

DINAMICA COGNITIVA: COMPLESSITÀ E CREATIVITÀ.

UN ELOGIO DELL'OBLIO

E. TITO ARECCHI

Dipartimento di Fisica, Università di Firenze

Istituto Nazionale di Ottica Applicata

Sommario

Un problema scientifico formalizzato in un codice prefissato è traducibile in un corrispondente problema computazionale. Chiamiamo *complessità (algoritmica)* la lunghezza in bit dell'istruzione più corta che risolve il problema. Il *caos deterministico* caratterizza un generico sistema dinamico, rendendo il corrispondente problema pesante sia sperimentalmente sia computazionalmente, in quanto occorre ri-assegnare le condizioni iniziali a intervalli temporali più brevi del tempo di perdita dell'informazione, che è il reciproco della velocità di perdita dell'informazione, detta *entropia di Kolmogorov*. Si può controllare il caos aggiungendo nuovi gradi di libertà, cioè rimpiazzando l'informazione persa con quella portata dai nuovi gradi di libertà. Ciò equivale a un cambiamento di codice.

All'interno di un singolo codice, cambiare ipotesi vuol dire fissare gruppi diversi di parametri di controllo. A ciascun gruppo assegniamo una probabilità a priori, che va confrontata con l'esperienza e trasformata in una probabilità a posteriori, mediante il teorema di Bayes. L'applicazione sequenziale della regola di Bayes non è altro che la strategia darwiniana nell'evoluzione biologica. La sequenza è un algoritmo che sceglie la massima pendenza nello scalare un monte delle probabilità. L'esplorazione di ipotesi plausibili termina una volta raggiunta la massima probabilità.

Cambiando codice (cioè il pacchetto di variabili e mutue relazioni che descrivono il problema) si può ricominciare a formulare nuove classi di ipotesi. Chiamiamo *creatività* il cambiamento di codice, che è guidato da suggerimenti non formalizzati nel codice precedente, e pertanto non accessibili al computer.

Chiamiamo *complessità semantica* il numero di codici distinti o *modelli*, che descrivono in modo soddisfacente una situazione. Si tratta di un concetto sfumato per due motivi: 1) la nozione «soddisfacente» è convenzionale, legata a un certo contesto culturale; 2) se anche si raggiunge un accordo sul punto 1), il numero di cambiamenti varia a seconda dell'ambiente entro cui è inserito il sistema in esame (il sistema è aperto).

Queste considerazioni sono illustrate con riferimento a un compito cognitivo,

partendo dalla sincronizzazione dei neuroni in un'area percettiva e discutendo il passaggio dalla percezione alla cognizione.

1. Introduzione

In questo contributo definisco il concetto di *complessità* e ne illustro il ruolo in un programma scientifico. Il titolo può sembrare strano: qual è mai la connessione fra *complessità* e *creatività*? Cercherò di formulare risposte soddisfacenti alla seguente domanda: *come trasformiamo le percezioni in messaggi carichi di significato?*

A questo riguardo scopriremo delle ambiguità non eludibili: nella misura in cui ci poniamo domande di senso, ci accorgiamo che possiamo arrivare alla stessa risposta partendo da percezioni diverse; viceversa, possiamo attribuire significati differenti allo stesso oggetto percepito, senza variarne la descrizione fenomenica; c'è una specie di *complementarità* fra fenomenologia e significati. Nasce allora il problema: *come delimitiamo l'ambito dei significati per avere descrizioni fenomeniche obiettive, cioè invarianti per tutti?*

Parliamo allora di *visione scientifica* e della sua obiettività, come raggiunta rinunciando a domande di significato. L'ideale sarebbe un unico codice di lettura del mondo. Se tutto fosse riconducibile entro quest'unico codice, qualunque problema potrebbe essere risolto da un computer sufficientemente potente. D'ora in poi designeremo come *macchina di Turing* tale computer; i suoi algoritmi dovrebbero farci risolvere qualunque problema; la *complessità (algoritmica)* di un problema sarà allora definita come la lunghezza in bit dell'istruzione più corta che risolve il problema. Vedremo come le molteplici definizioni d'uso corrente siano varianti di questa definizione fondamentale. Chiameremo *complicazione* una misura di complessità relativa a un preciso codice descrittivo.

Qui però interviene un fatto cruciale, di cui abbiamo preso atto solo da alcuni decenni. I processi dinamici che implicano il *caos deterministico*, pur essendo descritti da un codice relativamente semplice (un numero piccolo di equazioni non lineari accoppiate), perdono in maniera non sanabile l'informazione sullo stato iniziale da cui erano partiti, e pertanto possono approdare a tutto un ventaglio di possibili scenari. L'incertezza che avevamo esorcizzato rinunciando ai significati riemerge. Rimpiazzando l'informazione persa con nuova informazione proveniente dalle osservazioni, può essere opportuno ridefinire nuove variabili e regole di mutuo accoppiamento, cioè *cambiare codice di descrizione*.

Nell'ambito di un codice unico, la *complessità-complicazione* cresce al crescere della velocità di perdita dell'informazione. Invece, il cambiamento di codice può portare a una nuova descrizione a bassa complessità: è questo l'aspetto saliente della creatività scientifica.

Quando un fenomeno è descrivibile da codici diversi, parleremo di *complessità semantica* distinta dalla complessità *algoritmica*. È opportuno anettere un valore, o significato, a un codice nella misura in cui esso sia associato a gradi diversi di

complicazione. Vediamo dunque che una pluralità di codici equivale a quella pluralità di significati che avevamo creduto di eliminare dal discorso scientifico.

Nel cambiare codice, cioè nel formulare un nuovo modello, dobbiamo introdurre nuove ipotesi. Partendo da un certo gruppo di ipotesi, a ciascuna delle quali assegniamo con argomenti di plausibilità un certo grado di probabilità a priori, il confronto con i dati sperimentali privilegerà una di queste ipotesi rispetto alle altre; avremo cioè probabilità a posteriori più aderenti alla situazione da descrivere. È questo il teorema di Bayes, che è la base dell'inferenza scientifica, e che si è creduto di poter automatizzare dotando un computer di un programma corrispondente e facendone così un *sistema esperto*, cioè un sistema che elaborando un pacchetto di dati sia in grado di elaborare una *diagnosi* su un certo fenomeno.

La probabilità a posteriori *sceglie* l'ipotesi più opportuna per costruire un modello procedurale. La ripetizione sequenziale di: *formulazione di ipotesi – confronto con l'esperienza – selezione dell'ipotesi con probabilità a posteriori massima* sembra assicurare una procedura scientifica ottimale. Tale è in definitiva la strategia darwiniana di mutazione e selezione del carattere che meglio si adatta all'ambiente: l'adattamento (*fitness*) rappresenta la scalata del monte delle probabilità, verso valori sempre maggiori.

Vedremo però che la strategia bayesiana-darwiniana funziona solo in uno scenario con un singolo picco di probabilità. In presenza di picchi multipli – come peraltro accade in situazioni sufficientemente «complesse» (e qui usiamo il termine in senso qualitativo) – occorre fuoriuscire dall'automatismo del teorema di Bayes e avere il coraggio e la fantasia di inventare ipotesi non immediatamente in linea con la situazione finora osservata e spiegata, cioè apprestarsi a scalare un monte di probabilità diverso da quello fino ad allora seguito. Questa cesura con la procedura precedente rappresenta un cambiamento di codice descrittivo, e – se confortata dai dati sperimentali – è la base per la formulazione di una nuova teoria. Abbiamo dunque indicato dei salti discontinui non delegabili al computer: essi sono la base della *creatività scientifica* e rappresentano una scelta libera, non determinata dalla situazione precedente. Si tratta di un atto squisitamente umano, che non ha un equivalente nelle procedure di una macchina. Avendo dovuto introdurre uno scenario con molti picchi di probabilità, sembrerebbe naturale associare il termine *informazione* (che nella definizione di Shannon è una misura di probabilità) ai vari punti che si inerpicano su un picco particolare, e invece di chiamare *significato* il paesaggio complessivo con una molteplicità di picchi. Allora la *complessità semantica* può essere misurata dal numero di picchi distinti cui corrispondono descrizioni scientifiche diverse.

Finora ci siamo mossi su un terreno puramente epistemologico, descrivendo procedimenti cognitivi che ci permettono di affrontare situazioni disparate. Che valore ontologico dobbiamo attribuire al paesaggio a molti picchi di probabilità? Se l'immagine rappresentasse il mondo come esso è, indipendentemente da noi, avremmo allora trovato uno strumento per descrivere tutte le possibili teorie e quantificare la complessità semantica. Vediamo però subito che non è così. Noi siamo simultaneamente spettatori

ed attori; e le nostre azioni conseguenti alle nostre valutazioni cognitive modificano lo spazio delle probabilità prima vigente, in modo tale da dover considerare il nostro modello come descrittivo di un passato che ormai non esiste più.

Consideriamo ad esempio uno studioso di economia, che formula descrizioni dinamiche dei mercati finanziari. Se lo studioso ha elaborato una descrizione soddisfacente e ne informa altri agenti o se ne vuole servire per giocare in borsa, immediatamente il suo intervento causa reazioni multiple da parte degli altri agenti che alterano il paesaggio di probabilità.

Si potrebbe obiettare che in molte situazioni l'atto di misura dell'osservatore non altera sostanzialmente lo stato di fatto: se uno osserva le stelle non ne altera il moto. Ciò non è più vero nel mondo quantistico, dove, per effetto dell'*entanglement*, una perturbazione localizzata altera globalmente tutto il sistema. Effetti visibili di tal fatta sono oggetto di sperimentazione fra chi sta cercando di realizzare computer quantistici; ma tutto ciò va al di là di quanto si prefigge questo lavoro (non definisco con precisione l'*entanglement*: basti dire che è un legame inestricabile fra situazioni apparentemente in mutuo conflitto, come ad esempio gli stati di gatto vivo e di gatto morto nel famoso esempio di Schrödinger).

2. La visione scientifica

Con la scienza moderna, si tenta di ripulire ogni resoconto sul mondo da quelle ambiguità che richiedono una interpretazione; evitando quest'ultima, si fuoriesce dal soggettivo e si approda a descrizioni valide per tutti.

La complessità nasce dal tentativo di una costruzione logica del mondo a partire dalle «*affezioni quantitative*» in cui frantumiamo un oggetto di esperienza, piuttosto che «*tentarne l'essenza*» (G. Galilei, Lettera a Marco Welser, 1610). Supponiamo di scomporre una mela nelle sue *affezioni quantitative*, cioè nelle proprietà distinte che possiamo misurare: sapore, odore, colore, forma etc. Ad ognuno di questi *punti di vista* corrisponde un apparato di misura che fornisce un numero. Se abbiamo scelto un pacchetto adeguato di indicatori, la collezione ordinata dei numeri corrispondenti a ciascuna misura dovrebbe caratterizzare la mela.

L'approccio è motivato dalla convinzione che si possa raggiungere una descrizione completa di un oggetto di esperienza attraverso i suoi «*atomi di significato*» misurabili con apparati di misura che forniscono numeri ripetibili e indipendenti dagli stati dell'osservatore (esempio della mela: odore, colore, sapore, etc., oppure, più radicalmente, gli atomi di cui essa è composta). Ogni oggetto diventa così un elemento definito di un insieme di Cantor, espresso da una collezione ordinata di numeri utilizzabili come dati di ingresso in una teoria matematica che permetta di prevedere il futuro

Come vedremo, la procedura appena delineata può bastare per il rifornimento del magazzino di un mercato, però non esaurisce la realtà della mela. In effetti l'*intero* (la mela) precede le misure separate, che forniscono quantità derivate, non primitive: si appoggiano alla mela ma una volta enucleate (tecnicamente: *astratte*) non permettono

la ricostruzione logica della mela: non siamo mai sicuri che nel futuro non compaia un ulteriore *punto di vista* inedito che sia misurabile e che arricchisca la descrizione della mela.

Prendiamo l'esempio del genoma umano. Sono stati individuati 30.000 geni, cioè tratti di sequenze di basi che codificano proteine; ma unendo tutti i tratti si arriva all'1% della lunghezza totale del genoma; il 99% – non codificante – espleta funzioni che stiamo appena cominciando a chiarire. L'insorgere di correlazioni tra tratti diversi (per esempio un gene la cui espressione dipende dalle proteine codificate da altri geni) rende *complesso* il problema nel senso che non ne abbiamo esaurito il significato con il semplice elenco dei geni.

Gli esempi più convincenti vengono dalla linguistica. Prima nasce il linguaggio articolato, la poesia, le favole, e poi da queste si estrae un lessico. L'operazione inversa, partire dal vocabolario e programmare al computer un testo letterario, è fallimentare. In questo ordine di complessità rientra la traduzione da una lingua all'altra. Esistono macchinette per cavarsela alla stazione o al ristorante in Cina; ma per un testo elaborato, per esempio Omero, il traduttore non è automatico: piuttosto è un bilingue che vive le situazioni di Ulisse in Greco e le racconta in Italiano. Questo ruolo del traduttore è discusso in un bel libro: *Il grande codice (la Bibbia e la letteratura)* di Northrop Frye (Einaudi, Torino 1986); l'autore si pone il problema: come mai la Bibbia scritta in Ebraico (lessico di 5.000 vocaboli) può essere tradotta in Inglese (il cui lessico comprende 500.000 vocaboli)? La risposta è che una procedura come la traduzione letteraria non può essere condensata in un elenco di operazioni di macchina, e perciò non può essere automatizzata.

Siamo dunque in presenza di due accezioni di complessità: una, scientifica, legata alla difficoltà di fornire una descrizione soddisfacente, un'altra (che è di natura *semantica*), legata alla possibilità di attribuire significati diversi allo stesso oggetto di esperienza, a seconda del punto di vista da cui ci si colloca. Quel che disturba è l'apparente relativismo della seconda. Peraltro, la complessità semantica sembra essere più fondamentale; essa rappresenta il limite della presunzione ideologica della fisica dei componenti elementari di essere l'unica descrizione rilevante del reale: il cosiddetto *riduzionismo*.

Ci proponiamo di definire rigorosamente la complessità scientifica e di chiarire le procedure che permettano un consenso sulla selezione dei significati.

3. Complessità e caos

Nella risoluzione al computer di un problema, la complessità rappresenta il costo del calcolo; essa pertanto presenta due aspetti:

- *spaziale*: quantità di risorse di macchina da impegnare, una volta convenuto di riferirsi al programma più corto ($CA = \text{complessità algoritmica}$ di Chaitin, 1965);
- *temporale*: tempo necessario all'esecuzione, una volta scelto il programma più corto ($D = \text{depth (spessore) logico}$ di Bennett, 1985).

Esploriamo questo secondo aspetto. Sia t_1 l'unità di tempo necessaria per descrivere un singolo elemento, e si abbiano N elementi. Se le loro proprietà si sommano (come i fagioli in un sacco) cioè se l'oggetto composto di N elementi è cresciuto per *giustapposizione*, allora il tempo totale necessario T è:

$$T = N t_1$$

Se invece la crescita è per «*intussuscezione*» (uso un vecchio termine della filosofia della natura che vuol dire che gli elementi non si limitano a giustapporsi, ma si accoppiano fra di loro con modalità non prevedibili a partire da un'ispezione del singolo elemento; es. tipico: una collezione di vocaboli di un lessico, disposti convenzionalmente in ordine alfabetico nel dizionario, rispetto a come invece si organizzano nel verso di un poeta) allora il tempo totale non è la semplice somma dei tempi individuali, e addirittura l'estensione (*size*) N del sistema composto può figurare ad esponente:

$$T' = 10^N t_1$$

Vediamo alcuni esempi di T e T' . Sia $t_1 = 1$ microsec = 10^{-6} sec. Per $N = 10$, si ha: $T = 10 t_1 = 10^{-5}$ sec; invece, $T' = 10^{10} t_1 = 10^4$ sec, che è ancora praticabile (si tratta di meno di tre ore). Per $N = 100$, si ha $T = 100 t_1 = 10^{-4}$ sec; invece, $T' = 10^{100} t_1 = 10^{94}$ sec (molto più lungo dell'età dell'universo che è circa 10^7 sec/anno $\times 10^{10}$ anni = 10^{17} sec).

La complessità, scegliendo una delle due definizioni viste, è quantificabile, ma nell'ambito di un problema di calcolo, per ridurci al quale abbiamo ridotto il mondo a un solo codice di lettura: pertanto la diremo piuttosto *complicazione*, in quanto rappresenta l'impegno di una macchina che lavori all'interno di un codice descrittivo prefissato. Nell'ambito di una formulazione dinamica, genericamente è presente caos deterministico. Se prepariamo il sistema in uno stato iniziale, l'informazione corrispondente si perde nel corso del tempo. L'informazione persa per unità di tempo è un *invariante* dinamico, cioè una caratteristica intrinseca non legata a un cambiamento di rappresentazione; essa è detta *entropia metrica*, o *entropia di Kolmogorov K*.

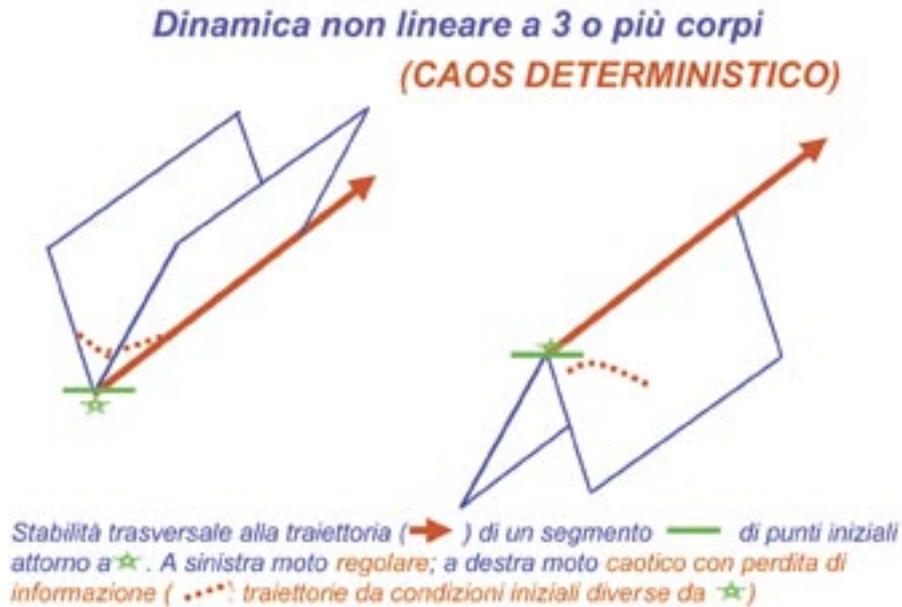


Figura 1 - Caos deterministico. La traiettoria (linea spessa con freccia) calcolata con una certa legge dinamica è unica per una condizione iniziale assegnata, qui indicata da una stella. Le coordinate del punto iniziale sono in genere numeri reali che tronchiamo a un numero finito di cifre; pertanto la condizione iniziale non è un punto euclideo, ma un segmento. Condizioni iniziali a destra o sinistra di quella ideale convergono verso la traiettoria calcolata o ne divergono a seconda che la stabilità trasversale sia di tipo valle (a sinistra) o colle (a destra). Poincaré fece vedere che a partire dal problema a 3 corpi (ricordiamo che Newton aveva risolto quello a 2 corpi!) può accadere genericamente la situazione di destra, che oggi chiamiamo caos deterministico, con perdita di informazione sulla scelta iniziale che sarà più o meno rapida a seconda della ripidità del colle.

K risponde alla domanda: se osserviamo il sistema per un tempo così lungo che l'informazione per unità di tempo diventi invariante, che informazione abbiamo sul futuro? Orbene, la risposta è:

1. $K = 0$ per un sistema *ordinato*; infatti, in un sistema deterministico, una volta esauriti i transitori, non si ha un extra d'informazione per un extra di osservazione;
2. $K \rightarrow \infty$ per un sistema *disordinato*; infatti, per quanto a lungo si osservi un moto browniano, non si sa mai dove il sistema andrà a un istante successivo;
3. K è *finito* per un sistema *caotico*; l'informazione si consuma gradualmente.



Figura 2 - Controllo del caos. Il caos può essere controllato dall'esterno, aggiungendo altre azioni dinamiche che – senza perturbare la traiettoria longitudinale – alterino la pendenza del colle. Nel caso percettivo, è plausibile pensare che i segnali top-down più opportuni siano quelli che assicurano una vita lunga alla dinamica collettiva dei neuroni. La selezione fra le possibili percezioni è perciò legata al loro tempo di conservazione dell'informazione.

In Fig.3 mostriamo come la CA di Chaitin o il D di Bennett variano in funzione di K . Nel primo caso è ovvio che per K tendente all'infinito (come per numeri *random*) non esiste un programma corto di calcolo, in quanto non c'è una guida per comprimere l'istruzione, ma occorre inserire in essa tutte le cifre, una dopo l'altra. Invece la durata di un programma che genera un numero *random* è cortissima: una volta scritta la (lunguissima, come appena detto!) istruzione basta comandare: *stampa!* (figura a destra). CA cresce linearmente con K : calcolare un numero *random* è costoso perché va dato cifra per cifra e non c'è una regola per prevedere le cifre successive; invece D raggiunge un massimo e poi torna a valori piccoli per numeri *random*: per questi, una volta scritto il programma, il tempo di esecuzione è minimo, in quanto non c'è una regola di evoluzione.

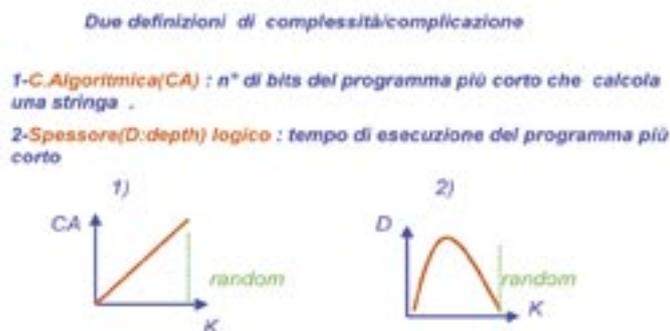


Figura 3 - Due definizioni di complicazione. 1) CA, complessità algoritmica (Chaitin, 1965) = numero di bits del programma più corto; 2) D (depth), spessore logico (Bennett, 1986) = tempo di esecuzione del programma più corto.

4. La complessità come complicazione, all'interno di un codice

Consideriamo situazioni fisiche che coinvolgono grandi folle di oggetti identici: gas in una cella, magneti elementari (spin).

Nel lavoro sull'*opalescenza critica* (1910) Einstein tiene conto delle interazioni fra le molecole di un fluido, non più viste come scorrelate. In condizioni normali, queste interazioni sono mascherate dall'agitazione termica, e le molecole sembrano indipendenti una dall'altra, come nel moto browniano. Se inviamo fasci di luce su una cella riempita con tali molecole (fig.4 a sinistra) la luce si inserisce fra una molecola e l'altra, e la cella è trasparente. Ma se ora aggiustiamo con accuratezza la temperatura vicino al valore critico, per cui l'agitazione termica uguaglia l'interazione molecolare, allora ogni molecola sente l'interazione dei vicini e si ritrova virtualmente distribuita su un volume molto maggiore di quello che essa occupava inizialmente; avremo, cioè, correlazioni a lungo raggio e la luce sarà riflessa da questi contorni del dominio di pertinenza di ciascuna molecola e perciò sarà sparpagliata un po' in tutte le direzioni, come farebbe una nebbia densa: chiamiamo questo fenomeno *opalescenza critica*. Se però abbassiamo ulteriormente la temperatura, il gas condensa in un liquido ad alta densità, le molecole saranno impaccate ciascuna attorno alla propria posizione e la cella sarà di nuovo trasparente. Dunque l'opalescenza critica caratterizza l'insorgere di notevoli interazioni mutue per cui non possiamo considerare più ogni molecola per conto proprio.

In modo concettualmente analogo, nell'emissione di luce i fotoni generati sono indipendenti se prevale l'emissione spontanea, ma per effetto dell'emissione stimolata si correlano fra loro: è questa la base concettuale del *laser* inventato qualche decennio dopo.

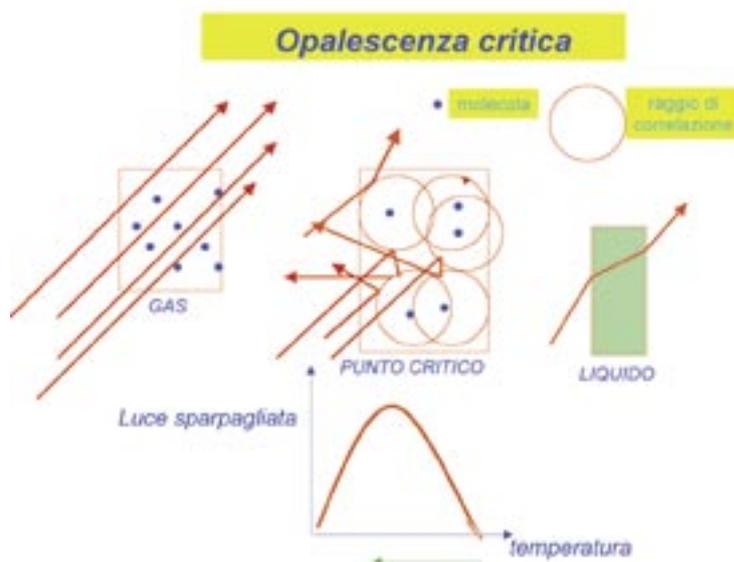


Figura 4 - Opalescenza critica. Ad alta temperatura le molecole del gas sono quasi indipendenti una dall'altra; se inviamo luce, questa intrude fra una molecola e l'altra, e la cella è trasparente (a sinistra).

Portando la temperatura vicino al valore critico per cui l'agitazione termica uguaglia l'interazione molecolare, ogni molecola si ritrova virtualmente distribuita su un volume molto maggiore di quello che essa occupava inizialmente: avremo cioè cospicui raggi di correlazione (cerchi nella figura centrale); la luce sarà riflessa dai contorni di questi volumi e perciò sarà sparpagliata in tutte le direzioni, generando opalescenza critica. Se abbassiamo ulteriormente la temperatura, il gas condensa in un liquido ad alta densità: il fluido si comporterà otticamente come un vetro e la cella sarà di nuovo trasparente.

L'opalescenza critica caratterizza l'insorgere di notevoli interazioni mutue per cui non possiamo considerare più ogni molecola per conto proprio. Così come D , l'intensità della luce sparpagliata passa per un massimo fra fase disordinata (gas = alto K) e fase ordinata (liquido = basso K).

Come secondo esempio, consideriamo il modello di Ising (1925). Esso consiste in una collezione di aghi magnetici elementari, o *spin*, ciascuno vincolato ad assumere una fra due sole posizioni spaziali, o col Nord in alto (si sta parlando dei poli Nord e Sud di una calamita) o col Sud in alto. Spazialmente avremo una collezione di aghi magnetici tutti verticali e con la freccia (che indica il polo Nord) puntata in su o giù.

Per effetto delle mutue interazioni, gli spin tendono a disporsi tutti paralleli in una configurazione ordinata, per esempio tutti in su (fig. 14 a sinistra, colorazione scura). Ciò accade solo a basse temperature; se si aumenta la temperatura, l'agitazione termica di ciascuno spin contrasta con l'interazione mutua spin-spin, finché si arriva a una situazione del tutto disordinata, in cui è impossibile trovare alcuna struttura (*su* e *giù* rimescolati finemente, figura a destra). Nel mezzo, c'è una temperatura critica T_c a cui le due azioni si compensano e si formano domini estesi di spin in su (colore scuro) e in giù (colore chiaro), con forti fluttuazioni nella loro estensione, dovute alla competizione fra ordine (mutua interazione) e disordine (agitazione termica). Queste fluttuazioni rendono onerosa la descrizione in quanto, ripetendo l'osservazione, non

si registra la stessa cosa. Se chiamiamo L il numero totale di spin, cioè l'estensione del sistema, D è basso per basse o alte temperature, mentre cresce con una potenza frazionaria di L per $T=T_c$.

$D=$ Depth of Ising Model

Lontano dal punto critico, lo spessore D è indipendente dall'estensione L del sistema (algoritmo di Metropolis)

Vicino a T_c i migliori algoritmi sono i cluster algorithm, per cui D cresce come L elevato alla potenza detta "esponente dinamico" (per il modello di Ising tridimensionale, è circa $z = 2$)

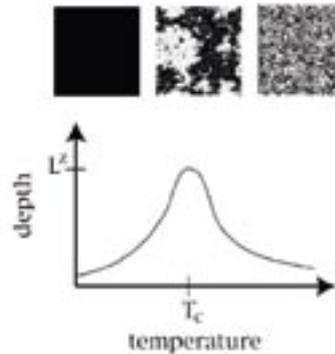


Figura 5 - Lo spessore logico del modello di Ising. Se chiamiamo L il numero totale di spin, cioè l'estensione del sistema, D è basso per basse o alte temperature, mentre cresce con una potenza frazionaria di L per $T = T_c$.

5. Riduzione di complessità per cambiamento di codice

La creatività scientifica è legata a un cambiamento di codice rispetto al programma riduzionistico, di assegnare posizione e velocità di ogni componente elementare. Si tratta di introdurre nuove variabili legate da nuove leggi. Ipotizzando nuove variabili che controllano il caos, riducendo K si introduce una teoria efficace che trascura le variabili veloci e descrive solo le variabili lente. In ciò si è guidati dalla nostra interpretazione dei segni (*semiosi*) ben al di là delle poche *affezioni quantitative* isolate dagli apparati di misura.

Uno dei padri dell'Intelligenza Artificiale (IA), Herbert Simon, aveva ideato un programma di computer, chiamato *BACON*, che, alimentato con i dati astronomici noti alla fine del 1500, estraeva le tre leggi di Keplero. Il successo sembrava fare dell'IA la spiegazione dei processi mentali. In effetti, *BACON* funziona solo perché si è nell'ambito di un codice unico. Appena $K > 0$, cioè si perde l'informazione dei dati iniziali entro un tempo $1/K$, occorre rimpiazzare l'informazione mancante con dati ulteriori. Ci sono due modi alternativi per ovviare.

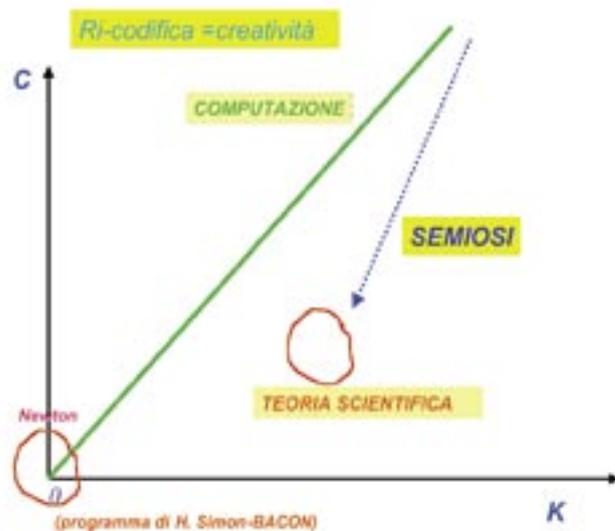


Figura 6 - Diagramma C-K. Confronto fra i modi di procedere di un computer e di un essere umano. Restando dentro un unico codice, come fa il computer, C cresce con K . Invece lo scienziato è opportunisto, e cerca il codice migliore per ridurre C . Ipotizzando nuove variabili che controllano il caos riducendo K , si introduce una teoria efficace che trascura le variabili veloci e descrive solo le variabili lente.

1. *Approccio riduzionistico*: si mantiene il piacere estetico delle leggi semplici (Keplero, Newton), ma trattando di $N > 2$ oggetti, si aumenta C proporzionalmente a K , cioè occorre aumentare massicciamente il numero di cifre con cui si assegnano le condizioni iniziali (aumentando sia la risoluzione dell'apparato di misura, sia la memoria che immagazzina i dati), per ovviare alla erosione di informazione dovuta a $K > 0$, oppure occorre ri-aggiornare le condizioni iniziali con nuove misure, come si fa con le previsioni meteorologiche, con i modelli finanziari, ecc. 2. *Teoria efficace*: si utilizza la perdita di informazione per introdurre nuove variabili legate da nuove leggi, per contenere la crescita di C , cioè si formula una nuova teoria, come è avvenuto per la *termodinamica*.

Se associamo la creatività a questa seconda procedura, allora dobbiamo prendere atto che essa è figlia di $K > 0$; questo spiega il sotto-titolo di questo lavoro: *un elogio dell'oblio*.¹

All'estremo di alti C e K , la dinamica delle singole molecole di un gas in equilibrio (tempi caratteristici: picosecondi) è rimpiazzata da una descrizione a poche variabili collettive, cioè dalla termodinamica (scale di tempi rilevanti: dal microsecondo in su).

Dal punto di vista della pratica scientifica, diciamo «complesse» le situazioni per cui non abbiamo ancora a disposizione una descrizione soddisfacente, esempio tipico: la *mente umana*. L'*economia* implica l'interazione fra un numero grande di agenti umani, quindi dovrebbe essere ancor più complessa, ma i modelli economici correnti sono più soddisfacenti dei correnti modelli cognitivi.

6. Il Teorema di Bayes - informazione e significato

Thomas Bayes amava giocare d'azzardo; pertanto elaborò una procedura per fare buon uso delle esperienze pregresse. In formula, il Teorema di Bayes è così espresso²:

$$\text{prob.}(ipotesi | \text{dato}) = \text{prob.}(\text{dato} | ipotesi) \times \text{prob.}(ipotesi) / \text{prob.}(\text{dato})$$

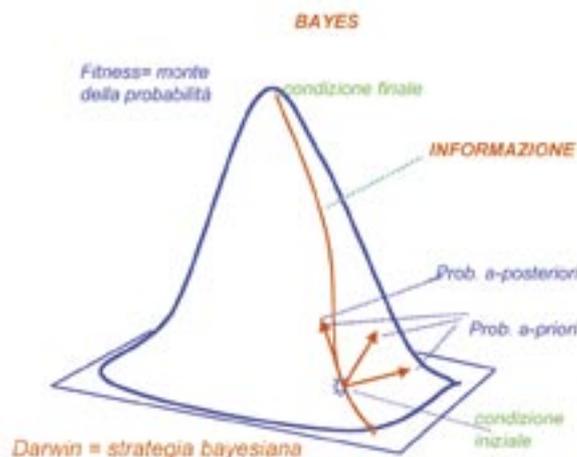


Figura 7 - Teorema di Bayes. Costruzione di modelli scientifici vieppiù adeguati agli esperimenti, per applicazione successiva del Teorema di Bayes. La procedura è una scalata del monte della probabilità. Ad ogni punto sulla curva che porta al picco, il contenuto di informazione dipende dal valore locale di probabilità. Si noti che la strategia darwiniana di selezionare la massima fitness dopo una mutazione è una implementazione sequenziale del teorema di Bayes.

Spieghiamone il significato. Il Teorema dice che la probabilità di una ipotesi, condizionata da (è questo il significato della barra |) un dato osservato, e che chiameremo probabilità *a posteriori*, è data dal prodotto della probabilità, che chiameremo *il modello*, che si verifichi il dato a partire da (condizionato da) una certa ipotesi per la probabilità dell'ipotesi (probabilità *a priori*: ci siamo inventati un pacchetto di possibili ipotesi a ciascuna delle quali annettiamo una probabilità a priori), diviso la probabilità di accadimento del dato sperimentalmente verificato (dobbiamo ripetere la misura su tutta una classe di dati per riuscire ad assegnare quest'ultima). Come mostrato in fig.7, se partiamo dalla condizione iniziale e formuliamo un ventaglio di ipotesi, il verificarsi di un certo dato ci farà scegliere quella ipotesi che massimizza l'espressione su riportata. Successive applicazioni del teorema equivalgono a una ascesa di un monte delle probabilità secondo la direzione di massima pendenza, finché si raggiunge il massimo.

La strategia evolutiva postulata da Darwin, come sequenze di *mutazione* e *selezione* del mutante che meglio si adatta all'ambiente (cioè, che ottimizza la *fitness*), è una applicazione del Teorema di Bayes, chiamando *fitness* il monte di probabilità da scalare. Questa procedura può essere formalizzata in un programma. Si può realizzare un *sistema esperto*, che elabori dati sperimentali estraendone una diagnosi; sarebbe utile in medicina, gestione aziendale, ecc.

La procedura è a-semiotica, cioè è agnostica rispetto ai significati; alcuni cultori di IA ne hanno fatto un paradigma del modo di operare della nostra mente. C'è però un

limite di fondo che blocca l'automatismo. Una volta raggiunto il picco, il programma si ferma; qualunque ulteriore insistenza sarebbe una catastrofica caduta in basso. Forse è quel che accade nelle estinzioni di massa di specie biologiche, senza bisogno di invocare il meteorite che estingue i dinosauri.

Prendere atto che esistono altri monti (fig.8) e che si può ricominciare la scalata altrove è un atto di creatività, corrispondente a una comprensione dei segni del mondo (semiosi) guidata da tutto il retroterra culturale e umano dello scienziato: operazione non delegabile a un computer. Quando si siano esaurite le possibilità legate alla scalata di un picco, il sistema esperto non sa come continuare, invece l'agente umano può intraprendere un'altra scalata, facendo un salto discontinuo rispetto alla procedura graduale prevista dalla formula di Bayes.

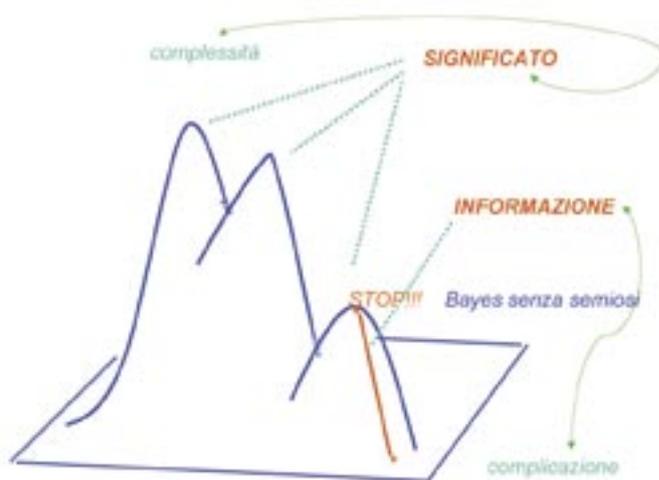


Figura 8 - Complessità semantica. Spazio delle probabilità con più massimi. L'ascesa verso un picco può essere automatizzata da un programma di scalata secondo il massimo gradiente. Una volta raggiunto il picco, il programma si ferma. Al significato corrisponde l'esistenza di più picchi.

Chiameremo *significato* il fatto che esistano più picchi. Esso va ben al di là dell'*informazione*, che nella definizione di Shannon quantifica l'evoluzione della probabilità nella salita su un versante. Possiamo identificare la complessità semantica con il numero di picchi, cioè di strategie di Bayes distinte che possiamo intraprendere; si tratta di un concetto non ben definito, in quanto il paesaggio a molti picchi varia man mano che varia il nostro grado di comprensione. La fig. 8 è puramente indicativa: in genere, operiamo in uno spazio con molte più di 2 dimensioni, e abbiamo molti più di 3 picchi. Inoltre, la nostra interazione intenzionale col mondo entro cui siamo immersi altera il numero di picchi.

7. Da complicazione a complessità

Abbiamo visto che in un universo chiuso, dove tutti gli oggetti sono già assegnati, non si ha complessità ma complicazione, nel senso che: (1) ogni problema è puramente sintattico, cioè implica la scelta e connessione di simboli già fissati, (2) le risorse necessarie

per risolvere il problema possono crescere in modo patologico con il numero di dati del problema. Se invece siamo aperti all'irruzione di elementi nuovi, non preventivati nel costruire un repertorio cognitivo, non possiamo ovviamente pretendere da demiurghi di conoscere il nostro mondo e l'altro (che d'ora in poi chiameremo *l'ambiente*), ma dobbiamo limitarci a registrare quali modifiche l'interazione con l'ambiente induca nel nostro mondo.

Il primo approccio, che possiamo chiamare statistico, consiste nel ritenere l'ambiente come un disturbo (*noise*) su cui siamo ignoranti; esprimeremo questa nostra ignoranza prendendolo a massima entropia, cioè all'equilibrio termico, e caratterizzandolo con una temperatura. All'interno del nostro sistema (quello su cui possiamo fare misure) l'accoppiamento con l'ambiente a una certa temperatura (che diremo bagno termico) provoca due conseguenze: *dissipazione*, cioè uno smorzamento della velocità iniziale; e *fluttuazione*, cioè oscillazioni permanenti a media nulla attorno al punto raggiunto di velocità nulla. Su questa falsariga si è costituita una strategia di Intelligenza Artificiale che possiamo chiamare *oracolo* (*oracle machine*). La macchina di Turing deterministica svolge solo operazioni sintattiche all'interno di un codice. Essa pertanto è limitata dal Teorema di Gödel che pone un limite alla procedura deduttiva a partire da un corpo di assiomi; la versione per macchina del Teorema di Gödel è³ che un computer equipaggiato con un programma definito non sa quando fermarsi per un generico pacchetto di ingresso. Supponiamo di aggiungere a una macchina di Turing un *disturbo* sufficiente a «saltare» su un programma diverso, cioè a cambiare «alfabeto» (tale è in effetti il passare dall'analisi di un sistema dinamico a N dimensioni a un sistema a N'=N+m dimensioni): In N+m potrebbe non incontrarsi la difficoltà che c'era in N. ciò permette di fuoriuscire dal limite di Gödel-Turing. Questa strategia di «immersione» di un problema in spazi a dimensioni differenti è la base dei calcoli «euristici» che guadagnano in ambito di applicabilità a scapito della accuratezza, che vien garantita solo entro un margine di errore; questo trucco di sporcare un calcolo deterministico con un disturbo permette di esplorare situazioni altrimenti non accessibili (le cosiddette *macchine di Boltzmann*).

La TM (macchina di Turing) è individuata da: $M = \langle \Sigma, K, \delta \rangle$ dove Σ è l'alfabeto di simboli, K l'insieme di stati, δ la funzione di transizione., la indeterministic TM «ha più di uno stato» ma sempre in numero finito e che la TM «probabilistica» permette «scelte random in un insieme finito di alternative»⁴. Invece l'aggiunta di m ulteriori dimensioni alla dinamica significa nuove equazioni del moto: questo porta la TM in una situazione nuova

$$M = \langle \Sigma^*, K^*, \delta^* \rangle ,$$

la costringe cioè a un'altra sintassi. Ogni sintassi sarà di un numero finito di Σ, K, δ , ma i salti di sintassi NON fanno parte del gioco.

Il secondo approccio consiste nel riconoscere la struttura ricca dell'ambiente e nel porci il problema semantico: possiamo dire cose significative del mondo esterno

attraverso le nostre acquisizioni percettive e cognitive? Cioè, possiamo fare salti «guidati» invece che casuali, come nella euristica su accennata? È questo il problema cruciale della creatività scientifica che abbiamo simbolizzato in fig.8. Orbene, la risposta è in fig. 2: per controllare il caos, dobbiamo aggiungere ulteriori gradi di libertà (le m dimensioni di cui sopra) ed m non è definito, perché un agente «semiotico» può decidere di compensare la perdita di informazione caotica in modo più o meno soddisfacente.

Diremo –come è oggi affermato dai cosiddetti «biosemiotici» – che qualunque vivente è semiotico perché fa operazioni di tipo fig. 2 (si noti, non semplicemente «omeostatiche» come nella fisiologia di un tempo, ma - lasciatemi introdurre un nuovo termine- «omeodinamiche») cioè si costruisce un suo mondo a prospettive stabili e, per far questo, deve alterare i gradi di libertà del sistema «se-stesso + ambiente», a differenza di un non-vivente, per cui si può fissare a-priori lo spazio dinamico entro cui evolve.

Elenchiamo qui di seguito quattro esempi di come «scienza creativa» voglia dire «introduzione di un nuovo codice» (*NON c'è programma di tipo BACON che possa far questo!*) e tre sfide aperte per la ricerca di questi anni.

Da complicazione a complessità: quattro teorie recenti

1. elettricità – magnetismo – ottica	equazioni elettromagnetismo (Maxwell)
2. tavola di Mendeleev	atomo quantistico (Bohr, Pauli)
3. zoo di 100 particelle elementari	quarks (M. Gell Mann)
4. leggi di scala in transizioni di fase	gruppo di ri-normalizzazione (K. Wilson)

Da complicazione a complessità: tre sfide aperte

1. leggi di potenza in fratture, mercati finanziari, DNA, reti metaboliche	?????? (vita, società)
2. incompatibilità fra QFT e GR (QFT=teorie quantistiche dei campi; GR = relatività generale)	?????? (cosmologia)
3. dinamica neuroni, glia, chimica neurotrasmettitori	?????? (mente, coscienza)

8. L'interpretazione dei segni – circolo e spirale ermeneutica

Nella dinamica cognitiva, il *segno* è qualsiasi cosa possa stimolare il controllo del caos di fig.2, portando un agente a una percezione o cognizione coerente (nel caso percettivo l'input è sensorio, nel caso cognitivo l'input è concettuale: una frase linguistica, il

risultato di una misurazione, ecc). I filosofi, a partire dall'antichità, hanno elaborato dottrine dei segni in modo più o meno implicito. Nell'epoca moderna, a partire dal sec.XVII con il *Tractatus de signis* di Giovanni da San Tommaso (Poincot) e a cavallo fra XIX e XX sec, con Peirce, il problema del segno è stato messo a fuoco. Per Peirce, affinché il segno serva deve entrare in relazione con un *oggetto*, e produrre nella mente del soggetto un *interpretante*: influenza triadica NON risolvibile in azione tra coppie. Questo processo interpretativo viene chiamato «semiosi».

Punto cruciale per noi: uno schema adattivo, con *feedback*, quale quello adombrato in fig.2, attinge a un repertorio pre-costituito oppure si arricchisce di novità? Risposta: lo spazio semantico non è precostituito ma si rinnova ad ogni esperienza. Il gioco dinamico fra *bottom-up* e *top-down* nel cervello fa sì che a pari stimolo esterno non corrisponda la stessa sequenza di *spikes* (cioè lo stesso treno di impulsi neurali); la sequenza pertanto non è come il codice a barre dei prodotti al negozio, dato *a priori*; e non ci sarà un futuro «grande inquisitore», cioè un *chip* impiantato nel cervello che possa leggere nel nostro privato e decodificare i nostri pensieri.

Illustriamo con due esempi recenti dalla neuroscienza. Inserendo elettrodi nei neuroni olfattivi di una locusta (G. Laurent, Caltech), si è verificato che la sequenza temporale di *spikes* che codifica un certo odore non varia in successive presentazioni dello stesso oggetto odoroso all'animale: questo sembra avere un repertorio limitato, che non evolve. Invece nel caso di un coniglio (W. Freeman, Berkeley) la successiva esposizione dell'animale allo stesso odore stimola distribuzioni di *spikes* differenti. Il coniglio si dirige in ogni caso verso l'oggetto odoroso (si tratta di una banana), ma la seconda volta aggiunge un elemento cognitivo in più: come a dire che si rende conto che si tratta di qualcosa che ha già sperimentato! Cartesio considerava tutti gli animali non umani come solo *res extensa*; forse, se avesse saputo di questi due esperimenti, avrebbe avuto qualche dubbio sul coniglio.

Se rappresentiamo i vari significati come punti di uno spazio semantico (fig.10), per la locusta – come per un robot – parleremo di *circolo ermeneutico*, cioè di connessione (linguistica o percettiva) fra elementi pre-assegnati e invariabili a cui si ritorna ineluttabilmente, per ricco che sia il repertorio accumulato. Invece per un coniglio (e a fortiori per un essere umano) parleremo di *spirale ermeneutica*, cioè di crescita cognitiva con l'esperienza. Si tratta di un programma *non finitistico*, in quanto non siamo confinati entro un insieme a numero di elementi finito: il ripresentarsi di un'esperienza non è mai uno stereotipo come per la locusta, ma ritorna arricchito dal nostro vissuto; come ben sappiamo quando rileggiamo un poeta o riascoltiamo un pezzo di musica a noi caro dopo molto tempo.

Il fatto che la spirale sia innescata da stimoli significativi (quelli per intenderci che stimolano gruppi di neuroni a organizzarsi in una rete sincronizzata) dà un solido fondamento realistico al nostro rapporto col mondo, escludendo il sospetto di relativismo.

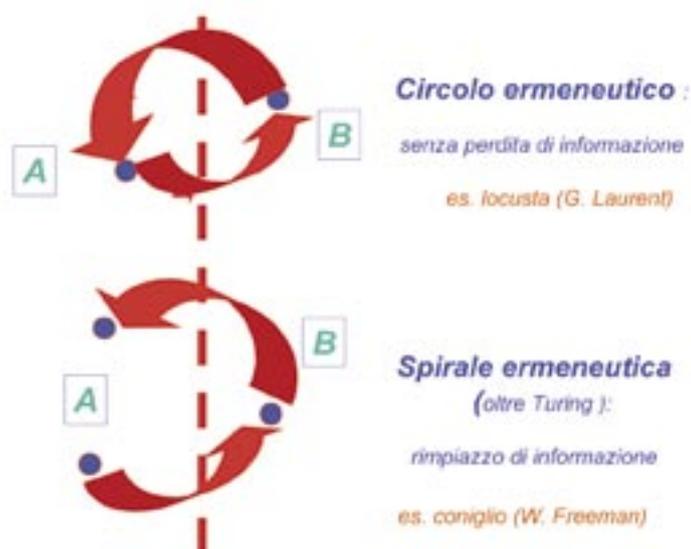


Figura 9 - Circolo ermeneutico e spirale ermeneutica. Il primo concetto, usato nelle scienze umane da H. G. Gadamer (*Verità e metodo*, 1960), riceve qui una formulazione precisa. Supponiamo di operare in uno spazio semantico finito.

Scelta una connotazione A1 per la parola A, un opportuno connettore mappa su un particolare significato di B. L'operazione inversa riporta ad A1, per cui si parla di *circolo*. Se ora i significati non sono congelati a priori, ma evolvono nella interazione con l'agente cognitivo, il ritorno avverrà su A2 differente da A1, in quanto nel frattempo l'agente ha vissuto una nuova esperienza: dobbiamo rappresentare i significati con un insieme *non finito* ma suscettibile di essere accresciuto nel tempo; la sequenza di operazioni cognitive si presenta ora come una *spirale ermeneutica*.

NOTE

¹ Un'analisi dettagliata di questa formulazione è stata presentata in F. T. Arecchi, «Complexity and emergence of meaning: toward a semiophysics», in *Complexity and Emergence*, a cura di E. Agazzi e L. Montecucco, World Scientific, Singapore, 2002, pp. 115-146.

² Rev. T. Bayes, «An essay toward solving a problem in the doctrine of chances», *Philosophical Transactions of the Royal Society*, 53 (1763), pp. 370-418.

³ A. Turing, «On Computable Numbers With an Application to the Entscheidungsproblem», *Proceedings of the London Mathematical Society*, 2/42 (1936), pp 230-265.

⁴ In conformità al *Dictionary of Algorithms and Data Structures* (a cura di Paul E Black, US National Institute of Standards and Technologies, 2006).

BIBLIOGRAFIA

Dato il carattere di questa presentazione, non è sembrato opportuno appesantirla con un apparato bibliografico. Sui vari aspetti della complessità, rimando ai lavori elencati sulla mia homepage: www.inoa.it/home/arecchi e, in particolare, a

F. T. Arecchi, A. Farini, *Lexicon of Complexity*, Studio Editoriale Fiorentino, Firenze 1996.

F. T. Arecchi, *Caos e Complessità nel Vivente*, IUSS Press, Pavia, 2004.

F. T. Arecchi, *Cognitive Dynamics: Complexity and Creativity*, DICE (in corso di pubblicazione).