

Intel·ligència Artificial

Ton Sales

La història comença amb Norbert Wiener, matemàtic engrescat a imaginar-se una nova ciència que reuniria matemàtics, metges i biòlegs en l'estudi de la intel·ligència. Wiener la va batejar, l'any 1947, amb el neologisme «Cibernètica». La nova ciència de Wiener havia de permetre comprendre com s'originen i s'executen les accions intel·ligents i així, un cop comprès això, podríem fer les feines intel·lectuals rutinàries als robots, tindríem fàbriques automatitzades i l'home s'hauria alliberat, per fi, del treball no creatiu. Un dels primers llibres publicats en català sota el franquisme va ser precisament un opuscle de Wiener («Cibernètica i societat», d'Edicions 62) on ho explicava en un to visionari molt seu.

La Cibernètica era un gran projecte, però va durar poc. Un lògic, Kleene, estudiant precisament les capacitats intel·lectuals del «cervell» que proposava la Cibernètica, va descobrir que, si el cervell humà fos efectivament com aquell, el llenguatge fóra absolutament impossible d'explicar. En d'altres paraules: el nostre cervell no pot ser de cap manera tan simple com el proposat, perquè si ho fos no s'entendria que nosaltres fóssim capaços de fer el que fem. Aquesta conclusió, feta pública l'any 1955, va suposar la ferida de mort per al programa cibernètic. I va donar lloc a l'episodi següent: la «Intel·ligència Artificial» (IA).

L'estiu del 1956, un grup d'universitaris americans, majoritàriament matemàtics, es va reunir a Dartmouth, New Hampshire, per comentar la situació i mirar de trobar una sortida endavant en l'estudi de la intel·ligència humana després de l'ensulsiada de la Cibernètica. Hi havia en John McCarthy (el mateix que ha estat a Barcelona fa poques setmanes),

Marvin Minsky, Herb Simon (futur Nobel) i altres. El seu plantejament era clar: tots ells tenien ordinadors a les seves universitats i alumnes entusiastes capaços de programar-los; de manera que van proposar-se fer equips i elaborar programes que fessin coses normalment considerades «intel·ligents» —tals com jugar a escacs, raonar sobre figures geomètriques o seguir l'esquema de raonament d'algú quan fa una demostració matemàtica—. No volien fer màquines ni teories; només programes. La idea era simple, i el projecte proposat (que McCarthy va batejar, sense donar a entendre res en especial, «intel·ligència artificial») incloïa el retrobament periòdic del grup, cada dos o tres anys, per avaluar els programes que cadascú havia construït.

L'èxit d'aquesta empresa ha estat fulgurant, més encara quan es té en compte que només era un programa temptatiu per desencallar una situació. A les reunions següents, cada equip aportava resultats sorprenents: programes que eren capaços de fer coses «intel·ligents», tals com, efectivament, les proposades (jugar als escacs, etc.); però, a més, els programes podien ser discutits i avaluats, se'n podien estudiar les limitacions i —sobretot— es podien posar en comú certs mètodes que havien funcionat bé. Amb el temps, fins i tot s'arribava a situacions tan sorprenents com l'ocasionada per Eliza, un programa que merament es proposava simular les reaccions d'un psicoanalista. L'efecte que feia en els que dialogaven amb el programa era fulminant. Tothom «veia» en el seu interlocutor un ser intel·ligent i fins i tot sensible. L'experiència il·lustrava clarament la facilitat que tenim de distòrcer la nostra percepció de la intel·ligència dels altres, i sobre la vacuïtat que poden tenir alguns intents

de «crear» intel·ligència a partir de màquines.

Cap a la fi dels seixanta, el cabal de tècniques de què es disposava era impressionant. Els diferents programes, una vegada passat el fort impacte inicial que tots ells provocaven en els observadors, potser no n'eren gaire, d'«intel·ligents», però eren una abundància, i relativament barata, matèria d'experimentació. D'aquesta època són els primers manuals sobre IA que sistematitzen per primera vegada un corpus científic com més va més sòlid.

Un dels grans èxits de la IA dels seixanta havia estat la creació de programes «raonadors»; la qual havia il·lustrat, de passada, com era d'enganyosa la nostra percepció intuïtiva de la intel·ligència humana, segons la qual la capacitat de raonar abstractament era vista com una facultat «superior» (i vet aquí que ara resultava més fàcil de materialitzar que d'altres facultats considerades generalment com a no tan humanes, com ara les de reconèixer cares, o olors). Doncs bé, per millorar la capacitat deductiva dels sistemes raonadors es va proposar concentrar els coneixements (a partir dels quals s'havien de deduir les conseqüències lògiques) a temes molt concrets, tancats. El resultat és que s'obtenien sistemes que raonaven si fa no fa com ho faria un expert humà en el tema. El sistema assimilava els coneixements rellevants, expressats en forma de «regles» (del tipus «si passa això aleshores fes allò») i en deduïa conclusions, o bé determinava accions a prendre. Els primers «sistemes experts», com es van dir aquests programes, tractaven de temes restringits com ara el diagnòstic de certes malalties a partir de símptomes (el Mycin) o la predicció de dipòsits de minerals a partir de dades geo-

lògiques (el *Prospector*). Doncs, aquest últim va ser catapultat a la fama periodística, i amb ell els sistemes experts en general (i de retruc tota la IA), perquè a primers dels anys setanta va predir —i ho va encertar— un jaciment de molibdè valorat en molts milions de dòlars.

Després d'aquest fet, la IA ja no va ser mai més aquella empresa científica progressiva i modesta que havia estat. Una vegada esdevinguda titular de diari, va sofrir una mutació. Empreses de totes les mides sorgien com bolets per oferir al mercat «sistemes experts» de tota mena. Amb tot plegat, els diaris van començar a anar plens d'«Intel·ligència artificial», com si aquesta cosa de títol altisonant i fatxada fos una invenció recent. I per si fos poc, els japonesos ho van acabar de complicar anunciant, a primers dels anys 80, una «cinquena generació» d'ordinadors, que serien essencialment sistemes experts santificats i generalitzats a tots els camps imaginables.

LA CRISI DELS TRENTA ANYS

Si els anys 1956-76 són de creixença i la dècada 1976-86 l'esplet de l'èxit, cap allà l'any 86 es respira una sensació d'atzucac. Què ha passat? Essencialment dues coses:

1) L'èxit dels sistemes experts sembla indicar que la idea de codificar els coneixements a base de regles i de deduir-ne conclusions (raonar) és un idea no sols bona sinó d'aplicació general. Ara bé, inevitablement sorgeixen alguns dubtes:

- Tots els camps del coneixement s'hi presten, a ser codificats per regles? Resposta [optimista]: sí, si ho fem amb molt de compte. (D'aquest «fer-ho amb compte» se'n diu «Enginyeria del coneixement»).

- En certs camps del coneixement (per exemple: el comportament lingüístic del parlat d'un idioma) ser «expert» obliga a retenir de l'ordre de vint, potser quaranta, mil regles, la qual cosa és

desastrosa pel que fa al rendiment d'un sistema que les hagi de tractar. Resposta [japonesa]: farem màquines ultraràpides que processin un nombre elevadíssim de regles per segon. Resposta complementària: organitzarem les regles de manera que s'hi pugui «navegar» per accedir ràpidament a qualsevol fragment del coneixement. (De les tècniques usades en aquesta organització se'n diu «Representació del coneixement»). Ara bé: si ja ens vèiem ofegats per una allau de regles, ara es tracta d'organitzar-les i controlar-les (afegint un nou problema al que ja teníem). Com prendre-s'ho? Qui treballa professionalment en la IA, enfrontat cada dia a la complexitat creixent dels programes, és fàcil que tingui la sensació d'haver arribat al límit de les possibilitats. I d'altra banda, les respostes «optimistes» semblen una crida a la fe, l'equivalent tècnic del copet a l'esquena. No és estrany, doncs, que més d'un hagi pensat que el que realment calia era fer un gir copernicà complet.

2) La segona cosa que ha passat és que es palpa una sensació de frustració no ja pel que fa a la construcció o complexitat dels programes sinó pels resultats que exhibeixen. Si els nostres programes són intel·ligents, això ho hauria de notar tothom. I, en canvi, la sensació dominant és que els nostres programes no en són, d'intel·ligents. La intel·ligència que mostren és, pròpiament i a tot estirar, només la del programador, i la inherent als mètodes usats en la seva construcció. I, a més, arreu apa-reix el molest fenomen de la «fragilitat» ('brittleness'): el programa «fa bondat» i sembla intel·ligent, però fàcilment arriba el cas que, perquè inadvertidament freguem o depassem els límits previstos d'aplicació o pel que sigui, el programa es comporta com un estúpid indissimulable o fa alguna bestiesa incomprendible (que demostra l'absència d'un mínim de sentit comú). (Per exemple, un sistema diagnosticador pot entestar-se a recomanar continuar mesurant la pressió del malalt tot i que aquest acabi de morir-se, o tot i que hi hagi un incendi a l'hospital, o pot quedar-se completament tallat, sense oferir cap

mena d'alternativa, quan passa algun imprevist, encara que sigui perfectament raonable, i l'alternativa senzilla i òbvia). En resum, és molt fàcil que el sistema «ensenyi el llautó». En aquests casos, la sensació de desconcert i decepció de l'observador humà (que imperceptiblement i inevitablement havia assimilat el comportament intel·ligent de la màquina o programa a l'esperable en humans en condicions semblants), acostuma a ser total, espectacular i difícilment recuperable. La resposta optimista a aquest problema («també els humans ensenyem el llautó quan estem fora del nostre ambient») ja es comprèn que no sigui gaire satisfactòria. I encara més: com va dir una vegada Roger Schank (un dels grans pontífexs de la IA) fa uns anys a Barcelona (en una conferència a la Generalitat), resulta penosíssim constatar la impressió d'estupidesa irrecuperable que fa veure un programa insistir milions de vegades a repetir un mateix resum de premsa (sense «avorrir-s'hi» ni «queixar-se'n»!) o la seva ineptitud a detectar la banalitat d'una conclusió que clarament no porta enlloc, o la incapacitat de «veure» l'interès d'algun fet nou... No, no hi ha res a dir: el programa només fa el que li han dit que fes; però difícilment es pot qualificar d'intel·ligent un comportament que més aviat recorda el d'un insecte repetint enfollit una mateixa acció demostradament estèril o pocasolta.

Davant de tot això, i de la desorientació i sensació d'impassa que pro-voca en els recercadors de la IA, la pregunta és: com sortir-ne?

Hi ha una sortida evident: la que consisteix a no veure que hi hagi cap crisi i tirar pel dret amb les solucions tècniques proposades: augment —i domini— de la complexitat creixent dels programes, augment de la capacitat i del rendiment de les màquines, i aprofundiment dels nous camps proposats («Enginyeria del coneixement» i «Representació del coneixement»). Aquesta és la sortida adoptada si fa no fa per tothom, ja que permet anar tirant. Però n'hi ha, almenys, dues més, que parteixen de

la hipòtesi de crisi. Són, d'una banda, el «Connexionisme»; d'altra, la d'un grup heterogeni d'investigadors europeus i americans insatsfets amb la situació actual.

A) EL «CONNEXIONISME»
(o «Xarxes neuronals», o «Neuro-informàtica», o «IA subsimbòlica»)

Parteix d'una idea bàsica: Reconèixer una cara li costa al nostre cervell una dècima de segon (= 100 mil·lisegons). Doncs bé: si les nostres màquines tan ultraràpides no són capaces —amb les nostres tècniques de programació ultrasofisticades— de fer una cosa tan senzilla i universal en tan poc temps i, en canvi, el nostre cervell ho fa sense problemes (i el nostre cervell, cal remarcar-ho, és una màquina molt lenta i que tendeix a cometre errors), això vol dir que hem de canviar de tàctica, almenys quan vulguem fer coses com ara reconèixer cares (o veus, o situacions). Com fer-ho? El cervell ens en suggereix una manera: creem un conjunt d'unitats (diguem-ne «neurones») i connectem-les massivament: tot seguit, fem funcionar el sistema resultant i experimentem amb els resultats. Doncs bé: els resultats són molt sorprenents. Hem pogut així construir sistemes («xarxes neuronals», se'n diuen, amb un abús de llenguatge) que sense ser gaire complicats són capaços de recordar moltes diferents imatges (i qui diu «imatges» podria dir «sensacions» o «situacions»), i fins i tot seqüències senceres d'imatges successives. Aquests experiments com a mínim aconseguen impressionar tothom que els veu. Ningú no sap si tot això porta en direcció a la intel·ligència, o si només és una sèrie de fenòmens que presenten una vaga analogia amb els fenòmens cerebrals. Sigui com sigui, algunes característiques dels sistemes connexionistes són noves:

- Memòria «distribuïda»: la memòria aquí ja no és un «receptacle» al qual s'accedeix sinó un conjunt d'«estats» que el sistema «recorda» i que és propi de tot el sistema, globalment, i no localitzable.

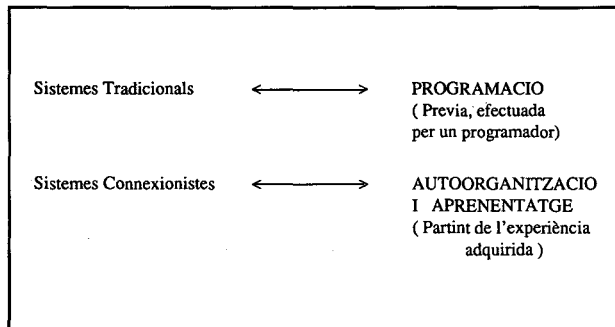
- Memòria «associativa»: l'accés a un record concret no es fa per adreçament (=dir on trobar-lo) sinó per contingut (=donar-ne un esquema bàsic, «evocar-lo»).

- Facilitat d'abstracció: el reconeixement actua com un procés d'autoorganització de la informació que accentua i aïlla trets bàsics d'una imatge.

- Paral·lelisme implícit en el procés: totes les unitats («neurones») col·laboren en qualsevol tasca.

- «Degradació gràcil»: el rendiment no cau pas en picat pel fet que una part —tot i important— del sistema deixi de funcionar. (A comparar —favorablement— amb la total intolerància a l'error que tenen els sistemes tradicionals i que provoquen la interrupció immediata i el bloqueig total del procés en curs.)

- Absència de «programa»: el sistema no funciona pas perquè se'l programi prèviament (de fet aquí no hi ha «programador») sinó que observa l'input, s'«autoorganitza», i actua en funció del procés viscut (que resta sempre obert a noves experiències). Aquestes propietats, absents dels sistemes tradicionals de la IA, configuren una metodologia (també una nova IA? o una anti-IA?) que potser és resumible en l'última característica citada. Es a dir :



El connexionisme no se sap si avança cap a alguna cosa, però almenys representa una alternativa a la IA tradicional. Aquells que consideren que el connexionisme no és una amenaça sinó un complement de la IA s'esforcen a fer compatibles els dos

enfocaments en un esquema únic que podríem descriure succintament com es veu a la fig. 2.

El problema, aquí, és com comunicar els dos camps a través de la barrera. La IA tradicional treballa amb PROGRAMES (que tenen accés a una memòria que és la dipositària del coneixement del sistema) i actua sobre objectes que són essencialment SIMBOLS, mentre que el connexionisme treballa amb XARXES (el diferents estats de les quals determinen el coneixement del sistema) i actua sobre ESTATS globals d'energia. Com s'ha de fer doncs per anar cap amunt (=convertir estats en objectes simbòlics)? i cap avall (=l'invers)? De moment ningú no ho sap veure. De fet, les habilitats de cada sistema semblen completament diferents i incompatibles.

B) LA RESTA:

No tothom és entusiasta del connexionisme, ni tampoc es pot dir que la gent de la IA sigui feliç d'haver de triar entre dues alternatives com si hagués de triar entre dues religions. Un element implícit i positiu del connexionisme és que trenca l'estat hipnòtic de pensar que la IA ha de ser vista necessàriament des d'una sola perspectiva (que és, encara, l'original de 1956: IA = fer programes). El connexionisme permet imaginar coses que no són ni tan sols «programables» en el sentit habitual. Només per això ja mereixeria ser considerat com un bon remei contra l'encarcerament. Però és que hi ha investigadors —no connexionistes— que creuen que

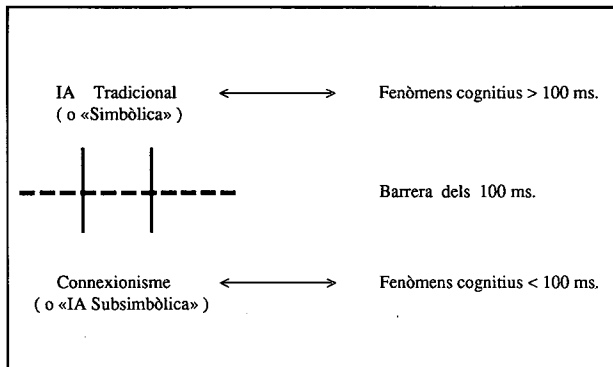
la «crisi» de la IA és una ocasió única per tirar endavant en la comprensió de la intel·ligència. En citarem tres: Steels, Hillis i Toffoli.

Luc STEELS, de la Universitat Lliure de Brussel·les, ofereix nous esquemes que permeten

treballar sota qualsevol metodologia: tradicional, connexionista o cap de les dues. Afirmar que el càlcul no és necessàriament simbòlic ni el contrari, i que sovint és l'efecte i resultat d'un procés cooperatiu. Exemple: una fila de formigues «calcula» automàtica-

proposa no capficar-s'hi gaire. Ah! i no es tracta pas d'un afer de formiguers, només: allà on hem dit formigues i pedres podríem haver dit robots i obstacles, o subjectes i situacions.

«genètics», dits així perquè —deixats anar— ells sols s'autoorganitzen o presenten espontàniament fenòmens tals com mutacions, seleccions, variació, creixement, supervivència o extinció que recorden extraordinàriament els comportaments evolutius clàssics de la Biologia).



Daniel HILLIS és el creador de la Connection Machine, un ordinador paral·lel ultraràpid, d'estructura semblant a un reticle. Hillis observa que la

L'americà **Tommaso TOFFOLI** i el seu grup de l' MIT treballa sobre un model de càlcul bastant diferent de l'habitual (si bé inventat per Von Neumann a la fi dels anys 40), els «autòmats cel·lulars», que permeten, com la Connection Machine, experimentar comportaments i dinàmiques de càlcul —incloent-hi els «genètics»— molt diferents dels corrents, amb resultats que no paren de sorprendre.

ment la trajectòria mínima entre el menjar i el formiguer, que sol ser —si res no ho impedeix— una recta; ara bé, si col·loquem una pedra que interfereixi la trajectòria calculada, les formigues la recalcularan espontàniament de manera que la nova solució eviti l'obstacle i continuï essent mínima. Aquí és ocios parlar de «quí» calcula la trajectòria: la calcula algun «programador»? es calcula «sola»? és realment un «càlcul»? o més aviat es tracta d'«autoorganització»? Steels

dinàmica de càlcul de la seva màquina presenta fenòmens curiosos: de vegades, en segons quins càlculs, l'activitat calculatòria es concentra en una part del reticle, en d'altres es distribueix a l'atzar o en oscil·lacions periòdiques, o caòtiques. Tot això suggereix implícitament que hi ha molt que no sabem de què és un «càlcul» i de les sorprenents maneres en què un càlcul es pot executar en màquines d'arquitectures no-convencionals. (I encara hi podríem afegir els càlculs

Cap on porta, tot això? No ho sabem pas. Només es pot afirmar que la IA encara té molt de camp per córrer, i l'última paraula encara no l'ha dita ningú. En tot cas, prepareu-vos per sorpreses. Com diu McCarthy, si la Física no ha resolt certs problemes fins al cap de cent anys d'estudiar-los, la IA —que Déu n'hi dó les sorpreses que ens ha donat en 34 anys— per què havia de resoldre els seus en menys?

Ordenadores personales e impresoras

olivetti