiència de l'expert humà que es vol emular.
8. Fals. Són els especialistes en intel·ligència artificial que ajuden a transferir l'experiència i el coneixement d'un expert humà a un sistema expert.
9. Lisp i Prolog.
10. Fals. Amb molt poques neurones s'obtenen bons resultats. La xarxa de Lang/Witbrock, per exemple, utilitza només tretze neurones.

## Bibliografia

### Bibliografia bàsica

**Cloksin, W.F.; Mellish, C.S.** (1981). *Programming in PROLOG*. Nova York : Springer-Verlag.

**Cortés, U. i altres** (1995). *Inteligència artificial*. Barcelona: Edicions UPC (Politext, 17).

**Cortés, U.; Sierra, C.** (1987). *LISP*. Barcelona: Marcombo-Boixareu Editores.

**Jackson, P.** (1986) *Introduction to Expert Systems*. Reading, Massachusetts: Addison-Wesley.

**McCorduck, P.** (1979). *Máquinas que piensan: una incursión personal en la historia y perspectivas de la inteligencia artificial*. Madrid: Tecnos, 1991.

**Wasserman, P.** (1989). *Neural Computing, Theory and Practice*. Nova York: Van Nostrand Reinhold.

### Bibliografia complementària

**Charniak, E.; McDermott, D.** (1992). *Introduction to Artificial Intelligence* (2a ed.). Reading, Massachusetts: Addison-Wesley.

**Feigenbaum, E.; McCorduck, P.** (1983). *The Fifth Generation*. Reading, Massachusetts: Addison-Wesley.

**Goldberg, A.; Robson, D.; Ingalls, D.H.** (1984). *SMALLTALK 80: The Language and Its Implementation*. Reading, Massachusetts: Addison-Wesley.

**Hancock, P.A.; Chignell M.H.** (eds) (1989). *Intelligent Interfaces: Theory, Research, and Design*. Amsterdam: North-Holland.

**Hofstadter, D.R.; Dennett, D.C.** (eds) (1985). *The Mind's I*. Harmondsworth, Middx: Basic Books Inc., Penguin Books Ltd.

**Kowalski, R.A.** (1979). *Logic for Problem Solving*. Amsterdam: North-Holland.

**Newell, A.; Simon, H.** (1972). *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall.

**Rich, E.** (1992). *Artificial Intelligence*. Singapore: McGraw-Hill, International Student Editon.

**Siklóssy, L.** (1976). *Let's Talk LISP*. Englewood Cliffs, N.J.: Prentice-Hall.

**Wertz, H.** (1986). *LISP: Introducción a la programación*. Barcelona: Masson.

