

Els somnis de Babbage

La tecnologia com a interlocutor

Jaume Ferrer i Rosera
juny 2004

Copyright/Copyleft © 2004 Jaume Ferrer i Rosera
S'atorga permís per a copiar, distribuir i/o modificar aquest document sota els termes de la Llicència de Documentació Lliure GNU, versió 1.2 o qualsevol versió posterior publicada per la Free Software Foundation. Es considerarà com a Seccions Invariants tot el document, no existint Textos de Portada ni Textos de Contraportada. Podeu consultar una còpia de la llicència a <http://www.gnu.org/copyleft/fdl.html>

Resum

Proposem una aproximació al fenomen de la imprevisibilitat en l'experiència d'usuari de tecnologia digital entesa com el resultat de l'augment de la complexitat de l'entorn social i tecnològic. Revisarem alguns conceptes clau com el de realitat, el d'interacció a temps real i el de consciència des del punt de vista de les interfícies i dels protocols de comunicació. Oferirem una breu introducció al software autònom i a l'estat de la recerca en intel·ligència artificial en matèria de sistemes multiagent i especularem sobre l'origen i l'impacte d'aquests sistemes socials en el context d'un món globalitzat.

Internet com a procés

Si les circumstàncies de la primera meitat del segle XIX haguessin permès que la proposta de màquina analítica de Charles Babbage hagués reeixit possiblement ara no percebríem de la mateixa manera el subtil lliandar que separa el món analògic del digital (els àtoms dels bits, en paraules de Negroponte).

Com hauria evolucionat la informàtica si s'hagués desenvolupat en paral·lel a la revolució industrial?. Potser podríem haver gaudit de prop d'un segle de hardware de la mida de camps de futbol dissenyat d'acord amb la visió del món ordenada i previsible com un rellotge suís que es derivava de les equacions de Newton i posada al servei de formes d'organització social significativament diferents de les actuals.

Potser fruit d'aquesta inèrcia mai no hagués estat possible desenvolupar un sistema de comunicació distribuït basat en paquets i en comptes d'internet tindríem alguna cosa més semblant al Minitel.

Imaginem, però, que tot hagués acabat sent semblant a com és ara. Evolucionant lentament des dels eixos i els engranatges impulsats per calderes de vapor fins a l'actual microelectrònica

portable i en xarxa, les tecnologies de la informació ens resultarien tan familiars com les de l'automoció i possiblement no usaríem certes etiquetes com ara 'realitat virtual', de la mateixa manera que no parlem "d'autolocomorealitat" ni "d'electrorealitat".

Pel contrari, la informàtica i les xarxes de comunicació digitals no tenen gaire més de seixanta anys d'història, la major part de la qual s'ha desenvolupat d'esquenes a la ciutadania. De fet, podríem afirmar que als països on les TIC gaudeixen de certa popularitat, fa trenta anys la gent pràcticament no les coneixia i només se n'ha fet un ús generalitzat a les darreres dues dècades.

Les TIC s'han desenvolupat en el context d'un paradigma científic i cultural segons el qual la visió del món, de l'espai, del temps, de la societat i de la cultura difereixen de la del segle XIX. Tot i que alguns d'aquests canvis han arrelat poc entre la gent, sí que han tingut efectes contundents a través de recerca aplicada i de la gestió, que ha transformat completament la forma d'analitzar el món i d'interaccionar-hi (Held et al. 2001).

En realitat, en certa manera Babbage sí va tenir èxit. Era un fervent defensor del lliure intercanvi d'informació entre nacions, sense limitacions, convençut de que això contribuiria a la prosperitat de tothom (Mattelart, 2001). Cent anys després, altres que pensaven com ell, i d'altres que tenien molt a guanyar si aquestes idees triomfaven, impulsarien els canvis en la legislació, l'economia, i la tecnologia que durien l'estat del món a la situació que coneixem com a globalització. El món digital, com qualsevol fenomen social d'un cert abast, beu d'arrels molt fondes que s'entrecreuen i es perden a les pregoneses del temps.

Internet es podria considerar més com una emergència sorgida al si de determinats processos socials que no pas com el resultat d'un determinat pla estratègic o de l'evolució lineal de la tecnologia (Abbate, 1999). D'altre banda, el seu potencial transformador és tan gran que també es converteix en un agent de canvi dels propis processos que l'han generada. Per tant, podríem parlar d'internet com d'un procés més que no pas com d'una infraestructura tecnològica concreta. A més de considerar-la com un producte o artefacte cultural (Hine, 2000), la naturalesa orgànica d'internet i del conjunt de xarxes de comunicació digitals adquireix tanta complexitat i rellevància que en podríem dir "l'artefacte" en singular. Malgrat que una part important de la població resta exclosa a l'altra banda de l'esclletxa digital (Castells, 2002), podríem afirmar que tothom acaba sent afectat directa o indirectament per allò que passa a la xarxa, per bé que jugar-hi un paper clau seria una qüestió d'oportunitat i de recursos. Tot i la innovació tecnològica que suposa internet, la seva principal importància rau en l'ús que se'n fa, per bé que aquest ús no sigui encara ni de bon tros majoritari.

És possible que la relativament recent irrupció de les TIC i el fet que encara ara se'n parli bastant més que no pas s'utilitzin hagi fet aparèixer una sèrie de mites i de percepcions distorsionades (Hine, 2000). A l'hora d'atansar-nos al món digital ens trobem com l'Àlicia mirant a través del mirall. Intuïm que la realitat que s'endevina allà al davant hauria de ser una perllongació de la que estem apunt d'abandonar però l'experimentem d'una manera diferent, entre d'altres coses perquè la imaginem diferent. Per començar a parlar sobre l'encaix entre experiència analògica i digital hauríem de definir el concepte mateix de realitat.

Realitat virtual

Reid utilitza l'expressió 'realitat virtual' per referir-se a l'experiència que suposa la comunicació mediada per ordinador (CMO) en grup, encara que només es tracti de comunicació textual en un entorn de tipus MUD ¹.

Si les vivències dins d'un MUD es poden considerar reals en tant que són viscudes amb plena intensitat i en tant que tenen repercussions, aleshores no té massa sentit etiquetar les relacions humanes al món analògic com a 'relacions socials reals'. Superant el tòpic del visor 3D i dels guants de dades, la noció de realitat seria independent de la tecnologia o mitjà utilitzat per mediar en la comunicació i rauria en la pròpia experiència (Reid, 1999). Podríem afirmar que si ho experimentem com a real és real. En tot cas potser podríem distingir entre experiències basades en estímuls generats de forma digital o analògica, entre experiències presencials i no presencials, online i offline, etc.

Com diu Pierre Lévy, “no es tracta de triar entre la nostàlgia d'un real datat i un virtual amenaçador o excitant, sinó entre diferents concepcions d'allò que és virtual” (Lévy, 1995). En aquest sentit, la realitat que construïm a partir de la informació que ens transmeten els mitjans de comunicació de masses es podria considerar tant virtual com qualsevol altre experiència.

Des del camp de recerca en comunicació interactiva, l'expressió 'realitat virtual' ha estat associada a d'altres significats més concrets i restrictius. Narcís i Roc Parés proposen una definició integradora que també té a veure amb l'experiència de l'usuari però que se centra en la seva naturalesa tecnològica i s'allibera d'alguns llastres pesants, com per exemple el concepte de simulació. Per 'realitat virtual' s'entendria una “interacció amb estímuls digitals generats a temps real”, és a dir, una interacció persona-ordinador on el sistema generaria respostes als estímuls canviants de l'usuari sobre demanda, sense haver estat prèviament processats i amb un retard tan curt que seria imperceptible (Parés i Parés, 2002). En certa manera podríem dir que la noció de realitat es podria associar a la d'immediatesa, no necessàriament en el sentit de comunicació síncrona de la CMO, sinó en el de resposta ad hoc, no definida i emmagatzemada prèviament.

En aquesta definició es contempla la comunicació digital com una qüestió d'interacció entre una persona i un sistema digital. En un sentit estricte tota CMO és una comunicació amb un sistema digital, atès que tots els estímuls i les respostes, ja siguin generats per un altre usuari humà o bé pel propi sistema digital, estan condicionats per les restriccions que imposa la interfície a través de la qual la persona i l'ordinador poden interaccionar i que implica la traducció analògic-digital-analògic.

La membrana informacional

El sistema sensorial humà també es pot entendre com una interfície que filtra els estímuls. En aquest sentit, la interacció social també es podria considerar com comunicació mediada pels nostres sentits i per les nostres estructures cognitives.

De fet, ens podríem plantejar si realment els humans interactuem a temps real amb el món que ens envolta. Medicions elèctriques experimentals del sistema nerviós semblen indicar que existeixen petits retards tant en el procés de presa de consciència d'un determinat estímulo com en el de presa de consciència d'una determinada resposta. Malgrat que els retards són prou llargs com per ser perceptibles, els ignorem sistemàticament. (Penrose, 1989).

Per al físic Roger Penrose, això tindria implicacions sobre la pròpia noció de realitat en tant que realitat construïda a partir d'informació 'procedent de l'exterior' de la pròpia consciència:

Existe, efectivamente, algo muy singular en el modo en que entra el tiempo en nuestras percepciones conscientes en cualquier caso, y pienso que es posible que sea necesaria una concepción muy diferente cuando tratamos de colocar las percepciones conscientes en un marco convencional de ordenación temporal. La consciencia es, después de todo, el único fenómeno que conocemos según el cual el tiempo ¡necesita fluir!. (...) El ordenamiento temporal que nos parece percibir es, estoy afirmando, algo que imponemos a nuestras percepciones para poder darles sentido en relación con la progresión temporal uniforme hacia delante de una realidad física externa.

Roger Penrose. La nueva mente del emperador. 1989

Alguns neurobiòlegs consideren la consciència com un procés dinàmic on cada cop que es produeix un estat conscient determinat es descarten molts altres estats possibles. Implicaria el processat de grans quantitats d'informació en molt poc temps mitjançant una xarxa de circuits sotmesos a adaptació i reconfiguració constant (Edelman i Tononi, 2000). L'arquitectura d'aquest sistema estaria encara molt lluny de la dels dispositius informàtics actuals, però en qualsevol cas requeriria l'existència d'una interfície per comunicar-se amb l'exterior.

Algunes interfícies cervell-màquina de caire experimental posen encara més de relleu el fet que allò que podríem anomenar “l'experiència del món real” implica l'existència d'un 'dins' i d'un 'fora' separats per una interfície de comunicació i l'establiment d'uns protocols de comunicació que assegurin una interacció consistent. En aquests sentit són especialment significatius els experiments realitzats amb elèctrodes implantats al cervell per controlar respostes, com per exemple el de la Universitat d'Emory, on l'activitat elèctrica del cervell aconseguia moure de forma controlada un cursor per una pantalla d'ordinador; o bé els experiments destinats a donar sentit a estímuls digitals, com el de l'empresa Dobelle, que permetia percebre llums i ombres a través de l'ull artificial d'una càmera que enviava impulsos elèctrics modulats per un ordinador al cervell d'un invident².

La imatge de la interfície i dels protocols com a membrana informacional que separa el dins del fora ens recorda les membranes biològiques, en constant interacció bioquímica, dels organismes unicel·lulars. De fet, els bacteris són un bon exemple de transacció informacional a gran escala. A través de les seves membranes intercanvien fragments de codi genètic que esdevenen plenament funcionals i els permeten implementar funcions metabòliques no contemplades en el seu codi original (Margulis i Sagan, 1986). En termes de societat de la informació, està clar que els bacteris ens porten milions d'anys d'avantatge.

Com afirma Haken (1999), parlar de societat de la informació per distingir-nos d'altres formacions socials pot ser molt etnocèntric, atès que en qualsevol societat s'intercanvia informació i que totes les formes de vida depenen de la informació.

Margullis defensa que aquests intercanvis no es produeixen de forma arbitrària, sinó que a la seva manera els bacteris trien entre diverses estratègies possibles, per tant prenen decisions i tenen en compte el context. No viuen en el buit sinó al si d'ecosistemes amb intenses interaccions entre individus que assoleixen sofisticades formes d'autoregulació. Aquests i d'altres sistemes d'intercanvi d'informació on parts simples i aparentment mancades de visió global i d'estratègia de grup es veuen involucrades en dinàmiques altament complexes i sovint ordenades ens proporcionen la imatge poderosa de la emergència, de l'ordre espontani sorgint del caos (Lewin, 1992).

La percepció del software com a entitat autònoma

Vista com una emergència, internet seria quelcom més que “la suma d'ordinadors que poden comunicar-se usant el seu llenguatge, el protocol TCP/IP” (Hine, 2000), seria una mena

d'ecosistema format per les diferents capes de software i de hardware que permeten la interacció social i alhora en són el seu resultat.

Succeeix, però, que el tràfic d'informació a internet està en la seva major part en mans de sistemes de software que negocien entre ells la ruta dels diferents paquets de dades de forma automàtica. D'altra banda, l'astronòmica quantitat d'informació disponible a la xarxa obliga als usuaris a utilitzar eines de cerca, les quals funcionen gràcies a agents de software que exploren constantment la xarxa i la cataloguen (knowbots). D'altres agents també exploren la xarxa, però amb la finalitat d'aprofitar-ne les vulnerabilitats, ensumar contrasenyes, obtindre informació privada o simplement autoreplicar-se i col·lapsar sistemes. Afortunadament, també tenim agents que vetllen per mantenir els nostres sistemes nets d'agents maliciosos, que es descarreguen periòdicament les signatures actualitzades de virus coneguts per poder detectar-los i que notifiquen a les seves bases respectives codis sospitosos.

Dins d'internet podríem afirmar que som algorismodepenents i, atesa la importància estratègica d'internet, podríem afirmar que també ho som fora d'internet. Sistemes bàsics directa o indirectament per a la nostra vida quotidiana com les finances internacionals, el transport o la sanitat operen gràcies a mecanismes de vigilància, anàlisi i gestió automatitzats que s'encarreguen de tasques rutinàries però vitals.

Tots aquests agents han estat desenvolupats i són gestionats per humans, però han estat dissenyats precisament per actuar amb un cert marge d'autonomia. Aquesta activitat acotada per humans en línies generals però no en els seus detalls implica un tràfic d'informació 'autònoma' que creix de forma exponencial. I com la ciència que s'ocupa de l'estudi del caos i dels sistemes complexos ens ensenya, quan els sistemes assoleixen un nivell de connectivitat crítica i es situen en el límit del caos poden aparèixer propietats emergents en forma de noves estructures ordenades (Lewin, 1992).

Sense entrar en l'especulació de si serà possible o no que del conjunt d'interaccions de la xarxa en sorgeixi una mena d'intel·ligència o consciència global³, és evident que per a la majoria d'usuaris/desenvolupadors de la tecnologia digital aquesta tendeix a manifestar de tant en tant comportaments aparentment arbitraris o incomprensibles que no sembla que responguin directament a cap finalitat funcional concreta.

Sherry Turkle ha observat com els nens tendeixen a considerar els personatges dels jocs d'ordinador com unes entitats híbrides a mig camí entre les màquines i els éssers vius, i especula sobre l'efecte que això pot tenir sobre la seva concepció d'allò que està viu (Turkle, 1995). L'experiència d'un usuari davant la seva màquina quan aquesta no es comporta com

s'esperava reforça encara més la sensació d'estar interactuant amb un sistema biològic aliè, no amb una màquina modelada per la cultura.

Diem que els nostres ordinadors “es pengen”, “s'infecten” o es “moren”. Aquestes expressions no són gratuïtes. Els estem convertint en éssers amb entitat pròpia. No és només per la relativa familiaritat que hi establim (no parlem així dels cotxes o de les cases) sinó perquè tendeixen a mostrar una imprevisibilitat moderada que els fa semblar vius. Aquest comportament lleugerament imprevisible pot ser intencional, com en el cas d'un joc d'ordinador per fer-lo més atractiu, o bé pot ser fruit de la pròpia inconsistència del disseny, com en el cas dels errors i de les inestabilitats aparentment inexplicables dels nostres sistemes⁴.

En certa manera, els espais socials creats al si de les xarxes es podrien entendre com una capa més, una capa que està per sobre de la capa d'aplicacions, i on trobem els pensaments, les expectatives i les reaccions humanes. En la mesura que la tecnologia s'anés tornant més imprevisible, els humans a més d'interaccionar entre nosaltres també estaríem interaccionant amb el conjunt de capes inferiors com si ho féssim amb qualsevol actor social.

Intel·ligència artificial i sistemes multiagents

Fins a quin punt podem considerar als agents de software com a interlocutors veritablement autònoms?. Podríem considerar que la interacció entre agents de software té alguna connotació cultural?, és a dir, podríem parlar de comunitats o de cultures estrictament cibernètiques?.

Per intentar respondre a aquestes preguntes caldrà fer un cop d'ull a l'estat actual de la recerca que s'ocupa de l'anomenada 'intel·ligència artificial' (IA). Aquest camp és molt extens i persegueix simultàniament interessos científics i d'enginyeria. D'una banda pretén comprendre com es produeix la percepció, el raonament, la comprensió, la comunicació o l'acció tant en màquines com en persones o animals. De l'altra, pretén desenvolupar a llarg termini sistemes que puguin fer totes aquestes coses, tot i que hi ha disparitat de criteris sobre el propi concepte “d'intel·ligència” aplicat a un sistema informàtic en el sentit de si serà *possible* o no (o té alguna utilitat) intentar simular o duplicar la intel·ligència humana (Penrose, 1989).

Existeixen diferents aproximacions a l'IA però la majoria d'exemples que veurem es basen en l'aproximació 'ascendent', que està molt lligada al concepte de 'vida artificial' entès com la dinàmica d'intercanvi de senyals que es produeix entre diferents parts simples i independents d'un sistema a partir de la qual sorgeixen propietats emergents. Un altre concepte bàsic és el d'agent.

Segons Björn Hermans, es configuren dues nocions d'agent, la feble i la forta⁵. Ambdues defineixen un agent com un hardware o, més habitualment, un software que reuneix les següents característiques:

Segons la noció feble:

autonomia: capacitat d'operar sense la intervenció directa d'humans o d'altres sistemes i capacitat de disposar d'alguna mena de control sobre les pròpies accions i el propi estat intern.

Sociabilitat: capacitat d'interacció amb altres agents i (possiblement) humans a través d'algun tipus de llenguatge de comunicació.

Reactivitat: capacitat de percepció de l'entorn i de trencar el seu estat de latència en resposta a certs canvis ambientals.

Proactivitat: capacitat de respondre desplegant comportaments dirigits a objectius que vagin més enllà de la simple reacció a l'ambient i prenent la iniciativa.

Continuïtat temporal: un agent està sempre executant un procés, ja sigui d'acció o de latència, i no es limita a executar un sol cop scripts que connecten un input concret amb un determinat output.

Orientació a objectius: capacitat de manipulació complexa i de determinar com subdividir una tasca d'alt nivell en subtasques més simples.

Segons la noció forta:

A més de les característiques anteriors, un agent hauria de tenir també com a mínim una de les característiques següents:

Mobilitat: capacitat de moure's per una xarxa digital.

Bondat: assumpció de que un agent no té com a objectiu el conflicte i que sempre intentarà fer allò que se li demana.

Racionalitat: assumpció de que un agent actuarà, sempre i quan això sigui possible, per tal d'aconseguir els seus objectius i no per evitar d'aconseguir-los.

Adaptabilitat: capacitat d'adaptar-se als hàbits, mètodes de treball i preferències del seu usuari.

Col·laboració: un agent hauria de tenir en compte que el usuaris humans cometen errors i no acceptar ordres que suposin objectius conflictius, que estiguin formulades amb ambigüetat, que

suposin un consum de recursos de la xarxa intolerable o que puguin ser perjudicials per a d'altres usuaris. Pel contrari, abans d'actuar un agent hauria de fer comprovacions preguntant a l'usuari o basant-se en un model d'usuari.

Com veiem, els agents són per definició sistemes orientats a la comunicació i als quals se'ls hi suposa certa autonomia. Podríem preguntar-nos si la noció de comunicació és la mateixa aplicada a un agent que a un humà.

Hakken (1999) afirma que els humans processem la informació d'una manera diferent que la resta d'organismes vius perquè en el nostre cas la informació resulta doblement 'culturalitzada', durant la percepció i durant el processat posterior:

1. Environmental data becomes semiotic – that is, data are 'cultured' in the process of being perceived; and
2. This semiotic information is often 'processed' again – compared further to complex symbol systems and evaluated according to existing 'knowledge'.

Més tard veurem algunes línies de recerca en matèria d'interacció agent-agent i agent-humà on els sistemes multiagent tenen control no només sobre el seu propi estat intern sinó també sobre el de les relacions que s'estableixen amb els altres agents i humans del sistema. En tant que aquests sistemes solen disposar també de mecanismes d'adaptació que els permeten aprendre de les experiències passades, amb l'argument de Hakken com a mínim podríem afirmar que aquests sistemes processen la informació d'una manera molt més semblant a la dels humans que no pas a la de la resta d'éssers vius.

Fins ara, possiblement els agents més famosos hagin estat els que es van enviar al planeta Mart abans de vehicles d'exploració de superfície. Atès que la comunicació entre Mart i la Terra triga uns quants minuts en arribar, els vehicles no podien estar teledirigits ja que a la Terra, igual que a la consciència humana, els esdeveniments arriben amb retard. Per tant, els humans enviaven directrius que després l'agent que controlava el vehicle havia de posar en pràctica enfrontat-se en temps real a l'entorn i prenent decisions que podien ser crítiques en el terreny inexplorat marcià.

L'exploració de Mart planteja algunes qüestions interessants. Els robots exploradors eren el resultat de la feina de múltiples equips formats per persones i sistemes digitals que havien fet allò en aquell moment i no en un altre gràcies a que factors que implicaven a la societat sencera (polítics, econòmics, tecnològics...) ho havien fet possible. Qui havia enviat realment aquells exploradors a un altre planeta? Uns tècnics amb noms i cognoms, una agència concreta o el

conjunt de la societat?. I en conseqüència, què mantenia el feedback amb els exploradors, processant la informació que enviaven, reconstruint a la Terra la realitat de Mart i retornant-los noves directrius?. Algú concret o una mena d'intel·ligència o consciència col·lectiva?.

L'exploració de Mart es basava en un diàleg entre dos sistemes autònoms. Els robots no eren una extensió dels tècnics de la NASA, estaven controlats per agents que tot i seguir les directrius marcades tenien autonomia, prenen les seves petites decisions i es comunicaven amb la Terra. Per a ells hi havia un sol interlocutor a l'altra banda. D'aquí se'n poden treure dues conclusions importants:

1. El conjunt d'equips humans i sistemes digitals que seguien la missió des de la Terra es comunicaven amb Mart com un sol interlocutor, per tant s'havien constituït com un sol organisme capaç d'interactuar amb l'exterior.
2. Els agents que eren a Mart interactuaven també com a interlocutors, no com a meres extensions tecnològiques, en la mesura que tenien la darrera paraula en el detall de la missió.

La tipologia d'agents més estudiada, però, no està a l'espai exterior sinó a internet i és la dels knowbots, que estan dissenyats per fer cerques d'informació a la xarxa. Alguns d'aquests agents impliquen l'existència de relacions socials indirectes, atès que els agents, a més de comunicar-se entre ells, es comuniquen també amb els usuaris. Hermans planteja una arquitectura en tres capes ⁶, que es podrien esquematitzar així:

comunicació humà 1 – agent 1
 comunicació agent 1 – agent 2
 comunicació humà 2 – agent 2

A la capa intermèdia és on es produeix la negociació estrictament entre agents, sense intervenció humana, durant la qual l'agent, com si es tractés d'un majordom cibernètic, intenta assolir els objectius fixats bescanviant informació amb altres majordoms.

Existeixen moltes tipologies d'agent. El Swedish Institute of Computer Science divideix el seu índex de recerca sobre agents en cinc categories:

Integració d'informació i informació compartida
 Resolució de problemes cooperatius i sistemes multiagent
 Agents portables
 Assistents personals, softbots i extracció de dades
 Simulacions i vida artificial

Com dèiem, ens centrarem en els sistemes multiagent, atès que ens interessen les seves possibilitats socials. Des de l'Information Sciences Institute de la Universitat d'Edimburgh⁷ es proposa l'acrònim RAP (group of Robots, software Agents and People) per definir l'estructura que haurien de tenir equips de treball on usuaris humans, software i hardware interactuessin estretament. Les aplicacions són tan diverses com optimitzar calendaris de reunions, ensinistrament, exploració espacial, operacions militars o gestió de catàstrofes, entre d'altres.

Lyengar i Truzskowski (2002), del NASA Goddard Space Flight Center, pronostiquen que els sistemes multiagent experimentaran un gran desenvolupament en les futures missions d'exploració del sistema solar per la reducció de costos i de riscos per als humans que això suposa. Proposen l'acrònim ASAP per sintetitzar el conjunt de característiques que hauran de complir aquests agents (Autonomia, Sociabilitat, Adaptabilitat i Persistència) que a grans trets coincideix amb la noció feble d'agent que comentàvem abans, i també assignen a cada una de les tasques relacionades amb les operacions espacials les característiques més concretes que cada tipologia d'agent hauria de complir. Es tracta d'una autèntica jerarquia social, amb líders, supervisors i treballadors. Aquests plantejaments impliquen resoldre dos tipus de problemàtiques: la interacció entre agents i la interacció entre humans i agents.

La interacció humà-agent està avançant en la direcció de proporcionar interfícies basades en llenguatge natural però també amb llenguatge no verbal (gesticulació, expressions facials...) per tal de millorar-ne la comunicació.

Des de l'Artificial Intelligence Center del Stanford Research Institute s'ha desenvolupat un llenguatge de comunicació entre agents (ICL) basat en reconeixement de llenguatge natural i dissenyat per a treballar en paral·lel amb d'altres canals sensorials com la veu o l'escriptura manual. L'estratègia de comunicació es completa amb els anomenats Facilitator agents, que s'encarreguen de resoldre situacions d'expressivitat complexa⁸.

En un sistema de múltiples agents on també intervenen humans cal que els agents disposin del que es coneix com autonomia ajustable, és a dir, la capacitat de cedir el control als humans. Sistemes com els RAP plantegen un problema relacionat amb el fet que no resulta fàcil que els agents cedeixin control quan al mateix temps s'estan coordinant entre ells, especialment quan l'execució de tasques a temps real és crítica. Per resoldre el problema que planteja l'autonomia ajustable s'han fet aproximacions diverses, com l'anomenada "estratègia de transferència de control", que intenta fer menys traumàtica la irrupció en escena dels usuaris humans (Scerri et al., 2003).

Una altra línia de recerca, que enllaça la comunicació persona-agent amb la d'agent-agent és el paper que hi poden jugar les emocions. Alguns sistemes basats en models d'avaluació cognitiva permeten concloure que introduint quelcom equivalent a les emocions es pot millorar la coordinació entre agents, atès que això permet a cadascun formar-se una imatge més clara de l'estat i de les prioritats del grup, i entre agents i persones, atès que les emocions proporcionen un alt valor comunicacional que no es pot suplir només amb llenguatge natural (Nair et al., 2003).

També s'han investigat i aïllat diverses patologies pròpies de la comunicació pura agent-agent. Jensen i Lesser (2002) n'han identificat algunes:

TOC (Tragedy of the Commons): quan un agent fa un ús excessiu d'un recurs compartit. Alguns investigadors suggereixen com a solució l'aplicació d'un sistema d'assignació de preus als recursos, és a dir, l'establiment d'una economia artificial.

Lock-in: s'observa que, tal com passa als sistemes biològics o a les societats humanes, la diversitat tendeix a beneficiar el sistema d'agents perquè els proporciona més potencial d'adaptació a les condicions canviats. El problema és que alguns agents tendeixen a adoptar les estratègies d'altres o a triar comportaments menys òptims com a resultat de perturbacions inicials. S'observa que alguns sistemes evolucionen en el sentit de premiar el conformisme desenvolupant una mena de pressió de grup.

Cycling: durant el procés de canvi d'estructures en resposta a canvis ambientals, alguns sistemes no aconsegueixen recuperar una situació d'equilibri i cauen en un canvi cíclic que consumeix molts recursos. És un procés equivalent a les carreres d'armaments entre països o l'evolució adaptativa de predadors i preses, que efectivament fan evolucionar les espècies però no introdueixen cap canvi net en la distància que les separa.

Blocking: un determinat agent pot quedar bloquejat perquè només un nombre concret d'agents de les seves característiques poden executar una tasca concreta i ja ho estan fent, o perquè la tasca que ha de realitzar un agent requereix que altres agents estiguin en un determinat estat que no es dona.

Per dur a terme aquestes recerques s'ha comptat, entre d'altres, amb eines com el MASS (Multi-Agent System Simulator)⁹, un entorn de simulació creat per posar a prova estratègies de coordinació/negociació entre agents, i el JAF (Java Agent Framework)¹⁰, una arquitectura que permet als agents d'una simulació operar de forma autònoma i comunicar-se entre ells tal i com passaria al mont real, ignorant que formen part d'una simulació.

Conclusió

Probablement encara és prematur parlar d'una cultura exclusivament cibernètica, però és evident que l'evolució del software autònom fa pensar que ja existeixen sistemes d'interacció prou rics que es podrien considerar autèntiques comunitats d'agents, o si més no, comunitats híbrides on els humans ja no són els actors principals.

La màquina analítica de Babbage, paradigma del domini de la raó i del determinisme, queda molt lluny dels sistemes d'acció simultània i distribuïda de tipus multiagent que actuen a les xarxes digitals d'avui en dia. No hauríem de subestimar les possibilitats dels agents, d'aquests bacteris del ciberespai, en el procés de modelat de les xarxes de comunicació.

És cert que la majoria dels sistemes multiagent comentats es basen encara en simulacions i estan dissenyats per a contextos molt concrets, però a la vista de com ha evolucionat la pròpia internet resulta altament probable que els sistemes multiagent acabin saltant a àmbits d'ús més general i es popularitzin. Un dels contextos on aquests sistemes podrien tenir èxit seria el de les simulacions socials basades en sistemes multiusuari massius, com per exemple The Sims Online, on podrien actuar com a moderadors facilitant la coordinació i la construcció d'espais socials entre els usuaris humans, o pel contrari exercir tasques de control i repressió. De fet, aquesta mena d'entorns es podrien convertir en autèntics laboratoris d'experimentació de relacions entre usuaris humans i agents.

Hem argumentat que la forma que tenim de construir la realitat no és significativament diferent dins i fora del món digital, i que la influència de les dinàmiques socials que tenen lloc a les xarxes de comunicació digitals té un abast global i va molt més enllà de l'experiència directa del usuaris connectats.

Al tractar els sistemes multiagent hem citat alguns estudis sobre organització social que ens podrien recordar els que van realitzar Reid (1994 i 1999) i Duval Smith sobre entorns socials humans online. També hem vist estudis orientats a resoldre l'encaix entre humans i comunitats d'agents. No estarien aquests investigadors analitzant la cibercultura més genuïna? Aquella on la tecnologia ja no és un simple mitjà o un faceta més de la dinàmica social, i adquireix la categoria d'interlocutor de ple dret?.

Com a éssers vius i com a éssers socials hem construït sofisticades membranes per interactuar de forma segura amb el món exterior. En la mesura que aquestes membranes s'han anat tornant més complexes han guanyat autonomia i imprevisibilitat, convertint-se elles mateixes en

una perllongació del món salvatge que preteníem deixar allà fora. Usant un terme extret de la pròpia IA, podríem dir que com a sistema, els humans patim de cycling. Durant el procés d'adaptació als canvis ambientals que nosaltres mateixos provoquem i que ens afecten, atès que el món és un sistema limitat, no aconseguim mantenir una situació d'equilibri i caiem en una mena de fugida endavant sense assolir un avantatge significatiu sobre el caos i la incertesa però malbaratant una gran quantitat de recursos durant el procés.

Entrelluquem, doncs, que més tard o més d'hora haurem d'investigar i de donar resposta a dues preguntes fonamentals: quina mena de relació establirem els humans amb sistemes digitals autònoms que ens acompanyaran a la feina o en l'oci i que permetran interaccions d'expressivitat complexa? Quin efecte tindrà això sobre la percepció de nosaltres mateixos i sobre l'organització de la societat? S'integraran dins de la nostra societat humana o romandran exclosos?. Quines emergències socials es derivaran d'aquesta relació i com afectarà això al desenvolupament de la pròpia tecnologia?.

NOTES

[1]: Multi-User Dungeon o bé Multi-User Domain, fòrums estructurats de xat en temps real basats només en text.

[2]: Per saber més coses sobre els experiments amb implants cerebrals es pot consultar:

<http://www.emory.edu/WHSC/HSNEWS/releases/feb99/022399brain.html>

<http://mouse.tercera.cl/2000/rep/04/21/03.htm>

[3]: El concepte de cervell global és un tema recurrent. De la mateixa manera que els sistemes complexos solen manifestar propietats emergents, sembla raonable creure que el conjunt d'interaccions socials que tenen lloc a la xarxa pot manifestar alguna mena d'emergència que alguns equiparen amb l'existència de pensaments a partir de l'activitat neuronal. Podeu ampliar informació a:

<http://pespmc1.vub.ac.be/SUPORGLI.html>

<http://pespmc1.vub.ac.be/GBRAIFAQ.html>

[4]: una altra exemple de la percepció popular de la informàtica com quelcom incert i no racional la trobem en alguns acudits, com ara el del químic, l'enginyer i l'informàtic:

“un químic, un enginyer i un informàtic van en un automòbil i es para el motor.

- Per l'olor segurament es tracta d'un mala combustió – raona el químic -. Caldria analitzar la composició de la benzina del dipòsit.
- Pel soroll jo diria que es tracta d'un problema de la transmissió – aventura l'enginyer -. Caldria desmuntar el motor i verificar que no hi hagi res mal ajustat.
- Tius, tius, tinc una idea! – exclama l'informàtic -. Què us assembla si sortim del cotxe i tornem a entrar?”

[5]: El treball sencer de Hermans sobre agents el podem trobar a:

<http://www.hermans.org/agents/index.html>

http://www.firstmonday.dk/issues/issue2_3/ch_123/

[6]: La descripció del model de tres capes es pot trobar a:

<http://www.hermans.org/agents/h42.htm>

[7]: El plantejament inicial dels RAP el podem trobar en un treball de Scerri et al. del 2002 que es pot llegir a:

<http://www.isi.edu/teamcore/tambe/papers/2003/rap-aamas03.pdf>

[8]: L'Artificial Intelligence Center és un lloc de referència en el desenvolupament de sistemes basats en agents. Es pot trobar més informació sobre el llenguatge ICL a:

<http://www.ai.sri.com/~oaa/>

[9]: Podeu trobar més informació sobre el MASS (Multi-Agent System Simulator) a:

<http://mas.cs.umass.edu/research/mass/>

[10]: Podeu trobar més informació sobre el JAF (java agent framework) a:

<http://shelob.cs.umass.edu/research/jaf/>

RECURSOS OFFLINE

ABBATE, J. (1999). **Inventing the Internet**.
The MIT Press: Cambridge.

CASTELLS, M. (2002). **The Rise of the Network Society**.
(Second Edition). Backwell Publishers Ltd. Malden.

DUVAL SMITH, A. (1999) **Los problemas de la resolución de conflictos en las comunidades virtuales** dins de Smith, M. A. i Kollock, P. (ed.) *Comunidades en el ciberespacio*. Barcelona. Editorial UOC. 2003.

EDELMAN, M. I TONONI, G. (2000). **El universo de la conciencia**.
Crítica. Barcelona. (2002).

ESCOBAR, A. (1994). **Welcome to Cyberia: Notes on the Anthropology of Cyberculture**.
Current Anthropology. Volume 35. Number 3.

HAKKEN, D. (1999). **An alternative to 'Computer Revolution' Thought**. [Cyborg@Cyberspace?](#): An
Ethnographer Looks at the Future. Ny/London. Routledge.

HELD, D.; Mc. GREW, A.; GOLDBLATT, D.; PERRATON, J.(2001). **Global Transformations**.
Polity Press. Cambridge.

HINE, C. (2000). **Internet as Culture and Cultural Artefact**. Virtual Ethnography.
SAGE Publications. Londres.

JONES, S. G. (1998). **Información, Internet y Comunidad: apuntes para una comprensión de la
comunidad en la Era de la Información** dins de Jones, S. G. *Cibersociedad 2.0*.
Editorial UOC. Barcelona. 2003.

LEWIN, R. (1992). **Complejidad**.
Tusquets Editores. Barcelona. 1995.

MATTELART, A.(2001). **Historia de la sociedad de la información**.
Paidós Comunicación. Barcelona. 2002.

MARGULIS, L i SAGAN, D. (1995). **Microcosmos**.
Tusquets. Barcelona. 1995.

NILSSON, N. J. (2001). **Inteligencia artificial**.
Mc.Graw Hill. Madrid.

PENROSE, R. (1989). **La nueva mente del emperador**.
Mondadori. Madrid. 1991.

PARÉS, N. i PARÉS, R. (2002). **Realidad Virtual**.
Graduat Multimedia. Fundació Politècnica de Catalunya. Fundació UOC. Barcelona.

REID, E. (1999) **Jerarquía y poder. El control social en el ciberespacio** dins de Smith, M. A. i Kollock,
P. (ed.) *Comunidades en el ciberespacio*. Editorial UOC. Barcelona. 2003.

RECURSOS ONLINE

Agents: Open Agent Architecture

<http://www.ai.sri.com/~oaa/>

Agents: diferents tipologies

Agent-Based Systems. Swedish Institute of Computer Science

<http://www.sics.se/isl/abc/survey.html>

Agents: AIRG Multi-Agent Systems Webliography

Llistat de recursos

<http://www.cs.wpi.edu/Research/airg/Agents-hotlist.html>

Agents: Multi-Agent Systems

Llistat de recursos

<http://www.aaai.org/AITopics/html/multi.html>

Cervell:

Interfície cervell-ordinador

<http://www.emory.edu/WHSC/HSNEWS/releases/feb99/022399brain.html>

<http://mouse.tercera.cl/2000/rep/04/21/03.htm>

Cervell global:

<http://pespmc1.vub.ac.be/SUPORGLI.html>

<http://pespmc1.vub.ac.be/GBRAIFAQ.html>

HERMANS, B. Intelligent software agents on the internet

http://www.firstmonday.dk/issues/issue2_3/ch_123/

JAF:

<http://shelob.cs.umass.edu/research/jaf/>

JESSER, D i LESSER, V. (2002). Social Pathologies of Adaptive Agents.

Department of Computer Science. University of Massachusetts.

<ftp://mas.cs.umass.edu/pub/jensen-aaaiss2002.pdf>

LEVINE, R. (2002). Welcome to The Sims Online, virtual reality with a white picket fence. Wired.

<http://www.wired.com/wired/archive/10.11/simcity.html>

LEWIS JOHNSON, W., KOLE, S., SHAW1, E, i PAIN2, H. (2003). Socially Intelligent Learner-Agent Interaction Tactics.

http://www.isi.edu/isd/carte/publications/aied03/AIED_submission8p.doc

MASS:

<http://mas.cs.umass.edu/research/mass/>

NAIR, R., TAMBE, M. i MARSELLA, S. (2003). The role of emotions in multiagent teamwork: A preliminary investigation

dins de FELLOUS, J. i ARBIB, M. (ed) "Who needs emotions: The brain meets the machine" MIT Press.

<http://www.isi.edu/teamcore/tambe/papers/2003/Nair-emotions.pdf>

NASA.

Anàlisi dels requeriments i funcions de sistemes multiagent en les futures missions.

http://www.decisionsciences.org/DecisionLine/Vol33/33_5/33_5research.pdf

REID, E. (1994). **Cultural Formations in Text-Based Virtual Realities**
http://cv.uoc.edu/moduls/UW04_63009_00056/web/main/m3/framelec2.html

SCERRI, P., JOHNSON, L., PYNADATH, D., ROSENBLOOM, P. Si, M., SCHURR, N. i
TAMBE, M.(2003). **Getting robots, agents and people to cooperate: An initial study**. AAAI Spring
Symposium
<http://www.isi.edu/teamcore/tambe/papers/2003/RAP.pdf>

TURKLE, S. (1995). **What Are We Thinking About When We Are Thinking About Computers?**
http://web.mit.edu/sturkle/www/routledge_reader.html

VAYREDA, A. i ARDÈVOL, E. (2004). **Comunidades Virtuales**. Fundació per a la Universitat Oberta de
Catalunya. http://cv.uoc.edu/moduls/UW04_63009_00056/web/main/m3/lec4.html