

SENYALS, SÍMBOLS I SOROLLS

per

JOSEP AGUILAR I MARTÍN

Directeur de Recherche au
Laboratoire d'Automatique et d'Analyse
de Systèmes du C. N. R. S. Tolosa de Languedoc
(Toulouse) França
i
Centre d'Estudis Avançats de Blanes del C. S. I. C.
(Girona) Espanya

FINALITATS DEL NOSTRE ESTUDI

Comencem per donar a Cèsar allò que és del Cèsar. El títol d'aquest article és la traducció al Català del d'un llibre de J. R. PIERCE publicat per Hutchinson and Co. a Londres el 1962. L'autor era un dels col·laboradors de C. E. SHANNON i féu en particular experiments sobre la informació dins la Música, ajudat per Betty SHANNON, muller del conegut "pare" de la Teoria de la informació. M'he proposat de parlar sota el mateix títol, i amb un punt de vista similar, només modificat per l'acumulació de resultats teòrics i tecnològics d'aquesta vintena d'anys. La finalitat del meu raonament és de *comprendre* l'estructura, i els mecanismes de la informació, i també d'ajudar a *construir* sistemes de comunicació eficaços, és a dir fidels, intel·ligents i fins i tot evolutius.

La lluita contra el soroll és la primera tasca per a obtenir sistemes de comunicació eficaços; aquesta lluita pot ésser purament passiva i donar lloc a problemes anomenats "filtratge"; també pot incloure mecanismes dits intel·ligents, i en aquest cas pren una forma activa utilitzant propietats estructurals dels senyals i de llurs generadors, com són les gramàtiques dels llenguatges, o la física dels emissors de senyals. Finalment, la lluita contra el soroll també pot utilitzar estratègies evolutives com són l'adaptabilitat i

l'aprenentatge, l'una consistent a tractar de seguir les modificacions eventuals de les fonts de senyals o missatges, i l'altre utilitzant la memorització de les situacions "viscudes" per a resoldre problemes futurs.

Aprofitaré aquest treball per establir una espècie d'"història natural" de la **Informació** en el sentit físico-matemàtic d'aquesta accepció.

ORGANITZACIÓ D'AQUEST ARTICLE

Tractarem la informació d'acord amb el títol, és a dir posant de manifest els seus elements, **senyals**, **símbols** i **sorolls**, així com els aspectes estructurals i mecanicistes que permeten d'establir les relacions entre aquests elements. L'estudi de la **informació** es pot justificar per la seva presència a tots els processos físics, naturals o artificials i fins i tot als éssers vivents; tanmateix vull fer aquí una defensa d'aquesta mena d'estudi i de l'absoluta necessitat per a la humanitat de posar totes les seves forces en el seu desenvolupament, i per això l'article començarà amb una *Introducció històrico-filosòfica*. Després parlaré dels *components de la informació* amb els seus aspectes estructuralistes i lingüístics. Un capítol ens permetrà de definir quantitativament la Informació, la seva transmissió i la complexitat del sistema, és a dir la *Mesura de la informació*. Els aspectes lligats al problema de l'*extracció de la informació* seran tractats sota el punt de vista estocàstic; i finalment dedicarem un capítol al mecanisme d'*aprenentatge i autoaprenentatge* on apareix la paradoxa dita de "l'estructuració pel soroll".

INTRODUCCIÓ HISTÒRICO-FILOSÒFICA

El món ha arribat a una riquesa tecnològica capaç de resoldre la majoria dels problemes materials de la Humanitat, altrament dit: disposem avui d'un **hardware** bastant bo. Malauradament, grans conflictes i cataclismes són sempre presents i hem de confessar que la meitat de la població de la Terra no arriba a menjar suficientment, la qual cosa és la primera necessitat vital.

Aquesta situació és deguda al mal funcionament d'aquest **hardware**, o sigui a defectuositats en matèria d'informació, de comunicació i de llur utilització; aquesta situació crítica de la Humanitat és assimilable a una deficiència palesa de **software**.

Els grecs desenvoluparen la cultura i la civilització, parts bàsiques d'aquest **software**, i ho pogueren fer perquè no tenien grans problemes ni de comunicació ni d'informació, gràcies a llur petit nombre, a la facilitat de comunicació per via del Mediterrani, i a la manca de desig de fer funcionar un sistema massa complex, i així mantingueren sempre l'individu com a base fonamental de llur món.

Més tard els romans, amb les conquestes de Cèsar, es trobaren en presència d'un territori extens i diversificat molt més complex. Aquest Imperi s'hauria esvaït i mai la civilització greco-romana no ens hauria arribat si l'emperador següent, August, no hagués intuït, amb geni, que calia fer participar els extrems més reculats de l'imperi en els mecanismes de govern de Roma mitjançant informacions i no pas presència material d'homes. Comprengué que el desplaçament de la informació té avantatges molt superiors al transport d'homes o de coses, i que era inútil que els tribuns d'Itàlica, prop de l'actual Sevilla, on ell nasqué, es desplaçessin contínuament a Roma, quan la informació podia transitar molt més ràpidament i econòmica per un sistema de correu amb postes de cavalls i relleus.

Aquest exemple donat per la nostra pròpia història mediterrània, i que portà la Pau i la Prosperitat durant un gran nombre d'anys a la major part del món conegut, dona un èmfasi molt particular a l'esforç indispensable sobre la tecnologia de la Informació i la Comunicació, única via per a resoldre els greus problemes que afecten avui el nostre món molt més complex.

L'optimisme ha d'ésser present quan constatem els grans avenços actuals en la transmissió i el tractament de la Informació, oberts tant per les innovacions tecnològiques de l'Alta Integració de circuits i components electrònics, com pels progressos indiscutibles en el tractament de la Informació que ens permeten de simular la intel·ligència i l'habilitat d'experts en tasques d'una complexitat cada dia creixent. Cada dia tenim a la nostra disposició millor **software**, i per tant més possibilitats de resoldre els problemes que hem esmentat.

1. ELS COMPONENTS DE LA INFORMACIÓ

Ferdinand de SAUSSURE, que pot ésser considerat el fundador de la lingüística estructuralista moderna, i fins i tot de les bases de tota la filosofia estructuralista incloent-hi Roland BARTHES i Claude LÉVI-STRAUSS, ens ensenya que el signe lingüístic és *arbitrari, lineal i diferencial*, i és la conjunció de **significat**, i **significant**; avui hem après a generalitzar lleugerament aquests conceptes i reconeixem un signe lingüístic generalitzat, o simplement "informatiu", cada cop que retrobem aquesta conjunció entre un element d'un Univers Semàntic i un element d'un Univers Semiòtic, com apareix a la figura 1. L'Univers Semàntic és una finestra del món real on es troben tots els conceptes significatius o que tenen sentit, o informatius, d'un camp de coneixements. L'Univers Semiòtic és el conjunt de **signes**, o manifestacions distingibles del món físic, amb una estructura d'associació que per extensió és anomenada "lingüística" o "gramatical".

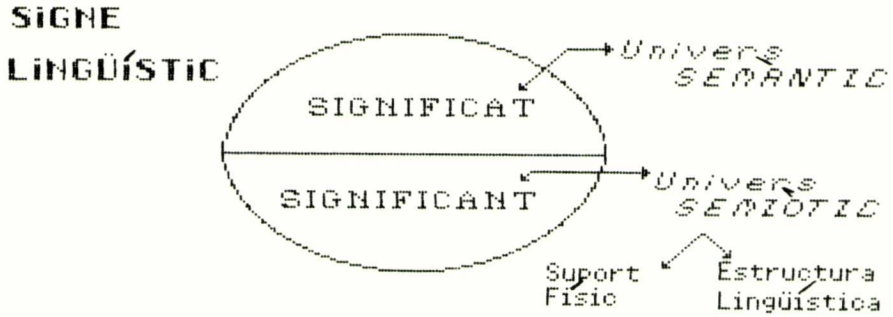


Fig. 1

L'associació del **significat** amb el **significant** és forçosament, a la majoria dels casos, totalment *arbitrària*, sense que això afecti la funció de comunicació. L'aspecte *lineal* es troba a la paraula oral i als senyals electromagnètics o electrònics, però menys a la paraula escrita, o als ideogrames o altres signes gràfics. Malgrat tot, la natura seqüencial de l'anàlisi visual pot fer dir que aquesta característica és quasi sempre present. La necessitat per al signe lingüístic d'ésser *diferencial* ve de la necessitat d'ésser intel·ligible, és a dir de poder separar-se d'un fons o substrat, consistent en la matèria suport físic del significant. La Teoria de les Catàstrofes, del matemàtic René THOM, dona un quadre analític a aquesta "diferenciació" obligatòria del signe: cal que existeixin variacions suficientment grans en un espai o en un temps suficientment petits perquè pugui aparèixer una "forma" perceptual. La famosa teoria de la "Gestalt" (forma, en alemany) tracta d'analitzar l'essència i l'existència d'aquestes formes en funció de formes suposadament preestablertes al nostre cervell, i fins i tot per extrapolació a la "consciència universal"; no entrarem ací en aquestes consideracions sibil·lines.

Per resumir direm que els senyals són els suports físics del signe, alhora que els símbols són els senyals dins una estructura lingüística. Podem trobar al *Tractatus* de L. WITTGENSTEIN la frase: "Cap signe no té significat tot sol", la qual cosa intueix l'existència de relacions entre signes, altrament dit de gramàtica, o més generalment estructura lingüística en sentit general.

Per il·lustrar aquests conceptes farem ací una construcció visual d'un univers informatiu a partir del Caos inicial. La nostra tasca serà, guardant totes les proporcions de talla i de respecte que cal considerar, similar al procés descrit al **gènesi** quan Déu, a força de separar o diferenciar matèries, acaba creant un món molt complexament estructurat i el qual, si bé globalment no hi ha encara ningú que li hagi trobat cap sentit objectivament, té sens dubte una multitud d'elements informatius dels quals aquest article és, si més no, un modest exemple. A la figura 2a podem veure la imatge d'un suport físic indiferenciat: es tracta d'un full verge on no s'ha produït cap "catàstrofe", en el sentit de R. THOM. Aquest full serà per a nosaltres una petita

part, modesta, del Caos inicial del qual Déu comença per separar la Llum de les Tenebres.

La figura 2b ens dona una imatge del mateix full de paper després d'haver estat sotmès a una allau de "catàstrofes", en el sentit que hom vulgui, és a dir després d'haver estat rebregat, i ben rebregat. Els plecs del paper són, cada un, un "sistema catastròfic" perceptible amb l'ull a causa de les variacions brusques de pendent relativament a la direcció dels raigs de llum. Ja tenim el suport físic dels senyals, o dels significants, però encara no són **sím-bols**.

Per fer la figura 2c hem extret de la figura 2b les arestes més perceptibles, i ens hem quedat amb una xarxa de ratlles similar a un mapa de carreteres o de trens. Alhora hem imposat, d'una manera arbitrària, a aquests senyals una estructura de tipus lingüístic, formada per un alfabet basat en els senyals perceptibles, al qual hem afegit dos símbols més, l'un inicial **S** i l'altre final **A**, i per a una millor representació hem posat als símbols naturals les etiquetes **P** si un segment està connectat pels seus dos extrems i **N** per al cas contrari. Finalment hem definit una gramàtica que consisteix a definir allò que és una frase correcta, en aquest cas qualsevol seqüència del tipus **SPPPPPPPPA**, amb la regla de reemplaçament de qualsevol **P** per una seqüència qualsevol **PPPPPPP**, o l'element nul que consisteix en la supressió d'una **P**. Per a una millor comprensió d'aquest procediment explicarem com poden ésser interpretats aquests símbols:

- P** = Passa
- N** = No passa
- S** = Sortida
- A** = Arribada

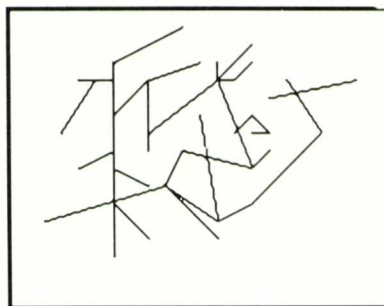
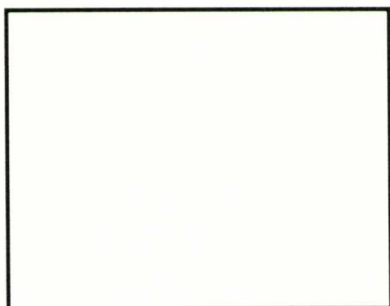
A la figura 2d hi ha representades dues frases, l'una correcta, o acceptable i l'altra inacceptable per la gramàtica. El missatge, o informació significativa transmesa per cada una d'aquestes frases, és el següent:

Per a la primera: "Es pot anar del punt **S** al punt **A** per aquest camí".

Per a la segona: "Aquest camí no permet de fer el viatge de **S** a **A**".

És fàcil de convèncer-se que aquesta informació pot ésser d'un gran interès per a un viatger, suposant que el dissenyador de mapes ha emprat els plecs del paper com a medi o suport físic dels seus signes.

Fins ara no hem tingut cap actitud hostil ni de la part del paper ni dels plecs obtinguts per la nostra acció "catastròfica". Suposem ara que al nostre petit simulacre de Gènesi també hem d'enfrontar-nos amb el Diable, tal com diuen que passà. Però, per tal de no posar el problema massa difícil per a les nostres possibilitats simplement humanes, no escollirem un Lucifer, portador de llum com diu la seva etimologia, i per tant un bon tros intel·ligent, no. Escollirem un bon Belzebuth, que no fa sinó enredar-ho tot sense gaire malícia, i veurem que ja ens donarà prou feina. La figura 2e és la imatge



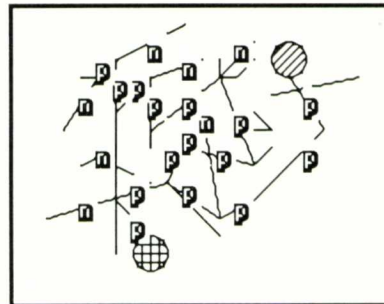
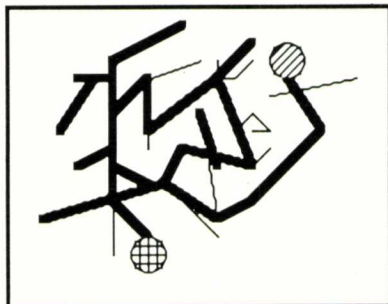
Semiòtica : suport homogeni

Semàntica : caos

(a)

(b)

conjunt de catàstrofes
signes



Semiòtica : selecció de senyals

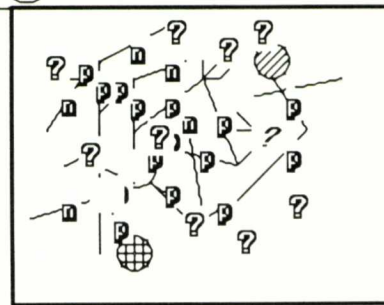
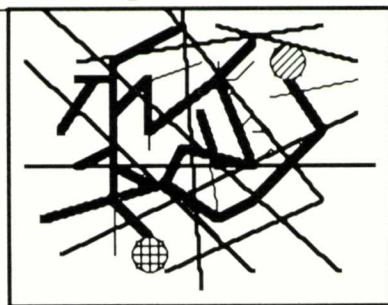
Semàntica : □=passa ■=no passa

(c)

(d)

frases de símbols

camins viables



Semiòtica : 

Semàntica : camins i soroll

(e)

(f)

frases ambigües

camins de viabilitat incerta

figures 2-a, 2-b, 2-c, 2-d, 2-e i 2-f
figura 2

del **soroll**, o com altrament diria Ton SALES, de la NOSA, per apropar-se al NOISE anglo-saxó. En el nostre univers el **soroll** és un conjunt de ratlles distribuïdes erràticament, raó per la qual li podríem dir Soroll blanc, que per als matemàtics vol dir de qualsevol manera, per oposició a un soroll que tingués una estructura predominant, és el cas de la llum blanca, on tots els colors s'hi troben, per oposició a la llum d'un cert color ben determinat.

La superposició del **soroll** de la figura 2e a la figura 2c, ens dona una nova classe de seqüències o frases que no sabem caracteritzar; són les frases *ambigües*. Els missatges precedents s'han transformat en missatges ambigus, i l'utilitzador eventual d'aquest mapa de camins ja no podrà tenir una confiança total en el sentit d'aquests missatges. Com lluitar contra aquest Dimoni anomenat **soroll**? En el capítol dedicat a l'extracció de la informació tractarem d'aquest tema, però ací és un bon lloc per a fer una anticipació. Si teníem d'antuvi una idea de com són els camins que busquem, és a dir de la impossibilitat per exemple de tenir angles massa tancats, o d'anar en una direcció parcial contrària al vent... és a dir si teníem una espècie de model generatiu dels camins, aleshores podríem determinar el grau de versemblança d'un camí, d'una seqüència o d'una frase dels nostres missatges. Al dibuix de la figura 2f han estat representats els camins trobats anteriorment, i es veu el camí inacceptable quan no hi havia soroll ara com a camí ambigu. Si afegim al nostre coneixement de l'estructura lingüística la restricció que consisteix a suposar improbables els camins que canvien massa ràpidament de direcció, podrem aleshores invalidar de nou el camí realment inacceptable, i hauréu sortit vencedors de la nostra lluita contra el Soroll.

2. LA MESURA DE LA INFORMACIÓ

No farem ací un curs sobre la teoria de la Informació, sinó que tractarem de comentar algunes de les fórmules que són sovint introduïdes per a mesurar la informació.

La **informació** d'un **símbol** dins un **missatge** ha estat definida per C. SHANNON com una funció positiva i creixent de la proporció entre la probabilitat del símbol a la seva recepció sobre la probabilitat a la seva emissió, és a dir:

$$IN[\text{símbol } x] = \left(\frac{\text{Prob}[x/\text{recepció}]}{f(\text{Prob}[x/\text{emissió}])} \right)$$

Per tant, quan un símbol arriba efectivament al receptor, la $\text{Prob}(x/\text{recepció}) = 1$. La funció $f(\cdot)$ escollida per Shannon, per raons lligades a la termo-

dinàmica i consideracions d'ordre lògic i fisiològic és el logaritme; per tant podem escriure

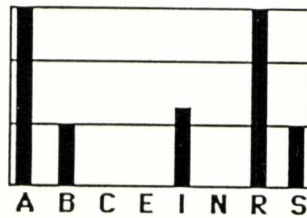
$$IN [x] = - \text{Log} (\text{Prob} [x/\text{emissió}]) = - \text{Log} (\text{prob} [x])$$

(Reemplaçant prob [x/emissió] per prob [x])

La informació mitjana d'un missatge serà calculada com a l'esperança matemàtica de la Informació per símbol, és a dir la suma ponderada per les probabilitats de cada símbol, sobretot l'alfabet:

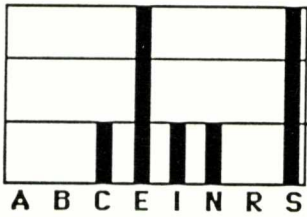
$$E [IN(x)] = - \sum_{\text{Alfabet}} \text{prob} [x] \cdot \text{Log} (\text{prob} [x])$$

Aquesta última expressió, per analogia amb la física i la termodinàmica, és anomenada **entropia** del missatge H (**missatge**). És una funció que té la propietat d'ésser màxima quan totes les probabilitats són iguals, i conseqüentment de valor $1/N$ si N és el nombre de símbols de l'alfabet. Cal remarcar tanmateix que missatges molt dissimilars poden tenir la mateixa entropia, i per tant que l'entropia és independent del valor semàntic, o sentit del missatge; a la figura 3 donem un exemple d'aquest fenomen.



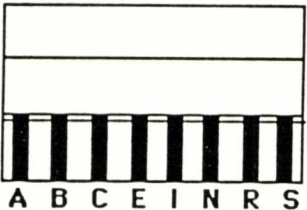
Missatge ARRIBARÀS

Entropia 1.46481638



Missatge ESSÈNCIES

Entropia 1.46481638



Missatge neutre i equiprobable amb el mateix alfabet

Entropia 2.07944154

Fig. 3

La Informació mitjana o Entropia pot ésser considerada la dissimilaritat entre un missatge concret, o una família de missatges, i un missatge hipotètic neutre o indiferenciat en el qual tots els símbols són equiprobables. Aquestes propietats de l'Entropia com a mesura de la informació la fa també inepta per a donar compte d'eventuals relacions d'imbricació, o jeràrquiques, que poden existir entre símbols, i que apareixen en formalismes com els de les gramàtiques formals o generatives.

Però, malgrat tot allò que la teoria de SHANNON deixa de banda, la seva potència com a eina quantitativa és suficient perquè s'hagi tractat d'estendre el seu camp d'aplicació, fins a la Biologia, l'Ecologia, i fins i tot a camps de la Filosofia.

Enunciarem ací uns teoremes bàsics d'aquesta teoria:

TEOREMA DE SHANNON:

La quantitat d'informació transmesa per un **canal** decreix, (no pot fer altra cosa que decreixer), d'una quantitat igual a l'ambigüïtat introduïda pel **soroll**. La figura 4 representa l'esquema d'aquesta situació i les fórmules corresponents a cada una de les entropies presents: Entropia abans transmissió $H(x)$, entropia pròpia del canal $H(x/y)$, que reflecteix l'ambigüïtat introduïda, i entropia del missatge a la recepció $T(x,y)$. Tindrem $T(x,y) = H(x) - H(x/y)$ i evidentment si el canal fos perfecte tindriem $H(x/y) = 0$ i $T(x,y) = H(x)$.

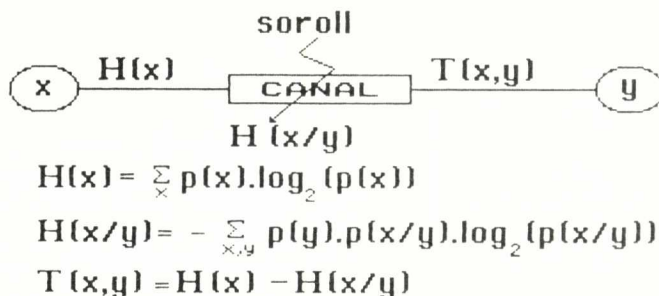


Fig. 4

TEOREMA DE VON FOERSTER (PRIGOGINE, ATLAN)

La quantitat d'informació total d'un **sistema** és la suma de la informació rebuda més la informació d'ambigüïtat introduïda pel soroll: és a dir

$$IN(x,y) = H(x) + H(x/y)$$

La quantitat total d'informació mesura el grau d'improbabilitat que el conjunt sigui l'obra de l'atzar. És per això que ha pogut ésser considerada una mesura de la complexitat.

La quantitat d'informació, per a un **missatge**, expressat en "bits" a partir de logaritmes de base 2, és el nombre mínim mitjà de símbols binaris necessari per a traduir-lo del seu alfabet original. En el cas dels **sistemes**, aquesta quantitat representa el mateix concepte necessari per a descriure la composició del sistema, i per tant representa la complexitat composicional exceptuant la complexitat funcional lligada a les relacions funcionals i dinàmiques entre elements.

Podem remarcar que el Canal perfecte no introdueix ambigüitat, i per tant no fa perdre informació a la transmissió, però que, vist d'una forma global, tampoc no introdueix informació al sistema, i no n'augmenta la complexitat. A la figura 5 hem representat els tres casos corresponents a la transmissió perfecta, la transmissió amb soroll i la transmissió nul·la, i paradoxalment és aquest cas el que pot ésser considerat com a més carregat d'informació, o més complex, pel fet que els dos subconjunts A i B poden ésser totalment diferents, alhora que en el primer cas l'un és una còpia exacta de l'altre, i no aporta cap informació.

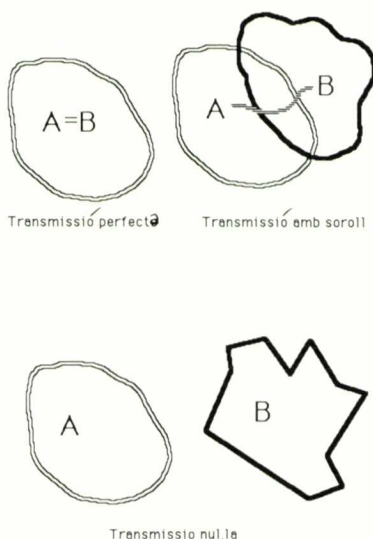


fig. 5

3. ELEMENTS SOBRE L'EXTRACCIÓ DE LA INFORMACIÓ

Per tal que un missatge pugui ésser informatiu cal que quelcom l'hagi elaborat segons unes regles comunes a l'emissor i al receptor; aquestes regles poden ésser una simple paleta de colors, on cada un porta una significació pròpia; aquest tipus de codi és fonamentalment pobre i no pot donar lloc sinó a un nombre limitat de missatges; aquesta és la raó per la qual els éssers intel·ligents es comuniquen mitjançant llenguatges articulats. Etimològicament "articulat" ve del llatí "artus" = "membre" o "tros", ell mateix provinent del grec "αρω" = "jo dispo". Consisteix doncs a jugar al *puzzle*, és a dir, a disposar trossos per a crear missatges.

Fent una ràpida anàlisi del llenguatge humà, trobem les dues articulacions següents:

1a. articulació o disposició dels símbols = sintaxi

2a. articulació o disposició dels sons físics = fonologia, dinàmica acústica.

A la primera articulació correspon la gramàtica de la llengua, i a la segona un generador de senyals, les característiques del qual són donades pel sistema físic que pot ésser la larinx humana o l'amplificador de senyals electrònics. Si ens restringim a considerar el soroll com un missatge generat de manera anàloga per una altra font de missatges, tindrem una ajuda per a discernir entre el missatge informatiu i els altres, que consisteix a tractar de detectar l'estructura de la font de cada missatge; si aquesta estructura correspon a la de la font dels missatges informatius, acceptarem el missatge, i el rebutjarem en cas contrari.

A la figura 6 representem l'esquema fonamental de la lluita contra el soroll basat en aquest principi. És aleshores necessari de disposar d'una rèplica o model del generador de missatges informatius. El corrector consisteix a crear l'acció corresponent a la informació triada correctament, és a dir on la part provinent de canals o generadors de soroll ha estat rebutjada.

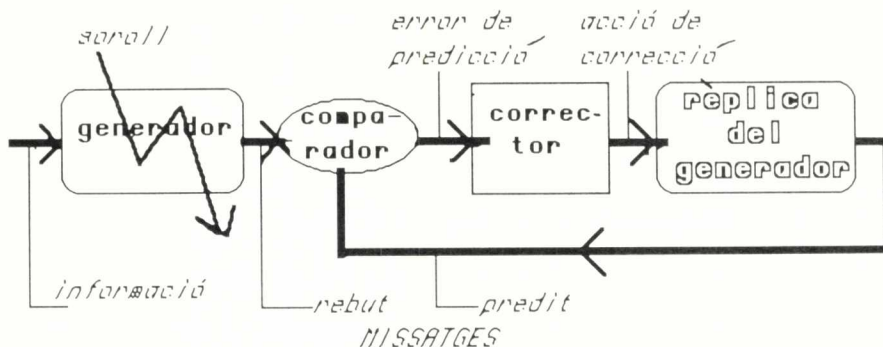


Fig. 6

Ací podem fer una remarca paradoxal: un missatge predictable no porta cap informació, i en canvi pot tenir sentit; en canvi el soroll pur és totalment impredecible però no té sentit, només porta informació sobre la seva pròpia realització.

El rol fonamental que tenen els models de predicció és palès en l'esquema de **predicció/correcció** que hem apuntat a la figura 6. És per això que creiem convenient d'aturar-nos a analitzar els models de predicció estocàstics, que són els que han estat estudiats d'una forma matemàtica més completa, i recordant que l'etimologia de la paraula "estocàstic" és el grec "στοχος" = "meta, o objecte a atènyer". D'una manera molt ràpida i superficial, i per a ajudar a comprendre les funcions relatives a cadascuna de les definicions més freqüents en la teoria dels processos estocàstics, direm que existeix un paral·lelisme entre la noció de sistema dinàmic o evolutiu dins el temps, i per tant parcialment predictable, i la noció de procés de MARKOV; l'estat d'un procés de Markov evoluciona de tal manera que quan hom coneix, o mesura, o observa, el present, els esdeveniments futurs no depenen de la trajectòria passada, tan sols depenen del present conegut. Per tant la sola observació del present és necessària per a fer pronòstics sobre el futur, i de res no serveix de tenir en compte el passat. La representació matemàtica d'aquests processos en el marc dels missatges quantificables pren formes diferencials:

$$dx = m(x, t) dt + g(x, t) dw$$

o incrementals:

$$x(t+1) = m(x(t), t) + g(x(t), t) v(t)$$

on $w(t)$ i $v(t)$ representen sorolls absolutament impredecibles o blancs.

La llei de probabilitat condicional de transició d'un estat x_0 a l'instant t_0 , cap a un estat x a l'instant t pot donar lloc a una densitat de probabilitat

$$p[x, t] = \text{pr}[x, t/x_0, t_0]$$

i satisfà una equació amb derivades parcials de tipus difusional:

$$\frac{\partial p(x, t)}{\partial t} = f\left(\frac{\partial p}{\partial x}, \frac{\partial^2 p}{\partial x^2}, x, t\right)$$

Aquests processos són uns models matemàtics d'utilització freqüent, però no sempre la realitat d'un senyal pot amotllar-se a aquestes hipòtesis. Per a generalitzar l'aplicació, sovint hom considera que un procés markovià exis-

teix d'una manera latent però que l'observació verament assolible n'és una funció $y(t) = h(x(t), \varepsilon(t))$.

La variable $\varepsilon(t)$ correspon a un soroll que ve a pertorbar la recepció correcta del missatge, i és anomenada sovint: soroll d'observació, o de mesura.

La figura 7 recull el mecanisme fonamental dels models de generació de senyals a partir de processos de Markov. A la part inferior de la figura hom ha tractat d'esquematzar el mecanisme de **predicció/correcció** que permet de lluitar contra els sorolls en aquesta situació: la seva explicació sintètica és la següent:

$$IN(t) = \{ \dots y(t-1), y(t) \}$$

gràcies a la seqüència es tracta de predir $\hat{y}(t+1)$; la diferència o comparació entre $\hat{y}(t+1)$ i el valor observat $y(t+1)$ permet de fer una acció sobre la transició de l'estimador $\hat{x}(t)$ a $\hat{x}(t+1)$, el qual permetrà de fer una nova predicció $\hat{y}(t+2)$, i així successivament.

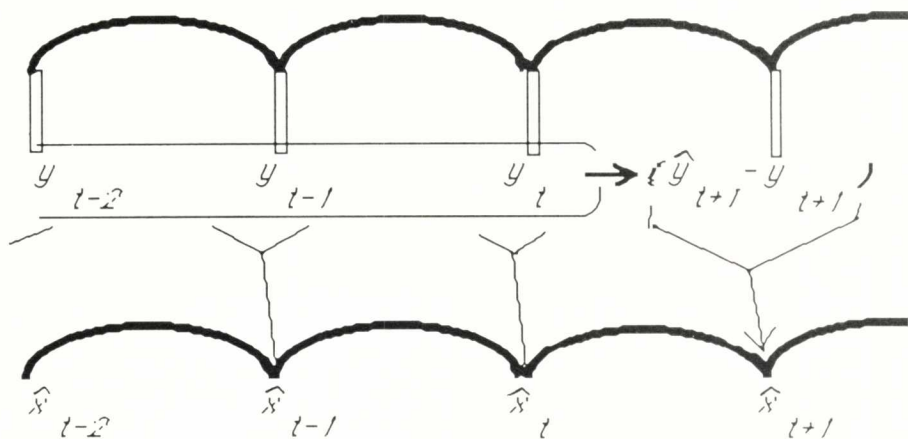


Fig. 7

En llenguatge probabilístic podem escriure que el pas de $x(t-1)$ a $x(t)$ ens és conegut per la probabilitat de transició: $pr[x(t) / x(t-1)]$, i que l'observació també correspon a una probabilitat condicionada pel procés markoviana $x(t)$, és a dir que podem escriure $pr[y(t) / x(t)]$. Amb aquestes notacions el problema de filtratge és resolt per una recurrència sobre les probabilitats mateixes que donen informació sobre l'estat $x(t)$ a cada instant, és a dir sobre

$$pr[x(t) / y(t), y(t-1), y(t-2), \dots y(t_0)] = pr[x(t) / IN(t)]$$

La fórmula de BAYES per a dos esdeveniments lligats,

$$\text{pr}[B/A] = \frac{\text{pr}[A/B] \cdot \text{pr}[B]}{\text{Pr}[A]}$$

ens permet d'escriure ací:

$$\begin{aligned} \text{pr}[x(t)/IN(t)] &= \text{pr}[x(t)/y(t), IN(t-1)] = \\ &= \frac{\text{pr}[y(t)/x(t)] \cdot \sum_x \text{pr}[x(t)/x(t-1)] \cdot \text{pr}[x(t-1)/IN(t-1)]}{\text{factor normalitzant la suma a 1}} \end{aligned}$$

Encara que sembli que aquesta fórmula recurrent no manifesta l'esquema fonamental de **Predicció/correcció**, es pot interpretar de la mateixa manera considerant que la predicció porta sobre la probabilitat de $x(t)$, i és donada per:

$$E[\text{pr}[(x(t)/IN(t-1))]] = \sum_x \text{pr}[x(t)/x(t-1)] \cdot \text{pr}[x(t-1)/IN(t-1)]$$

Quant a la correcció, consisteix en l'operació dita d'inversió bayesiana de la probabilitat de les causes després d'haver observat un dels seus efectes, i que es tradueix pel producte per $\text{pr}[y(t)/x(t)]$ seguit per una normalització. Com a conclusió a aquest capítol, volem mostrar que mitjançant el mecanisme elemental de **Predicció/correcció**, tenim una base per a la lluita contra el soroll que pot prendre aspectes passius, és a dir la simple eliminació dels senyals imprevistos, o aspectes adaptatius que consisteixen a tenir en compte les possibles variacions del canal de transmissió; aquesta situació no ha estat considerada explícitament ací, però és evident que, dins les fórmules probabilistes que hem exposat últimament, el terme $\text{pr}[x(t)/x(t-1)]$, gràcies al seu condicionament per $x(t-1)$, pot incloure modificacions conegudes al moment $t-1$, i conferir aleshores l'adaptativitat al sistema recurrent. D'un punt de vista algorítmic, aquesta introducció no posa cap problema important; tanmateix fonamentalment pot fer perdre el caràcter markovià a l'estimació de $x(t)$, i causar certes dificultats d'anàlisi.

4. APRENENTATGE I AUTOAPRENENTATGE

Què és aprendre? Acumular coneixements, sembla la resposta més natural. Si considerem que "coneixements" és sinònim d'"informacions", sembla que podrem aplicar aquests conceptes a qualsevol sistema artificial, i en particular als sistemes anomenats *intel·ligents* de la família dels ordina-

dors. Ràpidament ens adonarem que aquesta definició ens portarà a situacions inacceptables des d'un punt de vista tècnic i fins i tot pràctic i funcional. En efecte, les informacions que poden arribar a un sistema poden ésser tan nombroses que la seva capacitat serà aviat insuficient; d'altra banda, encara que arribem a emmagatzemar totes les informacions físicament, el procés de llur extracció per a ésser emprades serà cada cop més complex i lent, a mesura que el nombre d'informacions es faci més gran.

Aquestes reflexions ens fan dir que "coneixements" i "informacions" no són exactament la mateixa cosa. Coneixements són informacions estructurades, i, damunt de tot, resumides. El resum de la informació es defineix com a l'extracció dels caràcters essencials a la seva reconstrucció, tenint en compte l'ús que hom n'ha de fer ulteriorment. No es pot parlar d'aprenentatge sense finalitat o sense criteri previ d'agregació de la informació.

Per no estendre massa aquest article, ens limitarem a l'aprenentatge de conceptes quantitativs als quals poden ésser aplicades les operacions bàsiques de l'aritmètica, sumes restes i productes, així com comptatges d'esdeveniments. Mecanismes d'aprenentatge similars poden ésser introduïts en situacions no quantitatives, però les operacions matemàtiques necessàries necessiten una teoria més complexa. Limitant-nos al caire quantitativ, podem considerar l'acció de reemplaçar un conjunt de quantitats pel seu valor mitjà com una de les maneres més elementals d'efectuar un "resum", i per tant considerarem el mecanisme del càlcul de mitjanes un mecanisme fonamental de l'aprenentatge.

Si tenim una seqüència de valors $\{x(1), x(2), x(3), \dots, x(t)\}$ la mitjana serà

$$m(t) = \frac{1}{t} \sum_{s=0}^t x(s)$$

Correspon a una suma on tots els termes estan ponderats per $1/t$.

— Hom pot fàcilment establir una recurrència per calcular $m(t+1)$ a partir de $m(t)$ i de la nova informació o mesura $x(t+1)$:

$$m(t+1) = (t/(t+1)) \cdot m(t) + (1/(t+1)) \cdot x(t+1)$$

En aquesta nova fórmula cal remarcar dues coses: que la nova informació es ponderada pel factor $1/(t+1)$, i que no cal precisar la inicialització d'aquesta recurrència perquè $m(1) = 0 \cdot m(0) + 1 \cdot x(1) = m(1)$. És una equació recurrent que s'autoinicialitza amb el primer valor de la seqüència. Aquesta recurrència pot ésser escrita d'una manera que faci palesa la presència del mecanisme que ja hem apuntat de predicció i correcció:

$$m(t+1) = m(t) + (1/(t+1)) \cdot (x(t+1) - m(t))$$

Si considerem $m(t)$ un resum de la informació donada per la seqüència de x 's fins a l'instant t , podem dir que la predicció que podem fer sobre el futur valor $x(t+1)$ hom considera que serà igual a $m(t)$, és a dir, $\hat{x}(t+1) = m(t)$.

L'error de predicció, després d'haver mesurat $x(t+1)$, és doncs $(x(t+1) - m(t))$. En absència de nova informació, podríem considerar $m(t+1) = m(t)$, però coneixent que hem comès un error, corregirem aquesta predicció fent-hi repercutir una fracció d'aquest error, i en el cas de la mitjana, per a ésser coherent amb la seva definició cal que ponderem aquest error per un factor variable amb funció de t , $(1/(t+1))$; aquest valor de ponderació és decreixent, i per a instants molt distants de l'origen del temps de la mesura, és a dir per a t molt gran, agafa valors excessivament petits, i no transmet pràcticament informació.

Aquest fenomen pot ésser anomenat *envelliment* de l'aprenentatge, i tothom sap que com més vell es fa un ésser, més difícil li és d'aprendre noves coses. Inversament podem notar que les primeres observacions tenen un pes molt gran per a l'aprenentatge i podem anomenar aquest període inicial, *infància* de l'algorisme, que, com la infància humana, no té una data precisa per al seu acabament i transició al règim estable o madur.

El mecanisme d'aprenentatge que hem exposat a partir del càlcul de la mitjana es retroba idèntic si hom considera com a aprenentatge la determinació de la freqüència d'un esdeveniment, la qual cosa dona una idea molt propera a la de la seva probabilitat d'aparició, noció estretament lligada a la seva predicció. Si sobre t unitats de temps un esdeveniment ha ocorregut $n(t)$ vegades, direm que la seva freqüència és

$$f(t) = \frac{n(t)}{t}, \text{ i podem escriure } n(t) = \sum_{s=0}^t \delta(s)$$

on $\delta(s)$ és una variable que pren el valor 1 si l'esdeveniment ocorre a l'instant s , i 0 altrament. Per tant, es tracta d'una mitjana anàloga a la precedent, i que porta sobre la seqüència $\{\delta(1), \delta(2), \delta(3), \dots, \delta(t)\}$.

Podem escriure la mateixa fórmula recurrent de predicció/correcció que és:

$$f(t+1) = f(t) + (1/(t+1)) \cdot (\delta(t) - f(t)).$$

Amb aquest algorisme podem dir que, si un esdeveniment té la probabilitat d'ocórrer igual a P , a cada instant t tenim una estimació d'aquest valor, que és el resultat de l'aprenentatge, i que és donat per la freqüència $f(t)$ observada.

A causa del fet que el factor de ponderació $(1/(t+1))$ és decreixent i positiu, i que l'error de predicció $(\delta(t) - f(t))$ és en valor absolut inferior a 1,

hom pot afirmar que $f(t)$ és sempre convergent; aquesta propietat, que és necessària per a assegurar l'eficiència de l'aprenentatge, presenta l'inconvenient de produir aprenentatge encara que la seqüència d'observacions no correspongui a una veritable informació, és a dir que l'algorisme pot donar la sensació d'aprendre en absència d'informació.

La paradoxa de la generació d'informació espontàniament a partir de l'algorisme d'aprenentatge pot ésser exemplificada a partir de la situació següent: suposem que tenim un generador d'esdeveniments aleatoris amb una probabilitat determinable, per exemple generador de nombres aleatoris binaris, 0 o 1, i suposem que apliquem l'algorisme d'aprenentatge de la freqüència a l'esdeveniment (1), ell mateix generat a partir d'una probabilitat $pr(1) = f(t)$. Aquesta situació és representada per la figura 8. És un cas particular de retroacció probabilista. Aquest experiment pot ésser fet de la manera següent:

Tenim dos dipòsits de boles, l'un ple de boles blanques, i l'altre de boles negres en nombre pràcticament infinit; tenim una urna amb una bola blanca i una bola negra. Efectuem un tiratge aleatori de l'urna, i procedim després a l'acció següent: introduïm a l'urna la bola que ha sortit, conjuntament amb una altra del mateix color provinent dels dipòsits. Aleshores la composició de l'urna és de 2 boles d'un color i 1 de l'altre; tornem a fer un tiratge aleatori, i l'acció es repeteix indefinidament. El resultat d'aquest experiment és sempre un valor límit de la composició de l'urna, és a dir que el procediment convergeix; tanmateix convergeix cap a un valor aleatori, és a dir que tant es pot arribar a una composició límit de 0.5 boles blanques, com de qualsevol altre valor. La distribució de probabilitat d'aquests valors no és equiprobable, però presenta una zona molt plana al mig de l'interval (0,1).

A la figura 9 han estat representades diverses trajectòries de $f(t)$ obtingudes amb la mateixa inicialització, és a dir una bola de cada color. A la figura 10 hom dóna la forma de la distribució de probabilitats del valor límit, com a límit dels valors assolibles amb un nombre creixent de tiratges.

Aquest exemple algorísmic ens demostra que és possible de crear informació, o de manera equivalent estructura, tot i estant sotmès a una observació provinent d'una font on la informació és absent, font caòtica.

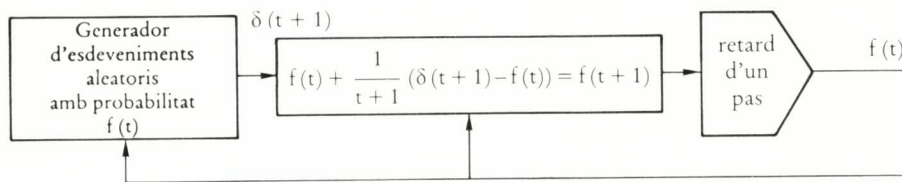
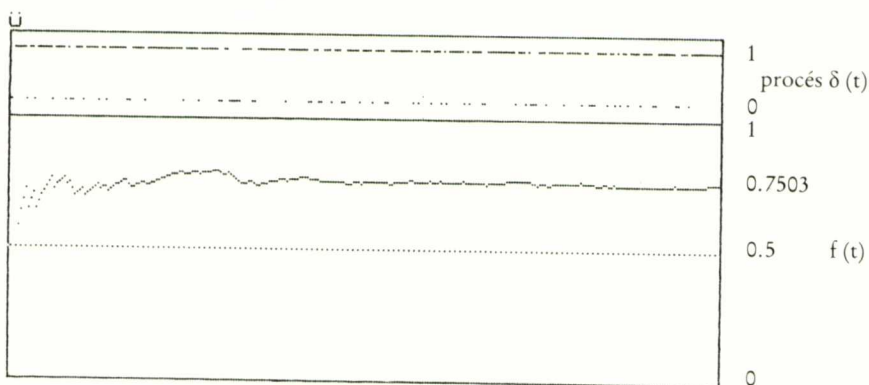
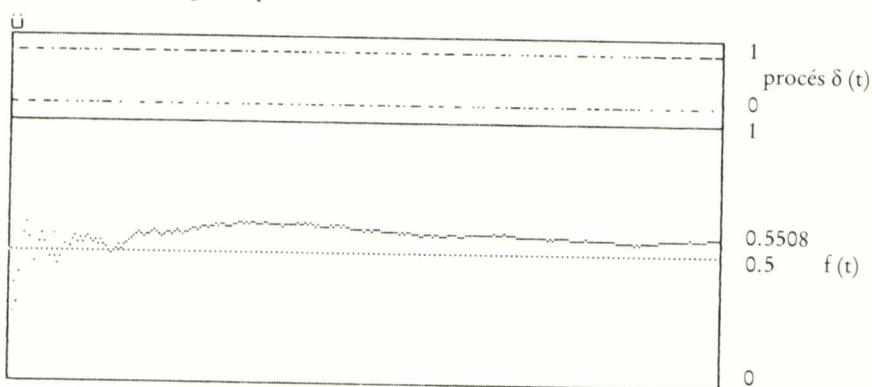


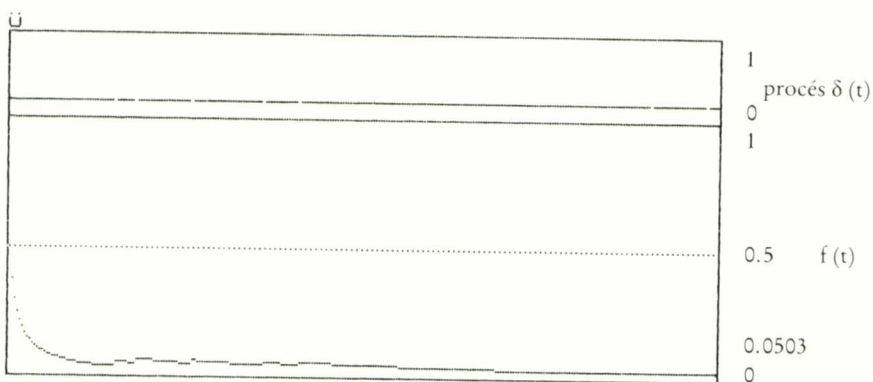
Fig. 8



Trajectòria convergent cap a 0.7503



Trajectòria convergent cap a 0.5508



Trajectòria convergent cap a 0.0503

Tres trajectòries de $f(t)$ per a la mateixa inicialització del procés $\delta(t)$ $f(0) = \frac{1}{2}$ i per tant prob
 $[\delta(t) = 1] = \frac{1}{2}$.

Fig. 9

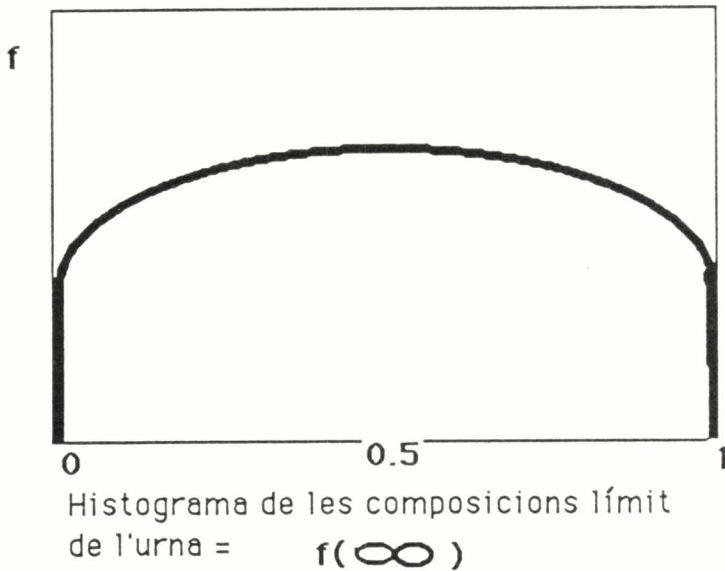


Fig. 10

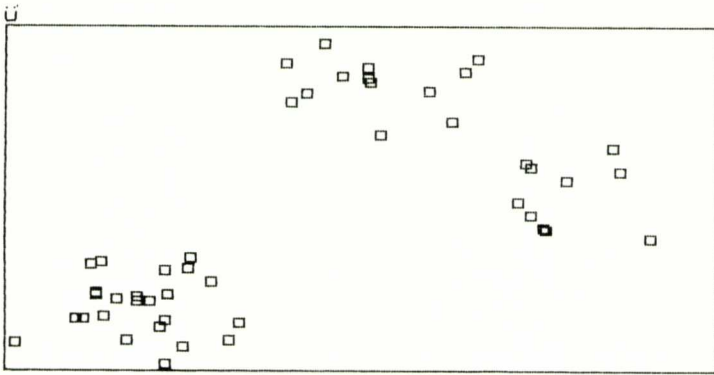
Recentment, matemàtics, biòlegs i informàtics s'han preocupat del fenomen que Henri ATLAN ha anomenat "ordre a partir del soroll". Aquest fenomen paradoxal, a més d'ésser present a bon nombre de problemes tecnològics, com el dels fregaments sòlids, o el del "dopatge" de semiconductors, dóna una hipòtesi per a la comprensió de la morfogènesi i fins i tot la cosmogonia o evolució primera de l'univers.

Hom pot, per aquest artifici, justificar i comprendre l'existència d'un món estructurat i parcialment comprensible que hagi pogut néixer del caos. Aquesta filosofia ens portaria a discussions teològiques que surten del nostre propòsit.

Per acabar aquesta exposició elemental d'allò que són els Senyals, Símbols i Sorolls, donarem una explicació de les figures 11 i 12.

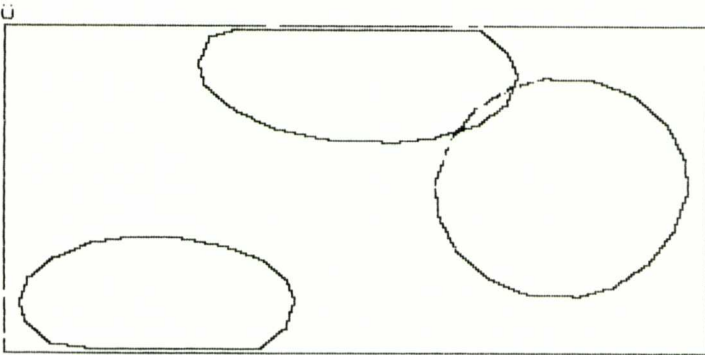
És una experiència numèrica basada sobre un algorisme d'aprenentatge que ha donat resultats en el cas de la descoberta de conjunts o conceptes dins espais representables per conjunts de punts geomètrics.

Tenim un algorisme d'aprenentatge capaç de detectar agrupaments de punts en un espai on d'una manera seqüencial van apareixent punts situats segons una llei aleatòria no homogènia. A 11a tenim el conjunt de 50 punts, i a 11b els grups de punts obtinguts i reconeguts per l'algorisme. Modifiquem la presentació del resultat introduint una llei de desplaçament dels punts que tenen tendència d'apropar-se en una certa proporció al centre del grup més proper existent al moment de llur aparició.



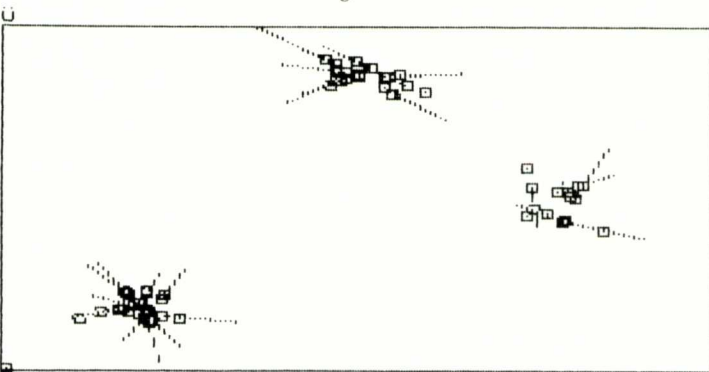
50 punts distribuïts en tres grups i aleatòriament.

Fig. 11a



3 grups reconeguts per l'algorisme de classificació iterativa amb autoaprenentatge

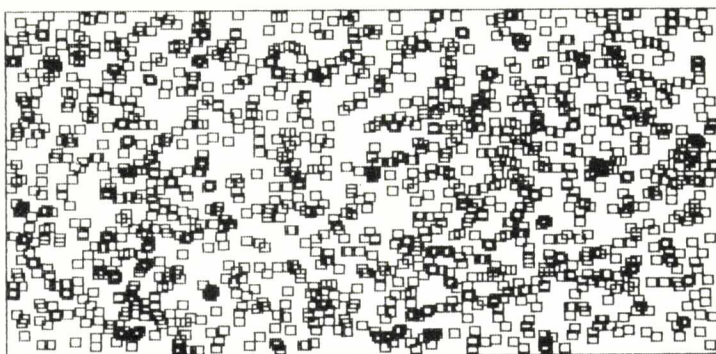
Fig. 11b



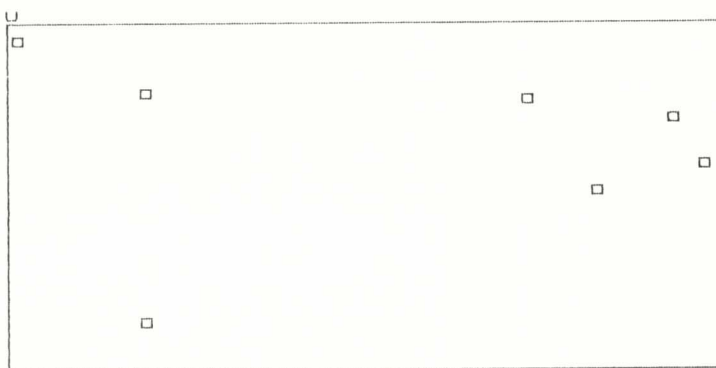
Agrupaments obtinguts afegint a l'algorisme de la figura 11b una llei de desplaçament cap al centre del grup més proper existent al moment d'aparició del punt.

Fig. 11c

Resultats d'aprenentatge significatiu.



1460 punts distribuïts d'una manera uniforme i aleatòria. Figura 12a.



Els 7 primers punts distribuïts d'una manera uniforme i aleatòria. Figura 12 b.



Grups constituïts aleatòriament pels 1460 punts de la figura 12a amb un algorisme d'agrupament cap al centre dels grups ja constituïts al moment d'aparició de cada punt.

Fig. 12
Resultat d'aprenentatge no significatiu.

A la figura 11c hom pot apreciar l'aparició de grups concentrats provinents de punts que s'han desplaçat per a formar-los. Ha estat representada amb una recta la trajectòria de cada punt des del lloc on aparagué fins a la seva situació final representada per un petit quadrat. Fins ací no hi ha cap paradoxa, ja que la situació real, figura 11a, posseeix una estructura ben clara d'inhomogeneïtat espacial.

A la figura 12a hem representat una altra situació on la font real no té cap estructura, és caòtica; la figura representa 1460 punts uniformement distribuïts a l'àrea considerada per a la simulació. Aquests punts han anat apareixent progressivament, i és evident que quan només eren uns pocs semblava que hi existís una estructura; és allò que hem volgut representar a 12b, on només hi ha 7 punts.

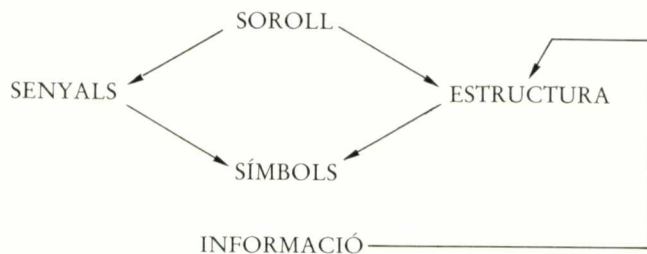
Aplicant el mateix algorisme que a la figura 11c, hom ha obtingut la imatge 12c, on podem apreciar grups ben constituïts que no reflecteixen cap estructura real. Si la història de la infància de l'algorisme hagués estat diferent, haurien estat obtinguts altres grups en altres llocs, però podem afirmar que sempre hauríem obtingut grups, és a dir estructura. És, doncs, aquest un exemple simulat del mecanisme de creació d'estructura, o informació a partir de soroll.

CONCLUSIÓ

Hem volgut en aquest article posar l'accent en un dels components del nostre món que anomenem **informació**, i que està estretament relacionada amb la intel·ligibilitat de la matèria, amb l'**estructura**. La **informació** és formada per **senyals**, **símbols**, i **soroll**, i tots tres components hi són igualment necessaris.

Les interaccions formals i filosòfiques d'aquests elements són molt complexes: **símbols** són **senyals** amb **estructura**, sense **símbols** no hi ha **informació**, sense **informació** no hi ha **estructura**. El **soroll** pot ésser a la base física de la generació de **senyals**, i fins i tot, com ho hem vist paradoxalment, creador espontani d'**estructura**.

Podem doncs meditar sobre l'esquema següent:



Sembla que la font primordial i original sigui el **soroll**, únic capaç d'inicialitzar el procés.

I tanmateix digué Albert EINSTEIN que Déu no juga als daus; només la fe ens ho podria assegurar.

BIBLIOGRAFIA

1. Teoria de la Informació

PIERCE, J. R. 1962. *Signals Symbols and Noise*. Ed. Hutchinson and Co. Londres.

SHANNON, C. E. and WEAVER, N. – 1949. *A Mathematical Theory of Communication* Urbana Univ. Pres.

2. Història

SUETONIUS TRANQUILLUS, C. – 0122. *Vida dels dotze cèsars*.

3. Lingüística, estructuralisme i filosofia

de SAUSSURE, F. – 1915. *Cours de Lingüistique Générale*. Ed. Payot – París 1973.

WITTGENSTEIN, L. – 1921. *Tractatus logico-philosophicus*. Traducció de J. M. Terricabras. Ed. LAIA. Barcelona 1981.

CHOMSKY, N. – 1957. *Structures Syntaxiques*. Traducció de M. Brandean. Ed. Seuil – París 1969.

BARTHES, R. – 1957. *Mithologies*. Ed. Seuil. París.

LEVI-STRAUSS, C. – 1949. *Structures élémentaires de la parenté* – París.

4. Matemàtiques de l'estructura

THOM, R. – 1965. *Propriétés différentielles locales des ensembles analytiques*. Séminaire Bourbaki 1964-1965, n° 281.

BRUTER, C. P. – 1974. *Topologie et Perception*. Ed. Maloine-Doin – París.

SALES, T. – 1985. Conferència en aquest seminari – Prada (Comunicació oral).

ATLAN, H. – 1979. *Entre le cristal et la fumée*. Ed. Seuil. París.

5. Matemàtiques Estocàstiques

STRATONOVICH, R. L. – 1960. Conditional Markov. Processes Theory. *Proba. Appl.* 5 - 156 - 178.

JAZWINSKI, A. H. – 1970. Stochastic processes and filtering theory. Ed. Academic Press. Nova York.