

# La biologia teorica come scienza della complessità

Renato Nobili

Dipartimento di Fisica "G. Galilei" - Università di Padova

E-mail: [rnobili@padova.infn.it](mailto:rnobili@padova.infn.it)

## I. ENTROPIA E INFORMAZIONE

### 1.1. Il costo termodinamico dell'informazione

L'intima connessione tra l'entropia di un sistema termodinamico e l'informazione che si può ricavare effettuando osservazioni e misure sul sistema stesso fu messa in luce da Leo Szilard nel 1929.

Per definizione, un bit d'informazione è l'eliminazione dell'incertezza associata a priori a due alternative equiprobabili. Se le alternative non sono equiprobabili, l'informazione  $I$  che si guadagna è inferiore a un bit. Precisamente si trova  $I = \log_2(1/p) = -\ln p / \ln 2$ . Si dimostra facilmente che  $0 \leq I \leq 1$ . Pertanto, quando si verifica mediante un processo di misura se il sistema possiede o no una certa proprietà si ottiene al più un bit d'informazione.

Una processo di misura ha un costo energetico minimo calcolabile in termini di aumento di entropia associato alla dissipazione energetica che accompagna inevitabilmente l'osservazione. La dissipazione è dovuta al fatto che l'acquisizione dell'informazione può aver luogo solo attraverso l'assorbimento da parte di un apparato rivelatore dell'energia trasportata da almeno un fotone. L'assorbimento consiste in definitiva nel fatto che l'energia del fotone viene degradata e dispersa in calore alla temperatura di fondo del cosmo attraverso un processo termodinamico irreversibile. L'aumento di entropia che accompagna l'eliminazione di un'incertezza tra due alternative che prima della misura venivano date come equiprobabili è almeno pari a  $\Delta S = k \ln(1/p) = k \ln 2 I$  erg/grado, dove  $k = 6.63 \times 10^{-16}$  è la costante di Boltzmann.

L'argomento di Szilard è il seguente: l'informazione guadagnata dall'apparato di osservazione potrebbe essere usata per diminuire l'entropia del sistema stesso di una quantità pari all'informazione acquisita (misurata in unità termodinamiche). Se il suo costo fosse inferiore a  $\Delta S$  il secondo principio della termodinamica potrebbe essere violato mediante una conveniente manipolazione del sistema. Pertanto nelle condizioni più favorevoli ogni guadagno d'informazione produce un aumento di entropia equivalente all'informazione guadagnata. Si comprende facilmente come una variazione positiva di entropia possa corrispondere a una perdita d'informazione.

Come è noto, l'entropia misura lo stato di disordine termodinamico di un sistema ed è massima per i sistemi in equilibrio termodinamico. Pertanto la "distanza" dello stato attuale di un sistema dallo stato di equilibrio termodinamico può idealmente misurarsi come differenza tra l'entropia massima  $S_{max}$  che lo stesso raggiungerebbe in un tempo sufficientemente lungo qualora venisse isolato in una scatola e quella  $S$  che il sistema avrebbe nell'istante iniziale del suo isolamento.

Per le ragioni spiegate sopra, questa variazione negativa di entropia  $\Delta S = S_{max} - S$  equivale a una quantità d'informazione  $I = (S_{max} - S) / (k \ln 2)$  relativa allo stato del sistema e può pertanto definirsi *informazione termodinamica* o *neg-entropia*. Poiché i processi informativi (scrittura, lettura ricodificazione di dati ecc.) comportano sempre aumenti irreversibili di entropia, si può comprendere come essi possano avere luogo solo in sistemi termodinamici sufficientemente distanti dall'equilibrio. L'universo è sede di processi informativi perché il suo stato termodinamico è lontano dall'equilibrio.

Secondo la termodinamica fondata da Clausius a metà dell'800, cioè quella basata sulle nozioni termometriche di temperatura e calore, l'entropia si presenta come una grandezza fisica distribuita tra le parti di un sistema termodinamico. Da questo punto di vista la stessa informazione potrebbe sembrare qualcosa di simile a una proprietà fisica della materia. Le variazioni infinitesime di entropia secondo Clausius sono definite da formule del tipo

$$dS_C = \frac{dQ_1}{T_1} + \frac{dQ_2}{T_2} + \dots + \frac{dQ_N}{T_N},$$

dove  $dQ_i$  sono quantità infinitesime di calore incamerate reversibilmente dalle  $N$  parti di un sistema che si suppongono trovarsi alle temperature  $T_i$ . L'entropia  $S_C$  di un sistema può essere calcolata come variazione di entropia totale necessaria per portare il sistema da uno stato in cui tutte le temperature sono prossime allo zero assoluto a uno in cui le parti del sistema hanno temperature  $T_1, T_2, \dots, T_N$ . Il passaggio logico che permette di calcolare l'entropia assoluta è possibile perché il terzo principio della termodinamica stabilisce che allo zero assoluto  $S_C=0$  per ogni sistema.

Una determinazione completamente diversa dell'entropia fu introdotta da Boltzmann verso la fine dell'800. Questo autore elaborò in modo originale le idee di meccanica statistica sviluppate da Maxwell fino a concludere che l'entropia poteva esprimersi come una funzione delle probabilità  $p_i$  dei possibili stati del sistema secondo la formula

$$S_B = -k \sum_i p_i \ln p_i.$$

Così, mentre l'entropia di Clausius si presenta come una funzione di grandezze fisiche oggettive del sistema, quella di Boltzmann si presenta come una funzione dei *valori di aspettazione* relativi ai possibili stati del sistema. Ora, questa "aspettazione" non può avere il senso di un contenuto oggettivo proprio del sistema ma di una quantità d'informazione posseduta da un apparato descrittore e manipolatore esterno al sistema. In altri termini,  $S_B$  rappresenta l'incertezza del soggetto descrittore esterno in relazione alle sue possibilità di osservare e manipolare il sistema stesso. Per questa ragione la relazione  $S_C = S_B$  non può essere intesa come un'uguaglianza tra due diversi modi di descrivere la stessa cosa ma come un'equazione tra due cose diverse.

Il significato informativo dell'entropia di Boltzmann traspare direttamente dalla formula che fornisce la *misura d'incertezza* introdotta da Claude Shannon nella sua teoria della comunicazione (1949). L'entropia (comunicazionale o informativa) di una sorgente di messaggi rappresenta l'incertezza aprioristica dell'apparato ricevitore, e l'informazione acquisita dal ricevitore viene definita come riduzione di quella incertezza conseguente all'acquisizione di un particolare messaggio o di una parte di esso. La formula che fornisce la misura shannoniana dell'incertezza  $H$  associata a un insieme di possibili messaggi di probabilità  $p_i$  è praticamente identica alla formula di Boltzmann e differisce da questa soltanto per fattore di proporzionalità  $1/k \ln 2$ , in questo modo che l'incertezza informativa venga misurata in *bit* invece che in erg/grado. Risulta pertanto evidente che la formula di Boltzmann non è altro che la misura dell'incertezza relativa alla descrizione "soggettiva" dello stato del sistema.

Leon Brillouin (1950) ha chiarito in modo completo il nesso tra entropia termodinamica ed entropia informativa con particolare riguardo ai processi di acquisizione e uso dell'informazione.

La teoria shannoniana dell'informazione riguarda essenzialmente la misura quantitativa dell'informazione, ad esempio la quantità di lettere alfabetiche usate per comporre una frase in un dato linguaggio, di cifre numeriche necessarie per codificare un dato sperimentale, di componenti di frequenza necessarie per generare un segnale elettrico, acustico o elettromagnetico ecc. Ciò non ha nulla a che fare col valore o il significato che l'informazione può avere per i suoi utenti in casi concreti.

In realtà non è lecito parlare di valore o significato dell'informazione, o dell'informazione in senso puramente qualitativo, se non in relazione agli usi che se ne possono fare. Se ne può parlare solo nel contesto di una descrizione globale del rapporto tra gli oggetti che "contengono" o "trasportano" un certo tipo di informazione e gli utenti reali o possibili di tale tipo di informazione. Questo stato di cose rende impossibile considerare l'informazione come una grandezza fisica incorporabile nella materia. In altri termini, la nozione d'informazione impone concetti e metodi che esulano da una visione puramente riduzionista dei fenomeni naturali; visione secondo la quale la conoscenza di un sistema, la definizione e l'interpretazione dei suoi comportamenti si riduce alla rappresentazione di come interagiscono le parti del sistema.

A questo punto dovrebbe essere chiaro che, se la nozione di informazione è essenziale per l'interpretazione dei fenomeni biologici, allora semplicemente non può esistere una visione riduzionista del vivente. Le proprietà di un organismo vivente non possono caratterizzarsi soltanto sulla base delle relazioni che si stabiliscono tra le sue parti ma anche dalle relazioni che l'organismo ha con tutte le cose dell'universo che lo circonda.

## 1.2. Entropia e indeterminazione quantistica

Quanto finora detto non basta a svelarci l'arcano della formula  $S_C = S_B$ . Per comprendere bene quale sia il legame che intercorre tra le due determinazioni dell'entropia bisogna uscire dagli schemi concettuali della fisica classica e familiarizzarsi con alcune nozioni di meccanica quantistica. Bisogna in primo luogo considerare che per la meccanica quantistica ogni *fenomeno fisico* è una corrispondenza che si stabilisce improvvisamente tra uno stato di un sistema osservato e uno stato di un apparato osservatore. Prima dell'evento la connessione tra gli stati dei due sistemi era affetta da una certa indeterminazione e bisognava dire che lo stato del sistema *osservato+osservatore* era una *sovrapposizione quantistica* di corrispondenze tra stati osservabili e osservazioni possibili. L'atto d'osservazione fa precipitare questa sovrapposizione di possibili corrispondenze in una sola di esse secondo una legge probabilistica. In altri termini, l'indeterminazione oggettiva del sistema osservato e l'incertezza soggettiva del sistema osservatore si annichilano insieme.

Come questo abbia a che fare con la termodinamica lo si comprende facilmente dal seguente esempio. Consideriamo una singola molecola posta in una scatola in equilibrio termodinamico a una certa temperatura. Si supponga che il volume interno della scatola sia divisibile in due parti uguali mediante un setto separatore inseribile lateralmente senza apprezzabile spesa di energia e in modo perfettamente reversibile. Secondo le leggi della termodinamica, l'inserzione del setto non può far diminuire l'entropia del contenuto della scatola, così come l'estrazione del setto non può farla aumentare. Tuttavia, secondo la *meccanica classica*, dopo l'abbassamento del setto la molecola si trova certamente da una parte o dall'altra parte della scatola e il volume dello spazio che la rinchioda è dimezzato. Poiché la termodinamica dei gas assegna alla molecola l'entropia di Clausius  $S_C = k \ln(V/V_0)$ , dove  $V$  è il volume della scatola e  $V_0$  quello della molecola, si dovrebbe concludere che  $S_C$  è diminuita di un fattore  $1/\ln 2$ , cioè di una quantità pari a un bit d'informazione.

L'entropia di Boltzmann  $S_B$  invece rimane invariata, perché la l'aspettativa di trovare la molecola da un lato o dall'altro è rimasta quella di prima. Solo osservando dove si trova effettivamente la molecola si può attribuire a  $S_B$  un valore inferiore a quello iniziale. È chiaro che a un certo punto le due determinazioni di entropia risultano in contraddizione.

L'interpretazione quantistica non genera invece alcuna contraddizione. Dopo l'inserzione del setto lo stato della molecola nella scatola è una sovrapposizione quantistica dei due stati "essere di qua" ed "essere di là" e bisognerà mettere in atto un processo di osservazione per verificare dove si trovi realmente la molecola; per far precipitare la sovrapposizione degli stati da una parte o dall'altra permettendoci di guadagnare l'informazione corrispondente. In questo caso l'uguaglianza  $S_C = S_B$  è sempre soddisfatta.

Questo semplice esempio ci fa capire tre cose molto importanti: *i*) l'entropia di Clausius è associata all'indeterminazione quantistica intrinseca dello stato del sistema e non all'incertezza dell'osservatore; *ii*) l'equazione  $S_C = S_B$  è una relazione tra indeterminazione "oggettiva" di un sistema osservato e l'incertezza "soggettiva" del sistema osservatore; *iii*) se il sistema osservatore viene considerato esso stesso un oggetto fisico allora l'incertezza soggettiva non è altro che indeterminazione dello stato quantistico del sistema osservatore.

In realtà tra sistema quantistico osservato e sistema osservatore vi è un'asimmetria fondamentale: a differenza dal primo, il secondo è sede di processi termodinamici irreversibili.

La descrizione quantistica del mondo fisico pone la questione della natura termodinamico-informazionale della misura e dell'osservazione in una nuova luce. Uno stato fisico è descritto da un'entità (ampiezza d'onda di probabilità) che rappresenta non tanto ciò che è direttamente osservabile in natura, quanto piuttosto la potenzialità di molteplici aspetti osservabili, i quali risultano spesso non osservabili simultaneamente. La natura e la struttura di questi aspetti dipendono dalla natura e dalla struttura dagli apparati macroscopici di misura e osservazione esterni al sistema osservato, cosicché nessun stato fisico avrebbe senso se non esistesse un livello macroscopico del mondo fisico nel quale gli apparati di osservazione sono dati come sistemi termodinamico-informazionali, in pratica sistemi termodinamici fuori equilibrio. Di nuovo, qui si presenta il fatto che l'aspetto informazionale di ogni porzione di realtà fisica - poiché tale è in ultima analisi il complesso delle misure e osservazioni di cui tratta la meccanica quantistica - è definito solo in rapporto a un sistema esterno: il mondo degli eventi termodinamici macroscopici, di cui ogni realtà microscopica è parte integrante e di cui lo stesso mondo macroscopico consiste.

### 1.3. Sistemi quantistici infiniti

Il problema della relazione tra mondo microscopico e mondo macroscopico, ovvero la spiegazione di cosa sia fisicamente un'osservazione quantistica, non trova soluzione nell'ambito dei sistemi fisici con un numero finito di gradi di libertà. In questo ambito nessun processo irreversibile può avere luogo. È pertanto importante che si sia potuto dimostrare (Hepp, 1974) che, ampliando la descrizione quantistica con l'assunzione che il sistema fisico possieda un numero infinito di gradi di libertà, si riesce a spiegare come possano aver luogo processi irreversibili, compreso il fenomeno della misura quantistica.

Hepp ha dimostrato che l'evoluzione temporale quantistica di un sistema con infiniti gradi di libertà può determinare la biforcazione dello spazio dello degli stati (spazio di Hilbert) in due spazi di disgiunti, che non possono più comunicare tra loro. Lo stato finale del sistema finisce nell'uno o nell'altro di questi spazi, con probabilità dipendenti dalle proprietà della sovrapposizione quantistica iniziale. Il principio di sovrapposizione quantistica non si applica più a questi due esiti finali che devono pertanto essere rappresentati come due possibilità reciprocamente esclusive. Queste biforcazioni devono considerarsi come due esiti macroscopici diversi di un processo di misura.

Il fenomeno descritto da Hepp dipende dal fatto che lo spazio della rappresentazione fisica di un sistema infinito non è uno singolo spazio hilbertiano, come nel caso dei sistemi con un numero finito di gradi di libertà, ma una *collezione infinita di classi non equivalenti di spazi disgiunti equivalenti* (von Neumann, 1936). Questa collezione ha persino la cardinalità del continuo. Questo *Grand Ensemble* di classi non equivalenti rappresenta la collezione dei possibili stati termodinamici macroscopici del sistema. È questo l'ambito in cui deve ricercarsi la base logica generale per una descrizione completa dei sistemi quantistici immersi in un mondo macroscopico di processi termodinamico-informazionali.

Questo stato di cose suggerisce che i fenomeni fisici ammettono un'interpretazione completa solo in una visione cosmologica dell'universo quantistico. Nelle singole rappresentazioni del *Grand Ensemble*, come pure nelle rappresentazioni delle porzioni finite

di universo, lo stato del sistema riveste il significato di qualcosa che è dotato di forma di esistenza *interfenomenica* puramente virtuale che non accede mai al livello della sua realizzazione fenomenica. Infatti la nozione di *realtà fisica*, caratteristica del livello fenomenico, si dà solo al livello delle classi non equivalenti del *Grand Ensemble* di von Neumann. In questo livello di descrizione infinitamente più complesso i modi di esistenza virtuale si differenziano e si multiforcano emergendo fenomenicamente quali processi termodinamico-informazionali macroscopici.

Attraverso l'emergenza fenomenica delle proprietà macroscopiche, la struttura quantistica dell'universo si presenta come il sostrato di un processo termodinamico-informazionale che evolve generando una successione infinita di eventi irreversibili che lo realizzano in forme sempre più complesse. In questa visione l'intero universo si presenta come sistema vivente in perenne evoluzione.

#### 1.4. Biologia e informazione

L'idea che la vita si basi su processi di accumulazione e uso d'informazione termodinamica (neg-entropia) è stata introdotta per la prima volta da Erwin Schrödinger (1947). La visuale secondo la quale non tanto l'energia, quanto piuttosto l'informazione sia l'autentica linfa vitale della natura, determinò una svolta storica nel pensiero biologico. Essa ebbe un'importanza cruciale per il progresso della biologia degli anni successivi per almeno tre ragioni:

1. Sistemi fisico-chimici termodinamicamente dissipativi e aperti sono capaci di auto-organizzazione (Prigogine, 1960). Questa proprietà si manifesta come tendenza di questi sistemi ad accedere a stati termodinamici lontani dall'equilibrio e perciò caratterizzati da una quantità di informazione termodinamica più o meno elevata.
2. L'introduzione della nozione di informazione da parte di Shannon fu concomitante in tutti i sensi alla nascita della cibernetica, o teoria del controllo automatico, ad opera di Norbert Wiener (1948). Un vasto repertorio di processi regolativi assicurano che i sistemi biologici, tutte le loro parti e le loro funzioni, siano strutturalmente stabili. Ciò costituisce anche una premessa per l'analisi delle fenomenologie critiche (Thom, 1975; Arnol'd, 1986) che caratterizzano il comportamento dei sistemi biologici, delle varie forme di stabilità strutturale e della perdita di tali condizioni.
3. Una descrizione puramente fisico-chimica può essere sufficiente a spiegare i dettagli di funzionamento dei sistemi biologici, ma non a caratterizzarne e spiegarne il comportamento in quanto organismi viventi. L'interpretazione delle eliche di DNA (acido desossiribonucleico) come memoria di programma (Crick e Watson, 1963) e delle sequenze di RNA (acido ribonucleico) come messaggi che codificano per proteine esula dalla fenomenologia fisico-chimica e coinvolge estensivamente nozioni tipiche della teoria dell'informazione, in particolare dei teoremi relativi alla codificazione ottimale.

Questi concetti sono indispensabili a una moderna visione delle scienze biologiche, per analizzare e interpretare correttamente la fenomenologia del vivente in tutti i suoi livelli e scaomparti: dalla biologia molecolare alla neurofisiologia. Tuttavia essi rispondono solo in parte alle questioni fondamentali che si pongono in ambito bioscientifico. In anni recenti sono stati rilevati i limiti e i difetti di una concezione puramente darwinista o neo-darwinista dell'evoluzione biologica ed è stata messa in evidenza che la tendenza all'auto-organizzazione dei sistemi biologici interagisce col processo di selezione naturale in forme complesse spesso contrastandolo (Kauffman, 1993). In concomitanza a ciò si è delineata la tendenza a interpretare alcuni aspetti della complessa fenomenologia della riproduzione cellulare in termini propri della *teoria degli algoritmi* e di quella dei *linguaggi* (Arbib, 1972; Fontana, 1992). Quali sono le basi concettuali di questi ampliamenti?

## II. I NUOVI CONCETTI DELLA BIOLOGIA TEORICA

### 2.1. Dualismo struttura-comportamento. Emergenza di qualità meta-strutturali nei sistemi complessi

In *The Theory of Self-Reproducing Automata* (1948-1956), von Neumann sollevò una questione di grande interesse teorico: *la descrizione della struttura di un sistema e quella del suo comportamento appartengono a due tipi logici diversi*. La differenza essendo simile a quella intercorrente tra il livello del "linguaggio" diretto, o di primo livello, i cui enunciati veicolano informazione sul mondo dei fatti naturali, e quello del suo "metalinguaggio", i cui enunciati veicolano informazione sulle proprietà del linguaggio del primo livello, assunto come mondo del discorso sui fatti. E' stato dimostrato che esistono enunciati meta-linguistici il cui significato non può tradursi in nessuna proposizione (finita) del linguaggio. Ad esempio il concetto di *verità* non è definibile nel linguaggio del primo livello (Tarski, 1936).

La distinzione tra struttura e comportamento è analoga a quella tra linguaggio e metalinguaggio. Esistono sistemi fisici la cui struttura è perfettamente descrivibile ma i cui comportamenti non sono descrivibili, ad esempio i computer. La teoria di Turing sulla imprevedibilità, in generale, della terminazione di un calcolo ricorsivo ne è l'esempio più noto. La teoria di Turing è basata sul seguente concetto fondamentale: solo un calcolatore capace di esibire comportamenti indescrivibili sarà capace di eseguire ogni calcolo che possa essere eseguito da un altro calcolatore, potrà cioè possedere la qualità dell'*universalità algoritmica*.

Nella fisica si assume che ogni sistema fisico finito sia in linea di principio descrivibile. Si tratta di una congettura ben fondata, dal momento che finora non è mai stato dato un controesempio. In linea di principio, anche una cellula o un organismo pluricellulare, ancorché complicati, sono descrivibili in termini finiti. Ma i comportamenti di un sistema descrivibile in termini finiti può non essere descrivibile in termini finiti. Quando questo accade, la complessità del comportamento risulta infinitamente maggiore di quello della struttura che lo esibisce; in altri termini, il comportamento non possiede una struttura definita o un modello strutturale. Non possiamo rappresentare un tale comportamento con un diagramma, un disegno o una collezione di disegni, non possiamo descriverne l'architettura. Per questa ragione dobbiamo parlare di carattere generalmente *meta-strutturale* dei comportamenti.

Si comprende allora che se un sistema ha capacità *costruttive* - nel senso che durante la sua evoluzione temporale esso è in grado di modificare cumulativamente lo stato di una struttura ambientale esterna - allora è possibile che la struttura che viene generata dal comportamento di questo sistema raggiunga in un tempo abbastanza lungo una complessità enormemente maggiore di quella del sistema costruttore stesso. Ciò spiega perché un ovulo fecondato sia capace di generare un organismo.

Un esempio molto elementare di questo stato di cose è dato dal famoso frattale di Mandelbrot: la struttura del programma di calcolo che lo genera ammonta a poche righe di codice, mentre la struttura del frattale che viene prodotta dal comportamento del calcolatore che esegue questo semplicissimo programma è indescrivibilmente complicata, più complicata del calcolatore stesso, la sua "forma" essendo intraducibile nel linguaggio geometrico.

Questo spiega anche perché il linguaggio della biologia si estende *necessariamente* oltre quello della fisica e della chimica e richiede nozioni come *ordine, organizzazione, informazione, memoria, messaggio, omeostasi, feed-back* ecc., che non sono definibili in puri termini fisici, cioè come proprietà fisiche della materia. Certamente, i sistemi biologici sono pur sempre sistemi fisici che possono ritenersi descrivibili almeno in linea di principio - per quanto una descrizione effettiva possa ottenersi con difficoltà. Il loro comportamento potrà essere descritto in termini di evoluzione fisica o chimico-fisica solo limitatamente ad alcuni aspetti, ma ben presto si arriverà a un punto gli aspetti interessanti del loro comportamento

non sono più riconducibile ai concetti e alle nozioni della fisica e della chimica. Ciò comincia ad accadere quando i processi biologici vengono meglio interpretati come forme di *circolazione, codificazione, memorizzazione, utilizzo di messaggi ecc.* - cioè mediante il linguaggio della teoria dell'informazione e del controllo - piuttosto che come semplici relazioni di causa-effetto tra entità o grandezze strutturali, cioè fisico-chimiche, del sistema.

Questo livello è soltanto il primo e il più elementare della descrizione biologica. Procedendo attraverso i gradini della complessità biologica, la sintassi deve arricchirsi di termini sempre nuovi e il pensiero interpretativo di nuovi concetti. In cima alla scala troveremo il problema del comportamento del cervello umano e della genesi del pensiero autoriflessivo, per spiegare il quale avremo bisogno di nozioni primitive e idee assai più evolute di quelle che sono sufficienti a descrivere il funzionamento dei cervelli più semplici.

## **2.2. I livelli della complessità biologica: algoritmico, descrittivo, computazionale...**

Nella letteratura si danno diverse nozioni di complessità: descrittiva, algoritmica, di processo ecc. Kolmogorov (1965) ha definito la *complessità algoritmica* di un oggetto matematico come la lunghezza del più piccolo programma capace di governarne la produzione. Stando a questa definizione la complessità del frattale di Mandelbrot ammonterebbe a poche decine di bytes. Ma se si osserva un'immagine di questo frattale dotata di buona risoluzione, magari "zoomandoci" dentro in vari punti ripetutamente, non si può non restare impressionati dall'enorme complessità dei dettagli. D'altronde, se non si conoscesse la formula che lo genera non sarebbe possibile risalire a essa neanche con l'ausilio del più grosso calcolatore del mondo: l'ignoranza dei dettagli, a qualsiasi ordine di precisione nella rilevazione dei dati dell'immagine permetterebbe solo la determinazione di un numero infinito di formule assai complicate, ognuna delle quali sarebbe capace di produrre lo stesso risultato a quel livello di precisione, ma non a livelli di precisione superiori. Bisogna dunque ammettere che c'è una notevole differenza tra la *complessità descrittiva* del frattale e la sua *complessità algoritmica*.

D'altronde, le sole nozioni di complessità algoritmica e complessità descrittiva non bastano a caratterizzare il fenomeno della generazione del frattale. Il processo di calcolo che lo produce richiede sequenze di operazioni logico-aritmetiche e la memorizzazione temporanea dei dati piuttosto semplici. Tuttavia esistono processi di calcolo ricorsivo in cui le operazioni del calcolatore e le modalità di accesso alla memoria di lavoro del computer si complicano sempre più mano a mano che il calcolo procede, e ciò in modo non direttamente correlato alle complessità algoritmica e descrittiva del risultato prodotto. Questo fa comprendere che la nozione di *complessità computazionale* si aggiunge alle altre due come nozione ulteriore.

In realtà la tipologia della complessità è essa stessa infinitamente complessa. Non potrebbe essere altrimenti dato che un'ipotetica trattazione esaustiva della nozione di complessità dovrebbe permettere quantomeno la classificazione di tutte le espressioni di complessità che si danno in natura e di quelle che potranno imprevedibilmente generarsi nel futuro. Chiaramente questo compito non sarebbe terminabile.

## **2.3. Sistemi dinamici a stati continui e a stati discreti: transizione al caos e universalità algoritmica**

È stato suggerito che i sistemi auto-organizzanti si collochino nella regione che separa il territorio dell'ordine da quello del caos. Non potrebbe essere altrimenti, dato che i comportamenti ordinati sono proprio quelli di complessità limitata. Ma quanto è "larga" questa terra di nessuno? E come può essere garantita la stabilità strutturale di un sistema in una simile collocazione? Il problema dell'esistenza di sistemi auto-organizzanti, strutturalmente stabili e suscettibili di evolvere verso illimitati livelli di complessità non trova soluzione nella teoria dei sistemi i cui gli stati evolvono deterministicamente in uno spazio continuo. La ragione è molto semplice: ogni minima perturbazione dello stato iniziale provoca

mutamenti macroscopici e divergenti dell'evoluzione di questi sistemi. In considerazione di ciò, è assai rilevante il fatto che i sistemi deterministici a stati discreti, ad esempio gli automi cellulari di Wolfram (1985), pur presentando notevoli analogie di comportamento con quelli a stati continui (esistenza di stati attrattori, cicli limite e attrattori strani) presentino un'importante novità: al posto del caos dinamico si manifesta l'universalità algoritmica.

#### **2.4. Replicazione e riproduzione. Costruttori universali. Sistemi di Post. Grammatiche casuali**

Alcuni automi cellulari (von Neumann-Burks, 1966; Nobili e Pesavento, 1996) ammettono configurazioni che possiedono l'*universalità costruttiva*, cioè la capacità di generare nello spazio degli stati qualunque configurazione di stati locali: una nozione che generalizza quella di universalità computazionale (Turing, 1936). Per von Neumann, l'universalità costruttiva è una condizione necessaria perché un sistema sia capace di autoriprodursi ed evolvere. Questo concetto introduce nelle bioscienze un punto di vista nuovo e avanzato. Ad esempio, sulla sua scorta si può dire che la biogenesi primordiale non può essersi basata su semplici processi di replicazione molecolare, perché questi non sono dotati di universalità costruttiva, ma devono essersi generate in condizioni equivalenti all'esistenza di una macchina dotata di potenzialità comportamentali infinitamente complesse.

Tuttavia l'universalità costruttiva da sola non è sufficiente a caratterizzare le proprietà biopoietiche di un sistema. Ad esempio, i sistemi deterministici a stati discreti, per quanto possiedano l'universalità costruttiva, non sono buoni modelli per la biologia essenzialmente perché *non tollerano difetti*: piccoli errori nella struttura del sistema producono in generale grandi variazioni nei comportamenti e i comportamenti interessanti dipendono in modo così critico dai dettagli della struttura che non è pensabile che questi sistemi siano suscettibili di evoluzione nel senso darwiniano del termine.

La ricerca di costruttori universali strutturalmente e comportamentalmente stabili si colloca pertanto necessariamente nel territorio delle fenomenologie algoritmico-stocastiche. Per questi sistemi la teoria è solo agli inizi. E' naturale pensare che i sistemi biologici reali rientrino in questa categoria. La natura "algoritmica" dei sistemi biologici reali può essere riconosciuta nella struttura discreta dello *spazio delle proteine e degli acidi nucleici*, cioè nel fatto che gli "stati" del sistema sono rappresentati da ordinamenti di monomeri (aminoacidi, acidi nucleici) entro polimeri (proteine, DNA, RNA) e formano pertanto un insieme discreto di stati. Il fatto che questi oggetti interagiscano in ambiente termochimico assicura il carattere stocastico del processo. Gli insiemi autocatalitici, e gli ipercicli teorizzati da M.Eigen (1986, 1992) si collocano in questo contesto. Ma sono garantite per questi sistemi le condizioni di universalità costruttiva e di stabilità strutturale e comportamentale?

Circa la natura dei sistemi biologici reali si può dire di più: il carattere di stringhe (aperte o chiuse) rende le proteine e acidi nucleici simili a sequenze di lettere alfabetiche, cioè a parole di un linguaggio. I "significati" veicolati da tali "parole" possono identificarsi nelle funzioni enzimatiche, e in generale operazionali, che le stringhe (generalmente avvolte su se stesse in modi complicati) sono in grado di svolgere nell'ambiente biologico. Gli algoritmi adatti a descrivere questi sistemi assomigliano a quelli della teoria dei linguaggi, in particolari ai sistemi di Post (Post, 1943), per i quali è assicurata l'universalità computazionale.

Come Turing ha ricondotto la nozione di macchina che effettua un calcolo a quella di una sequenza di operazioni elementari, quali sono eseguite da una *macchina di Turing*, così Post ha ricondotto l'aspetto sintattico di una generale teoria logica a un insieme molto ridotto e semplice di operazioni su stringhe. Ciò che si richiede all'agente di calcolo per eseguire queste operazioni è (Aiello *et al.*, 1979):

- a) essere in grado di scandire una stringa di simboli e riconoscere in essa eventuali occorrenze di una data sottostringa;
- b) eliminare certe sottostringhe di una data stringa, ricordando le parti rimanenti;



c) ricompone una nuova stringa, inserendo certe sottostringhe costanti in certe posizioni.

La rilevanza di questo approccio alla teoria dei linguaggi formali sta nel fatto che a ogni sistema di Post è possibile associare almeno una macchina di Turing equivalente. Così rimane dimostrata l'universalità computazionale dei sistemi di Post. E' chiaro che un'estensione di sistemi di Post nella direzione della creazione di un sistema dotato di universalità costruttiva rende questi algoritmi degli eccellenti candidati dei processi biologici elementari. Ed è anche chiaro che si richiederebbe a questo fine una trattazione statistica di processi di Post stocastici.

Recenti simulazioni al computer hanno dimostrato che *grammatiche casuali* che in qualche modo implementano queste proprietà tendono ad evolvere verso insiemi globalmente autocatalitici, auto-organizzativi e capaci di adattarsi a perturbazioni ambientali (W.Fontana, 1992; S.A.Kauffman, 1993).

Certamente, questi concetti, insieme a vari altri che sono presentati dagli altri relatori, sono oggi necessari per analizzare e interpretare correttamente la fenomenologia del vivente, e ciò a tutti i livelli: dalla biologia molecolare all'ecologia. Tuttavia essi rispondono solo in parte alle domande che si possono porre in ambito bioscientifico. In anni recenti si è delineata la tendenza a interpretare alcuni aspetti della complessa fenomenologia della riproduzione cellulare in termini propri della "teoria degli algoritmi" e della "teoria dei linguaggi". Del resto, l'idea di von Neumann che un sistema dotato di capacità autoriproduttiva ed evolutiva deve necessariamente contenere un "costruttore universale" e un "calcolatore universale" è di tal genere.

Padova, Febbraio 2004.

#### BIBLIOGRAFIA

1. Aiello, M., Albano, A., Attardi, G. e Montanari, U. (1979). *Teoria della computabilità, logica, teoria dei linguaggi formali*. Ed. ETS, Pisa.
2. Arbib, M.A. (1972), "Automata Theory in the Context of Theoretical Biology", in *Foundations of Mathematical Biology*, vol.2, pp.141-215, Academic Press, New York.
3. Arnol'd, V.I., Varchenko, A. et Goussein Zadè, S. (1986), *Singularités des applications différentiables*, voll.1,2, Ed. MIR, Moscow.
4. Ausiello, G. (1975), *Complessità di calcolo delle funzioni*, Boringhieri Editore, Torino.
5. Brillouin, L. (1950), *Science and Information Theory*, Academic Press, New York.
6. Crick, F.H.C. (1968), "The origin of the genetic code", *J.Mol.Biol.* **38**:367.
7. Dyson, F., *Origins of Life*, Cambridge University Press.
8. Eigen, M. (1986), "The Physics of Molecular Evolution", *Chemica Scripta*, **26B**, 13-26.
9. Eigen, M. (1992), *Gradini verso la vita*, Adelphi Editore, Milano.
10. Feldbaum, A. (1973), *Principes théorétiques des systèmes asservis optimaux*, Ed. MIR, Moscow.
11. Fontana, W. (1992). "Algorithmic chemistry" in *Artificial Life II: A Proceedings Volume in the Santa Fe Institute Studies in the Sciences of Complexity*, C.G.Langton et al. Ed.s, vol 10, Addison-Wesley.
12. Hepp, K. (1975). *Acta Physica Helvetica*.
13. Holden, A.V. (1986), *Chaos*, Manchester University Press.
14. Kauffman, S.A. (1993), *The Origin of Order - Self-organisation and Selection in Evolution*, Oxford University Press, New York.
15. Kolmogorov, A.N. (1965), "Three Approaches to the Definition of the Concept of the Amount of Information", *Problemy Peredachi Informatsii*, **1**, 193-302 MR. 32#2273.
16. von Neumann, J. (1936), On infinite direct products. *Compositio Mathematica*.

17. von Neumann, J. (1966), *The Theory of Self-Reproducing Automata*, edited and completed by A.W.Burks, University of Illinois Press.
18. Nobili, R. e Pesavento, U. (1996), "Generalised von Neumann's automata I: A Revisitation.", in *Artificial Worlds and Urban Studies*, E.Besussi e A.Cecchini Ed.s, DAEST Convegni 1, Venezia. Programmi di simulazione e altre informazioni disponibili nel sito <http://alife.santafe.edu/>
19. Post, E. (1943), "Formal reductions of the general combinatorial decision problem", *Am.J.Math.*, **65**, 197-268.
20. Prigogine, I. and Nicolis, G. (1977), *Self-Organization in Nonequilibrium Systems*, John Wiley & Sons Ed., New York.
21. Rizzotti, M. (1991), *Materia e vita. Big Bang ed evoluzione del vivente.*, UTET, Torino.
22. Rosen, R. (1967), *Optimality Principles in Biology*, Butterworths Ed., London..
23. Schrödinger, E. (1945), *What is life? The Physical Aspects of the living cell*, Cambridge Univ. Press, London; trad. it.: *Che cos'è la vita?*, Sansoni Ed., Firenze (1947).
24. Schroeder, M. (1991), *Fractals, Chaos, Power Laws*, Freeman & Comp., New York.
25. Shannon, C.E. and Weaver, W. (1949), *Mathematical Theory of Communication*, Illinois University Press, Urbana.
26. Szilard, L. (1929), traduzione in inglese dal tedesco: "On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings", *Behavioral Science* (1964).
27. Tarsky, A. (1936), "Der Wahrheitsbegriff in den formalisierten Sprachen": traduzione in inglese in *Logic, Semantic, Metamathematics*, 152-278, Clarendon Press, Oxford (1956).
28. Thom, R. (1969), "Topological Models in Biology", in *Towards a theoretical biology*, vol.3, Waddington Ed., Edinburgh University Press.
29. Thom, R. (1975), *Structural Stability and Morphogenesis*, W.A.Benjamin, Inc.
30. Turing, A.M.(1936), "On computable numbers, with an application to the "Entscheidungsproblem", *Proc. London Math. Soc. Ser. 2*, **42**, 230; **43**, 544E.
31. Wiener, N. (1948), *Cybernetics*, John Wiley, New York.
32. Wolfram, S. (1986), *Theory and Applications of Cellular Automata*, World Scientific Pub., Singapore.