

Systema Naturae, 2003, Vol. 5, pp. 237-306

MENTE & COMPUTAZIONE

Ignazio Licata

Istituto di Cibernetica Non-Lineare per lo Studio dei Sistemi Complessi

Via Favorita 9, 91025 Marsala (TP)

e-mail licata@programmazione.it

*A Lucyna Rodziewicz,
cmok!*

Sommario

Si propone un'analisi comparata dei modelli simbolici e sub-simbolici della mente, in relazione alla nozioni di Turing-computabilità, di amplificazione d'informazione e di sistema dissipativo. Si mostra che una contrapposizione tra le due classi di modelli costituisce un errore epistemologico. Si introduce il concetto di computazione naturale e se ne discutono le caratteristiche. Una relazione tra i modelli dell'attività cognitiva può essere delineata utilizzando la nozione di emergenza nei sistemi logicamente aperti. Si discute il ruolo che possono giocare i modelli quantistici per la costruzione di uno scenario teorico unitario nella scienza della mente, con particolare riguardo agli aspetti computazionali.¹

- Indice**
- 1 - I Processi Cognitivi tra 'mente' e 'cervello'
 - 2 - SST e PDP tra Simboli e Sistemi Dinamici
 - 3 - Sistemi, Informazione e Dissipazione
 - 4 - Automi Infiniti e Turing-computabilità
 - 5 - Computazione Naturale nei Sistemi Fisici e Biologici
 - 6 - Emergenza ed Apertura Logica
 - 7 - Teoria Quantistica e Processi Cognitivi
 - 8 - Bibliografia.

¹ Il dialogo interdisciplinare richiede che l'ascolto dell'interlocutore e delle sue ragioni sia tutt'uno con l'adozione di una prospettiva critica verso il proprio sistema concettuale. Questo lavoro, scritto da un matematico per i biologi, tenta di assolvere a quest'impegno attraverso uno stile *dialogico*, mantenendo comunque rigorosi i riferimenti fisico-matematici.

1 - I Processi Cognitivi tra ‘mente’ e ‘cervello’

La questione dello studio matematico dei processi cognitivi può essere effettuata utilizzando due classi di modelli che sembrano avere ben pochi punti in comune per quello che riguarda la struttura formale e gli assunti di base: (1) I modelli di tipo *simbolico-cognitivista*, nati nell’ambito dell’eredità cibernetica e sistemica negli anni ’50, costituiscono il nucleo concettuale degli strumenti teorici dell’Intelligenza Artificiale (IA) *classica*, sviluppati, tra gli altri, da A. Newell e H. Simon, J. McCarthy, M. Minsky, N. Chomsky e, più recentemente, raffinati e sostenuti da H. Putnam, J. Fodor e Z. Pylyshyn; (2) I *modelli connessionisti*, basati sui *sistemi distribuiti* e sulle *reti neurali* (RN), nascono negli ‘80 nell’ambito dello studio dei sistemi complessi e della matematica non-lineare e hanno subito un impetuoso sviluppo; tra i tanti nomi bisogna qui ricordare almeno J. Hopfield, T. Kohonen e P. Smolensky.

Le due classi sembrano divergere soprattutto in relazione alla cosiddetta *plausibilità biomorfa*, cioè alla possibilità di connettere questi modelli con i risultati delle neuroscienze, problema particolarmente rilevante in un periodo in cui nuove tecniche di *Brain Imaging* non-invasive permettono una conoscenza piuttosto dettagliata del sistema nervoso. I problemi che si delineano sono dunque di duplice natura. Da una parte dobbiamo analizzare le caratteristiche generali delle due classi di modelli, cercando di delineare i punti di contatto possibili o le irriducibili diversità; dall’altra è necessario chiedersi se, e in quale misura, le acquisizioni delle neuroscienze possano davvero autorizzare la supremazia di una tipologia modellistica sull’altra, ossia se la *plausibilità biomorfa* sia effettivamente un requisito necessario e sufficiente per la costruzione di una teoria generale dei sistemi cognitivi.

L’esistenza di un dibattito antichissimo sulle relazioni tra mente e cervello, o tra mentale e fisico, sarebbe già una buona ragione per dubitare della validità scientifica di un problema siffatto. Ci limiteremo dunque a qualche premessa epistemologica, in modo da fissare una volta per tutte il valore delle argomentazioni ‘biologiche’ in merito alle scienze cognitive.

Come vedremo più avanti, i limiti del cognitivismo tradizionale consisterebbero nello studio di una mente *astratta* e *asettica*, descrivibile in modo esaustivo utilizzando opportune strutture simboliche legate tra loro da un sistema di regole formali. Ne consegue che anche l’ambiente esterno deve essere in qualche modo *incorporato* nel sistema cognitivo

come rappresentazione, altrimenti sarebbe impossibile per la mente ricevere dall'ambiente una qualsiasi informazione ed elaborarla. Benché tutto questo possa sembrare artificioso, si noterà che ha una buona corrispondenza con l'idea che ci si fa solitamente della ricerca psicologica, dove si suppone proprio l'esistenza di un irriducibile soggetto-osservatore che descrive il mondo e opera in esso in relazione alle proprie strutture mentali. Sia nella psicologia 'tradizionale' che nell'approccio simbolico alla mente, si suppone l'esistenza di quello che J. Fodor ha denominato efficacemente *mentalese*, un livello di descrizione dei processi cognitivi che non necessita di alcun riferimento alle attività del cervello. Si badi bene che né gli psicologi né i cognitivisti hanno mai sostenuto una non-correlazione tra l'attività mentale e i processi cerebrali, ma si sono limitati piuttosto a difendere l'astrattezza dei propri modelli attraverso la loro evidente utilità pratica; in definitiva, nessuno di noi, mentre parla o gioca a scacchi o cammina scegliendo un percorso ottimale, può avere alcun accesso alle attività neuronali che *supportano* tali attività!

Per quanto riguarda invece l'aspetto cognitivamente *neutro* di questi modelli, sono ormai classiche le critiche di T. Nagel e di J. R. Searle, che possono riassumersi dicendo che una rappresentazione simbolica non può, per definizione, contenere alcun *peso semantico* in grado di rendere conto delle componenti individuali della cognizione, come di quelli emozionali e motivazionali. Non è il caso qui di riprendere le argomentazioni del pipistrello o della scatola cinese, o le loro innumerevoli varianti; basterà ricordare che i difensori del cognitivismo hanno ribattuto essenzialmente che le caratteristiche semantiche di un processo cognitivo possono sempre essere ricondotte ad un particolare schema simbolico che darebbe al *mentalese* anche quelle peculiari caratteristiche di importanza, profondità e priorità senza le quali non si dà alcuna reale esperienza mentale. In altre parole, anche se a un ordine diverso e di più alta complessità, sarebbe un po' come negli scacchi, dove alcune regole, ad esempio la sicurezza del re, hanno la priorità rispetto ad altre, pur importanti, come il vantaggio materiale. In pratica però questa idea ha trovato un'applicazione piena soltanto nell'ambito dei cosiddetti *micro-mondi*, come appunto quello del gioco degli scacchi (dove sintassi e semantica coincidono!), e dei *Sistemi Esperti*.

Il pregio di aver individuato dei modelli formali per la memoria (G. Sperling, memoria iconica), l'attenzione (teoria dei filtri di Broadbent e Treisman), il linguaggio (modello della produzione linguistica

di Chomsky), etc., non deve farci però dimenticare i limiti reali del cognitivismo classico, che consistono soprattutto nell'incapacità di spiegare la formazione e la flessibilità di ogni *mentalese*. Anche i modelli meno *statici* di strutture simboliche, come quelli suggeriti da M. Minsky nel suo famoso *La Società della Mente*, in cui si postula l'emergenza del *mentalese* dalla cooperazione di una gran moltitudine di agenti cognitivi (paradigma nel quale rientrano parecchie teorie e sistemi, come la *mente modulare* di Fodor o l'ATC di Anderson) appaiono molto più adatti alla progettazione di un'intelligenza artificiale che alla spiegazione esauriente dell'intelligenza naturale. In particolare, ci sarebbe nel cuore concettuale del cognitivismo un vizio razionalista di fondo, per cui verrebbe penalizzata nei vari modelli la ricchezza delle interazioni tra l'organismo, la mente e l'ambiente.

A questo tipo di critiche il connessionismo sembra rispondere proprio con la *plausibilità biologica*. Nei modelli connessionisti, infatti, gli stati e i processi mentali sono rappresentati da pattern neuronali dinamici, e non è necessario dunque postulare alcuna rappresentazione simbolica *a priori* di un concetto o di una regola, ma questi sono il risultato di complessi equilibri spazio-temporali connessi alle interazioni organismo-ambiente, e al *feedback* tipico di ogni processo biologico, e dunque anche dei processi cognitivi. E' evidente, in questo contesto, l'importanza dei processi percettivi, la cui complessità temporale, o profondità logica, sfuggiva alle possibilità dell'approccio simbolico. Infatti una delle proprietà peculiari delle RN è proprio quella di poter affrontare regimi transitori di tipo ordine-disordine in tempi molto brevi. C'è una confluenza naturale, dunque, tra le RN e le teorie psicologiche fondate su un'attenta considerazione del contesto biologico, come quelle di J. J. Gibson e J. Piaget. Il *mentalese* di Fodor qui non è più un super-linguaggio concettuale di tipo innato, ma piuttosto qualcosa che emerge dall'attività neuronale in virtù della sua *plasticità*. In questo senso si parla degli approcci connessionisti come *sub-simbolici*.

In termini più rigorosamente matematici, quella non-linearità che è connessa ai processi caotici e di auto-organizzazione tipica dei sistemi complessi, veniva invocata implicitamente nell'approccio cognitivista-simbolico e nell'IA 'classica' dalla cooperazione sinergica tra i vari agenti cognitivi, ma in pratica era difficile, se non impossibile, ottenere genuini effetti non-lineari da strutture simboliche così *rigide*, e con domini di input-output definiti in modo formale come le stesse regole di produzione. Nei modelli connessionisti, al contrario, il parallelismo e la non-linearità, con

la possibilità di ottenere situazioni complesse e spesso imprevedibili, sono alla base stessa del formalismo, nello spirito di una ispirazione biologica che era già nel primo storico modello di rete neurale formale ideato da McCulloch e Pitts nel 1943. Mentre ogni modello simbolico necessita di una serie di assunzioni che ne giustifichino di volta in volta le scelte nella rappresentazione, nei modelli neurali esiste una variabile che potremmo dire, in tutti i sensi, naturale, nella valutazione e validazione del modello, ossia l'efficacia sinaptica, che caratterizza la *flessibilità* della rete. Queste caratteristiche, inoltre, rendono le RN piuttosto stabili ed efficienti rispetto a dati incerti, rumore, errori, laddove invece i sistemi simbolici si mostrano molto delicati e sensibilmente critici a input *anomali*.

Anche in questo caso, la medaglia ha un rovescio non indifferente. Alla potenza del connessionismo nel trattare efficacemente situazioni di attivazione associativa tra pattern neurali tipica delle relazioni organismo-ambiente fa fronte, nei modelli neurali, il difetto di quella capacità di *astrazione* che è invece propria dei sistemi simbolici, capacità che permetterebbe di rendere conto non soltanto dell'esistenza di un livello simbolico-rappresentazionale del pensiero, ma anche delle tipiche proprietà di manipolazione di questo, come ad esempio l'*iteratività* e la *ricorsività* (ossia la possibilità di costruire espressioni simboliche nuove a partire dalle primitive), l'*ereditarietà* (cioè la caratterizzazione semantica in termini di proprietà ben definite di un concetto e dunque dei suoi derivati), e infine la possibilità di una *strutturazione gerarchica* dei processi cognitivi.

Ad esempio, una rete neurale è estremamente veloce nel riconoscimento di un pattern in ingresso, ma riconoscimento e apprendimento sono due processi ben distinti. Mentre il primo richiede un apparato sensoriale in grado di agire velocemente, l'apprendimento implica la capacità di fornire descrizioni strutturali del pattern acquisito, identificare in esso relazioni, e tutta una serie di abilità che potremmo chiamare di *secondo livello* rispetto al semplice riconoscimento, e che non possono essere ottenute semplicemente a partire dalle caratteristiche fisiche che legano i processi di adattamento rapido tra input in ingresso e formazione di un pattern neurale. In altre parole, così come sono trattate solitamente, le RN mancano di *spessore*, hanno difficoltà a mostrare livelli e stratificazioni, e dunque sono incapaci di supportare abilità autenticamente concettuali, compresa la dimensione semantica, che può vedersi, in fondo, come l'espressione dell'importanza e della priorità di un'informazione e del suo valore d'uso rispetto ad altre in un certo contesto di relazioni.

Questi limiti hanno portato negli ultimi anni a un crescente lavoro di *complessificazione* dei modelli neurali. Le ricerche più interessanti vanno in direzione della definizione di gruppi neuronali *specializzati*, in accordo con le idee di G. Edelman e altri sulla formazione e la fisiologia del sistema nervoso. In questo modo, la rete nella sua globalità rinuncia a un certo grado di efficienza per guadagnare in versatilità. In termini fisici, si tratta di un sottile problema di bilancio tra energia ed entropia parziali in un multi-sistema, che aspetta ancora una piena soluzione. Questi nuovi modelli di RN modulari hanno permesso pure un certo successo nello studio di alcuni disturbi o deficit del comportamento e del sistema nervoso (*patologizzazione del modello*).

Le differenze tra i modelli cognitivisti e quelli connessionisti hanno dunque messo in ombra i numerosi aspetti che li accomunano. Entrambi sostengono una posizione *funzionalista* degli stati mentali, anche se in modo leggermente diverso. Com'è noto sin dai lavori fondamentali di D. M. Armstrong, il funzionalismo sostiene l'indipendenza dei processi mentali dal loro *supporto materiale*, e li considera come *sequenze di stati governate da leggi che ne controllano l'ordine* (H. Putnam). Nel caso dei sistemi simbolici del cognitivismo, il funzionalismo è la naturale conseguenza delle relazioni che ogni simbolo ha con gli altri a un dato livello d'astrazione fissato dalla struttura generale del sistema. E' chiaro che una posizione del genere è all'origine del programma della vecchia IA *forte*, perché se i processi cognitivi possono essere studiati a un livello formale che è sostanzialmente autonomo rispetto ad ogni implementazione possibile, allora il discorso unifica sistemi naturali e artificiali.

Può sembrare meno ovvio che anche i modelli connessionisti rientrano nell'alveo della concezione funzionalista, poiché si sostiene spesso che in un modello neuronale gli stati mentali sono sostanzialmente la geometria e la dinamica spazio-temporale della rete stessa, ed è proprio quest'aspetto che ne giustificherebbe una maggiore *plausibilità biologica* rispetto ai modelli simbolici. In realtà, sostituendo nella proposizione originale di Putnam al termine *stato* il termine *pattern*, si può notare come un modello connessionista è tanto *asettico* e *astratto* quanto un modello cognitivista, ed è per questo motivo che in fondo è possibile rivolgere molte critiche analoghe a entrambi le classi di modelli. Non è un caso se le applicazioni autenticamente 'biologiche' delle RN sono piuttosto esigue se confrontate al gran numero di applicazioni tecnologiche dove vengono utilizzate come potenti dispositivi di calcolo parallelo. Queste

caratteristiche appaiono ancora più evidenti se si considera la complessità di struttura e di funzionamento del cervello come ci viene rivelato dalle neuroscienze. Una sinapsi *reale* è un processo assai più complesso, basti pensare al ruolo dei modulatori, e l'assemblaggio sinaptico è un classico processo fenotipico, dove un insieme di caratteristiche vengono fissate una volta per tutte sulla base del rapporto organismo-ambiente. Bisogna poi considerare anche un ricchissimo bagaglio di caratteristiche elettrochimiche intrinseche delle cellule nervose, che dipende in gran parte dalla componente genetica. Inoltre il sistema nervoso è in relazione profonda con il sistema immunitario ed endocrino: la cognizione è un fenomeno *incarnato e ambientale*! La chiave della ricerca sul Sé e sulla *coscienza* risiede in ultima analisi proprio in queste caratteristiche uniche di ogni processo fisico e mentale, ed è solo una *cattiva coscienza* metodologica che porta a chiedersi se, e in quale misura, un modello teorico può *descrivere* la coscienza. In definitiva, nessuno chiede alle neuroscienze di spiegare perché una lesione cerebrale in una stessa zona produce effetti differenti in soggetti diversi, ed è altamente improbabile che si possa fornire un tipo di spiegazione di questo genere!

La coscienza pone un problema analogo: sicuramente è possibile costruire un modello scientifico della coscienza come *costrutto della scienza cognitiva*, e vedremo che questo è strettamente connesso ai processi di amplificazione e dissipazione dell'informazione, ma la *coscienza fenomenologica* è piuttosto lontana dalle caratteristiche richieste a un fenomeno per essere oggetto di indagine scientifica, pur essendo una legittima e interessante questione speculativa.

Dobbiamo allora riconoscere che la *plausibilità biologica* di entrambe le classi di modelli è un fatto relativo, legato più che altro al tipo di utilizzazione che abbiamo in mente per quel modello. Possiamo, al più, parlare di modelli *bio-inspired*, e fatto ciò non curarci troppo a lungo del problema, o addirittura considerarlo tale. Il costruttore di modelli matematici di sistemi cognitivi non deve in alcun modo sentirsi vincolato da un progetto naturale che l'evoluzione ha messo a punto in milioni di anni. L'obiettivo più realistico che ci si può porre oggi è la costruzione di una teoria generale di tipo formale in grado di dialogare con le scienze del cervello e con la progettazione di sistemi artificiali, e capace di fornire un quadro di riferimento per potersi orientare sull'uso di modelli parziali e/o ibridi, sulle relazioni tra abilità senso-motorie e competenza concettuale, sull'economia neurale globale e modulare.

Il problema è ben noto in fisica teorica, dove i modelli possono essere inquadrati in due grandi classi, *macroscopici* e *microscopici*. I primi fanno uso di un ristretto numero di variabili di stato e si occupano di una descrizione ‘globale’ del sistema, mentre i secondi sono principalmente rivolti alle proprietà dei suoi costituenti. In questo caso il vero problema è stabilire una connessione tra i due livelli di descrizione, proprio com’è accaduto tra termodinamica e fisica statistica, interpretando l’entropia come informazione microscopica. La questione della validità dei modelli simbolici e sub-simbolici non può essere posta in termini di confronto *naive* con i dati neurofisiologici, ma – come suggerisce la comune vocazione funzionalista – bisogna cercare un panorama unitario tra i due livelli descrittivi basato sul concetto di informazione e sulla classe di complessità che il modello esibisce. Come vedremo, questo può essere fatto nell’ambito della teoria dei sistemi logicamente aperti e dell’emergenza, e può trovare un formalismo unitario utilizzando i concetti della teoria quantistica dei campi.

2 - SST e PDP tra Simboli e Sistemi Dinamici

Una descrizione generale dei sistemi dell’IA classica e della scienza cognitiva può essere fatta utilizzando la nozione di *Sistema Semanticamente Trasparente* (SST), un sistema strutturato di dati e regole che può essere specificato completamente in termini di relazioni sintattiche tra le sue componenti. Se ricordiamo la definizione classica di informazione (C. Shannon, 1948), possiamo concludere che ogni SST è una struttura macroinformativa, costituita cioè da un insieme di sottosistemi della stessa natura – si pensi ai *frames* di M. Minsky – in grado a sua volta di produrre informazione in risposta a una determinata sequenza di input. Quando specifichiamo *sottosistemi della stessa natura* intendiamo qui fare riferimento a una proprietà degli SST indicata spesso come *proiettabilità*: si intende che ogni struttura del tipo SST può sempre essere analizzata fino al più minuto elemento senza perdere mai le caratteristiche semantiche ereditate dal dominio di partenza. Essendo gli SST sempre riconducibili a un’analisi informazionale, si vuole indicare che il gioco delle relazioni sintattiche tra i vari elementi viene rispettato a ogni livello d’analisi della struttura, e in nessun punto possono apparire relazioni di nuovo tipo. Se guardiamo a questa proprietà in modo critico, possiamo riconoscere in essa le caratteristiche per cui i modelli rappresentazionali-simbolici sono

considerati *rigidi*, ossia funzionano bene in un dominio ben definito. Il punto essenziale da tenere a mente è che la possibilità di analisi unitaria degli SST non deriva dal particolare sistema scelto o dalla sintassi adottata, ma dal concetto classico di informazione.

I modelli connessionistici, invece, rientrano nell'ampia classe dei *Sistemi PDP* (*Parallel Distributed Processing*), e sono caratterizzati da un gran numero di unità che si attivano in parallelo secondo il paradigma sub-simbolico. Le RN sono sicuramente i rappresentanti più fecondi e potenti della classe PDP, e su questi concentreremo la nostra attenzione, benché, a rigore, tra i sistemi PDP rientrino anche sistemi ibridi, come quelli utilizzati nell'*IA distribuita*, e gli *Automi Cellulari*. I sostenitori del connessionismo argomentano che i caratteristici *schemi emergenti* dei sistemi PDP *rompono* la trasparenza semantica, e tutti i vantaggi di questo tipo di modelli nascono proprio dal *non* avere la proprietà della proiettabilità. A ben guardare, l'emergenza e la proiettabilità vanno proprio in direzioni diverse; la proiettabilità è una caratteristica *top-down*, mentre le proprietà emergenti sono tipicamente *bottom-up*. Scrive Smolensky:

Nel paradigma sub-simbolico (...) le entità interpretabili semanticamente sono pattern di attivazione di un grande numero di unità del sistema, mentre le entità manipolate da regole formali sono attivazioni individuali delle cellule della rete. Le regole prendono la forma di regole dinamiche di attivazione, completamente differenti dalle regole di manipolazione di simboli (Smolensky, 1987).

Questa caratteristica di *opacità semantica* dei PDP è stata chiamata da A. Clark *Olismo Informativo*, riferendosi alla capacità di un modello connessionista di elaborare anche dati incerti o incompleti in domini non ben specificati. Si è discusso molto delle relazioni tra queste due classi di modelli, e per un certo periodo i connessionisti, forti anche dell'apparato concettuale e matematico che accomuna PDP e sistemi dinamici non-lineari, hanno proposto di identificare i simboli dei sistemi ST come *approssimazioni* dell'attività PDP, che si situerebbe dunque a un livello in qualche modo più *fondamentale*. Esistono ormai parecchi lavori teorici e sperimentali che vengono portati a sostegno di questo tipo di relazione approssimata tra PDP e SST basata sui sistemi dinamici (S. Kelso, R. Beer, J. Gallagher, R. Port, T. Van Gelder, L. Smith, E. Thelen).

Gran parte della ricerca sperimentale è basata sullo studio dell'attività

neuronale in presenza di stimoli periodici, quasi-periodici e aleatori, utilizzando tracciati EEG. Riportando i dati in modo opportuno in uno spazio delle fasi, è possibile studiare il comportamento di gruppi neuronali in termini di attrattori. L'analisi dei dati permette di distinguere il mero rumore dalle strutture caotiche tipiche dell'attività neuronica. Questi stati sono estremamente sensibili anche a piccole fluttuazioni, e dunque si può pensare che l'attività neurale caotica rispecchi una sorta di *apertura* altamente sensibile alla possibilità di ricevere nuovi stimoli. Può avvenire che uno stimolo opportuno provochi una transizione di fase inducendo l'emergenza di stato ordinato, periodico o quasi-periodico, una *struttura coerente* che corrisponde alla *cristallizzazione* legata ad un processo di categorizzazione, ossia ad un *simbolo* o *concetto* nel linguaggio dei sistemi ST. Questo è stato dimostrato in modo del tutto generale dai lavori matematici di Matsumoto e Tsuda (1984), e applicato allo studio dei processi cognitivi da S. Grossberg, W. Freeman e I. Tsuda.

Nel modello di Freeman, basato sui suoi studi sul bulbo olfattivo del coniglio, gruppi di unità tendono a manifestare comportamenti oscillatori spontanei di differente ampiezza e frequenza. L'interazione tra questi gruppi finisce per produrre un processo caotico globale che però conserva la memoria delle oscillazioni da cui è stato originato. Quando arriva uno stimolo esterno (nel caso degli esperimenti di Freeman un odore), questo gioca il ruolo di un input che trasforma il comportamento caotico in un processo oscillatorio ordinato, che è l'*etichetta* della categoria a cui lo stimolo appartiene. In tal modo lo stimolo viene riconosciuto ed entra a far parte della memoria del sistema. Può anche accadere che lo stimolo esterno non venga riconosciuto e modifichi lo stesso l'attività e la forma geometrica degli attrattori; in questo caso l'attività caotica che viene a configurarsi individua uno stimolo attentivo.

Per quanto l'idea di interpretare i costrutti simbolici come componenti *frozen* dell'attività neurale abbia una sua plausibile attrattiva, di per sé non risolve il problema di costruire un solido ponte concettuale tra le due classi di modelli, e un'analisi superficiale mostra che una posizione di questo tipo corrisponde, in fondo, a un nuovo tipo di *riduzionismo neurale*. La seduzione di quest'immagine, però, può portare qualcuno a chiedere se davvero è necessario costruire questa correlazione, e se non è più facile, a questo punto, intendere definitivamente i costrutti simbolici come attivazioni più o meno transitorie di un insieme di processi PDP. Ogni forma di riduzionismo, in assenza di vincoli (condizioni limite), non

permette di selezionare od ottenere la descrizione macroscopica di un sistema, e della sua struttura interna, a partire dal comportamento e dalle proprietà delle unità costituenti.

Nel caso della scienza cognitiva, è impossibile, allo stato attuale delle cose, poter utilizzare un'idea come quella delle componenti *congelate* per selezionare un costrutto simbolico rispetto ad altre architetture possibili. Si noti che qui non si tratta di difendere l'ampia modellistica della psicologia cognitiva, che pure ha un bagaglio prezioso di verifiche sperimentali, quanto di riconoscere che essa poggia sul fatto che noi utilizziamo costrutti simbolici, e ci affidiamo ad essi, per scambiare informazioni con altre menti, che risiedono in cervelli la cui struttura è soltanto grossolanamente simile alla nostra! Esiste dunque un *livello di secondo ordine* che *deve* trovare un posto preciso in una scienza cognitiva completa e coerente.

Da un punto di vista *pratico*, poi, l'idea di risolvere la correlazione tramite la teoria dei sistemi dinamici incontra delle difficoltà concettuali e matematiche tutt'ora irrisolte. Poichè si tratta di un problema di carattere 'tecnico' che potrebbe trovare valide soluzioni in un futuro non troppo remoto, non ci soffermeremo qui sul fatto che l'analisi di sistemi collettivi non-lineari a molte variabili, dunque ad elevata dimensionalità, è piuttosto difficile da fare, e soltanto pochi casi particolari, dove è possibile identificare dei parametri d'ordine, offrono una reale possibilità di successo. Più importante, per i nostri obiettivi, è far notare che la scelta delle variabili collettive nello studio di un sistema dipende da quello che del sistema vogliamo sapere, ed è dunque un classico problema della cosiddetta *cibernetica del secondo ordine*, dove il ruolo dell'osservatore è parte integrante dell'analisi del sistema. Cade così in parte l'immagine ingenua di un'analisi delle micro-dinamiche neurali che porta a fissare in modo univoco i livelli di secondo ordine. Tanto più che nello studio dei sistemi dinamici è spesso necessario fare ricorso a un set di variabili che possono essere sensibilmente lontane dalla struttura del sistema, sia in senso morfologico che funzionale. Nella gran parte dei casi, queste variabili sono connesse ad aspetti significativi dell'ambiente esterno. Scrive van Gelder:

...nella sua pura forma, la spiegazione dinamica non fa riferimento all'effettiva struttura del meccanismo di cui si sta spiegando il comportamento. Essa ci dice come evolvono nel tempo i valori dei parametri del sistema, non cosa ne è del modo in cui il sistema stesso è costituito, che fa sì che quei

parametri evolvano secondo quella modalità specificata. Si tratta di esplorare la struttura topologica delle dinamiche del sistema, ma questa è una struttura interamente differente da quella del sistema stesso (van Gelder, 1991).

Anche le neuroscienze cognitive, pur dando un'importanza fondamentale al livello *fine* dei comportamenti neurali, non possono fare a meno di un livello di descrizione sistemico, volto a individuare mappe di organizzazione interna di tipo funzionale, come aree di attivazione di risorse, in grado di spiegare l'occorrenza delle strutture di soccorso che si attivano dopo un danno cerebrale e identificare circuiti integrativi di alto livello. E queste descrizioni utilizzano un linguaggio di tipo rappresentazionale e modulare. E' necessario dunque integrare i vari stili d'indagine e i diversi livelli di descrizione.

La vera vocazione dei sistemi dinamici nello studio dei processi cognitivi è quella di individuare, in situazioni ben precise, tramite la scelta di opportune variabili, il tipo di relazione profonda tra mente e ambiente, rivelando così che la cognizione è un *processo esteso*, dove può essere realmente difficile distinguere cosa appartiene ad una e cosa all'altro. L'atto cognitivo appartiene sempre alla *mente-nel-mondo*. Del resto, è proprio l'esistenza di cervelli *reali*, con neuroni *reali*, che ci richiede una descrizione di ciò che in questo processo appartiene alla mente e al suo funzionamento. Il vantaggio di una scelta metodologica non-esclusiva è quello di ottenere da ogni livello descrittivo un vincolo sugli altri. Ad esempio, ogni spiegazione basata sulla scelta di variabili collettive globali legate alla relazione organismo-ambiente dovrebbe poi sempre poter essere specificata a livello di micro-dinamiche neuronali, e queste, data la loro complessità, richiedono una sorta di *rappresentazionalismo minimo*, più o meno 'trasparente', per una comprensione sistemica del loro ruolo funzionale.

3 - Sistemi, Informazione e Dissipazione

Per un'analisi comparata dei sistemi ST e PDP è molto più utile rinunciare ad ogni forma di riduzionismo ontologico, e concentrarsi invece sul concetto di informazione e sulle caratteristiche computazionali. Bisogna qui chiarire che molte delle discussioni sui SST sono state rese confuse da una certa ambiguità proprio nell'uso del termine *computazione*. Ci sono

almeno tre livelli d'analisi di un sistema simbolico (come il *General Problem Solver* di Newell e Simon e le sue numerose varianti):

- (a) Il livello che definiremo *a scatole e frecce*, diagrammi utilizzati per operare un'analisi dei *flussi informativi* in entrata, in uscita ed interni al sistema; questo livello corrisponde a quello che D. Marr aveva definito propriamente *computazionale*, e di primo livello, visto che è a questo punto che si progetta lo schema logico del sistema.
- (b) Il livello della *rappresentazione e del processo*, in cui si sceglie la strategia per realizzare la struttura logica definita al punto (a). E' questo il livello in cui, in sostanza, si scelgono gli algoritmi e le strutture dati più efficaci.
- (c) Il livello dell'*hardware*, quello che realizza fisicamente la computazione, e che nella teoria SST è ben distinto dal resto, proprio in virtù delle sue caratteristiche 'astratte' che postulano un *mentalese*, descritto logicamente al punto (a), e algoritmicamente al punto (b).

Il punto (a) può sollevare qualche perplessità terminologica, perché porta a identificare il livello logico *macroscopico* della struttura del sistema con quello computazionale, che invece dovrebbe situarsi, come vedremo, a un livello molto più profondo, se dobbiamo dare al termine *computazione* il suo significato formale, sulla linea dell'analisi classica della *Turing-computabilità*. E' possibile ideare sistemi ST con macrostrutture logiche assai diverse tra loro, ma uguali per quel che concerne le caratteristiche del singolo *passo* di computazione. Il punto (b) è sicuramente quello che ha creato più confusione, poiché ha portato a identificare i SST con l'architettura seriale della *Macchina di von Neumann*. Questo equivale a dire che la differenza più rilevante tra SST e PDP è la serialità. E' sicuramente vero che le operazioni tipiche di un sistema simbolico (ricordiamo: iterazione, ricorsività, ereditarietà, strutturazione gerarchica), sono essenzialmente seriali, ma questo non dovrebbe essere considerato un tratto caratteristico dell'intero SST.

E' possibile ipotizzare sistemi simbolici che non debbano passare per il *collo di bottiglia* di von Neumann. Inoltre, un sistema di computazione parallelo può essere simulato per mezzo di una macchina seriale (esempio tipico, con una *LISP Machine*), a patto di non considerare rilevante la velocità di elaborazione (che è comunque una variabile rilevante, se consideriamo che molte prestazioni cognitive devono essere valutate come risposta ambientale efficace). Quello che è davvero significativo nei sistemi

PDP, e costituisce la differenza, è piuttosto l'aspetto cooperativo, più che il mero parallelismo, ossia la non-linearità, e il tipo di complessità che entra in gioco quando consideriamo l'attività di una rete neurale rispetto a un SST. In una *società della mente* come quella ipotizzata da M. Minsky, l'aspetto cooperativo ed emergente dell'interazione tra agenti è spesso evocato come sorgente di comportamenti adattativi, flessibili, capaci di cogliere ed elaborare gradazioni di significato attraverso rappresentazioni distribuite e memorizzazioni superposizionali, ma non viene mai indicato un modello effettivo per realizzare tutto questo senza l'adozione di un genuino *stile* PDP. Per questo è piuttosto difficile sostenere, come qualcuno ha tentato, che i PDP sono *soltanto* un modo diverso di trattare strutture come i *frame* o gli *script* di R. Schank. Il punto non è tanto *serialità vs parallelismo*, ma piuttosto risiede nella linearità dei sistemi seriali rispetto alla non-linearità che può essere facilmente esibita dai PDP; il connessionismo è utile proprio perché introduce un modo *realmente* nuovo di *produrre* informazione.

Altre osservazioni interessanti sono poste dal punto (c), quello relativo all'implementazione hardware. Nei sistemi connessionisti questo livello non può essere distinto in modo netto da quelli relativi al *software*; per essere precisi, il livello del *programma* in una rete è tutt'uno con le condizioni iniziali strutturali della rete, e da ciò deriva l'*opacità semantica*. Questo fatto ha conseguenze *fisiche* e *concettuali* rilevanti. Nei sistemi simbolici la distinzione che si può operare tra *software* e *hardware* è un'eredità del *Cartesian cut* tra mente e corpo, il ché da un punto di vista fisico significa relegare ogni tipo di *processo dissipativo* al livello della *ferraglia*.

La relazione tra informazione ed entropia ci suggerisce immediatamente che questa caratteristica deve avere un'importante contropartita nel *modo* in cui l'informazione viene processata negli SST. Se assumiamo valida la fisica classica (*statistica di Maxwell-Boltzmann*) come scenario teorico di riferimento, possiamo individuare due tipi di sistemi elaboratori di informazione:

- (1) *Sistemi conservatori di informazione*, caratterizzati dalla conservazione dell'energia e descritti dall'invarianza dell'hamiltoniana, e dunque privi di dissipazione.
- (2) *Sistemi dissipatori di informazione*, dove l'entropia aumenta e l'informazione si *disperde* tra i livelli energetici del sistema, fissando un *destino asintotico* prevedibile come attrattore a punto fisso (a meno del rumore di fondo).

Non soltanto negli SST il concetto di informazione viene *sradicato* dalla sua naturale origine fisica, ma l'assenza di dissipazione ci fornisce preziose informazioni sul tipo di processo elaborativo nei sistemi simbolici. Essi, propriamente, non producono nuova informazione, nel senso – qui intuitivo – di sequenze sintattiche imprevedibili a partire dall'analisi della struttura del sistema.

Benché nei modelli più raffinati l'intero *meccanismo* può assumere aspetti piuttosto complicati, il lavoro che fanno gli SST è più simile ad una *riorganizzazione* dell'assetto formale in risposta ad un certo ingresso, in modo analogo a ciò che fa un programma di scacchi quando gioca con un giocatore umano. È questo il motivo per cui, se esistono due vie per il matto, una in due mosse, e un'altra in più mosse con sacrificio, il giocatore umano tende a scegliere quest'ultima, perché, in un certo senso, più *creativa* e *spettacolare*, mentre il programma sceglie la prima: non c'è alcun motivo formale per giustificare la via più *lunga*.² Ma è comprensibile invece che un sistema cognitivo naturale, *tarato* sull'adattamento, ritenga in qualche modo più *intelligente* la risposta meno *comprimibile algebricamente*, proprio per il suo alto *grado di imprevedibilità*, che è, di fatto, un possibile vantaggio dal punto di vista evolutivo e biologico.

Tutto ciò suggerisce che la dissipazione *intrinseca* alla struttura del sistema computazionale gioca un ruolo chiave nella produzione di nuova informazione. Naturalmente, questa condizione da sola è necessaria ma non sufficiente; la dissipazione infatti evita ogni forma di *rigidità*, ma di per sé non può garantire alcuna *novità*, se non quella legata alla casualità e al rumore. Per avere un sistema amplificatore di informazione è necessario introdurre un nuovo elemento, la *non-linearità*. Un'analisi sistemico-informativa individua due classi fondamentali di amplificatori, entrambe individuate dalle soluzioni dell'equazione:

$$(3.1) \quad \frac{dI}{dt} = \frac{1}{V} \frac{dV}{dt}$$

dove I è l'informazione processata dal sistema, t il tempo e V il volume dello spazio delle fasi. L'equazione (3.1) è basata sul *Teorema di Shaw* che connette l'informazione prodotta al volume nello spazio delle fasi. Questo teorema

² Devo questa osservazione e l'analisi di situazioni di questo tipo al giocatore e giornalista di scacchi Yuri Garrett.

è nato nell'ambito della teoria dei sistemi dinamici, ma i suoi risultati sono del tutto generali. Anche nei casi in cui può essere effettivamente difficile scegliere delle variabili per mappare sperimentalmente lo spazio delle fasi di un sistema, questo è teoricamente sempre possibile. In virtù della (3.1) otteniamo due gruppi di sistemi:

- (a) *Sistemi asintoticamente prevedibili*: sono sistemi in cui l'informazione cresce con il tempo, poi va progressivamente rallentando, secondo una funzione polinomiale di t , e dopo un regime transitorio il sistema si *assesta* in una struttura ben definita dal punto di vista spazio-temporale. A questa categoria appartengono i *sistemi dissipativi*, come quelli studiati da Turing (modello della morfogenesi) e dalla scuola di I. Prigogine (sistemi di reazione-diffusione), da Lotka-Volterra (modello preda-predatore), etc. Si tratta di sistemi auto-oscillatori che evolvono verso attrattori tipo ciclo-limite. Per questi sistemi è sempre possibile individuare un parametro d'ordine che pilota il sistema verso una situazione di auto-organizzazione, seguendo un'*equazione del tipo Ginzburg-Landau*; questo è uno scenario universale, ossia indipendente dalle caratteristiche peculiari del sistema, in modo analogo a quanto avviene nelle transizioni di fase. E' possibile dimostrare che per questi sistemi esiste un *valore critico del parametro d'ordine*, e per valori superiori a quello critico il sistema comincia a presentare brusche instabilità. Si tratta dunque di sistemi in grado di mostrare una certa capacità di comportamenti organizzativi, ma piuttosto limitata nella varietà. Strumenti utili per l'analisi di questi sistemi sono quelli della *Sinergetica* di H. Haken, mentre si possono ottenere dei ritratti topologici tramite la *Teoria delle Catastrofi* di R. Thom.
- (b) *Amplificatori esponenziali*: si tratta di sistemi in cui sia I che dI/dt crescono con il tempo, e il volume di fase e l'informazione sono funzioni esponenziali del tempo. Il *teorema di Smale* (1966) ci fornisce le condizioni più generali per avere un sistema di questa classe: la non linearità e che il numero di variabili di stato n sia $n \geq 2$. Questi sistemi sono *strutturalmente instabili*, e appartengono alla classe dei sistemi caotici deterministici. Com'è noto, questi sistemi sono caratterizzati da *attrattori strani*, e sono estremamente sensibili alle condizioni iniziali e al contorno. Diversamente dai sistemi della classe (a), che tendono a *resistere* piuttosto bene agli stimoli esterni, negli amplificatori esponenziali ogni fluttuazione esterna, anche

piccola, può innescare comportamenti nuovi e imprevedibili a lungo termine. Proprio da questo tipo di sensibilità ad ogni input deriva il nome di amplificatori. L'intero processo può essere descritto tramite l'*entropia di Sinai-Kolmogorov*, e si può dimostrare che sistemi di questo tipo subiscono una tipica deformazione nello spazio delle fasi: lungo una direzione si ha una contrazione, mentre le altre mostrano un forte incremento. Intuitivamente, è possibile dire che il processo di dissipazione è parte integrante della *costruzione* di nuova informazione.

I sistemi PDP hanno dunque tutte le proprietà richieste per produrre informazione imprevedibile, non riconducibile all'analisi della struttura del sistema e delle unità componenti, e legata alla *dinamica* e alla *termodinamica* della computazione a un livello profondo, non banalmente connesso con i dettagli dell'implementazione. Com'è noto, le RN rispecchiamo pienamente questo insieme di caratteristiche. E' possibile infatti definire una *funzione di Lyapunov*, legata all'energia del sistema, e utilizzare, sotto opportune condizioni, il *rilassamento* della rete verso un certo attrattore per risolvere problemi di ottimizzazione. Considerazioni del tipo di quelle da noi presentate, hanno spinto alcuni studiosi, come A. Clark e B. Smith, a proporre un'ipotesi detta *congettura della classe di equivalenza di algoritmi*. Si sostiene, in pratica, che ogni sistema SST deve essere compatibile con *almeno* un modello PDP, il che equivale a dire che l'esigenza di un *rappresentazionalismo minimo* deve fare i conti con le proprietà dinamiche di un sistema PDP. La differenza, rispetto a un'ingenua riduzione dei SST ai PDP, consiste nell'idea di *classe di equivalenza*, e non esclude dunque che un sistema SST possa essere *confrontabile* in capacità e prestazioni con un PDP, pur utilizzando un'architettura radicalmente diversa.

4 - Tesi di Church-Turing, Computazione e Continuità

Il problema posto da congetture come quella della classe di equivalenza è legato alla possibilità di confrontare SST e PDP in relazione a macro-strutture di informazione. L'esame sistemico proposto suggerisce che un SST è un sistema conservatore d'informazione essenzialmente lineare, mentre un PDP nella sua forma più generale dev'essere un sistema amplificatore esponenziale non-lineare. Nulla vieta di pensare di poter ottenere un sistema lineare come limite, sotto opportune condizioni, di

un sistema non lineare; ma la modellistica attuale, nella sua varietà, non permette di dare un significato univoco all'idea espressa nella congettura delle classi di equivalenza, che si deve limitare a indicare un'esigenza più che un *set* rigoroso di vincoli.

Spostiamo adesso la nostra analisi informazionale al livello *microscopico* della computazione. Qui troviamo nuovi problemi, molto più *sottili*, legati a un dibattito matematico profondo, quello dell'alternanza, o complementarietà, tra *continuità* e *discreto*. Il legame storico tra lo sviluppo dei metodi classici in IA e i sistemi di elaborazione dell'informazione rende naturale l'identificazione tra SST e *sistemi digitali*. Questa identificazione ha radici antiche, che partono dal *programma formalista e assiomatico* di Boole, De Morgan, Frege, Peano, Russell e Whitehead ed Hilbert (1850-1930 circa), e arriva all'implementazione delle funzioni del *calcolo dei predicati del primo ordine* mediante la costruzione di unità di commutazione in grado di descrivere l'insieme delle *funzioni dell'algebra booleana*, ad opera di C. S. Peirce, C. Shannon, McCulloch e Pitts (il lavoro pionieristico di Peirce è dei primi anni del '900; il lavoro di Shannon del 1937, quello di McCulloch e Pitts del 1943). Dal punto di vista formale, ed indipendentemente dai suoi successi ed insuccessi *pratici*, il programma della prima IA si configura come la possibilità di poter effettivamente realizzare il programma di ridurre l'intera matematica alla logica, secondo il celebre *incipit* de *I Principi della Matematica* di Bertrand Russell:

La matematica pura è l'insieme di tutte le proposizioni della forma 'p implica q', dove p e q sono proposizioni che contengono uno o più variabili, né p né q contenendo costanti che non siano costanti logiche. Le costanti logiche sono concetti che si possono definire in funzione di: implicazione, relazione di un termine a una classe di cui è membro, nozione di 'tale che', nozione di relazione, ed ogni altro concetto implicito nella nozione generale delle proposizioni della forma precedente. Oltre a questi, la matematica 'usa' un concetto che non fa parte delle proposizioni che essa considera, vale a dire la nozione di verità (B. Russell, 1903, Cap. I, § 1).

La prima proposizione del trattato di Russell può leggersi come l'affermazione che *l'intera matematica è un sistema ST*, completamente descrivibile tramite sequenze sintattiche di relazioni a-semantiche. I risultati di K. Gödel e A. Turing durante gli anni '30 misero in crisi profonda questo progetto, per cui si potrebbe sostenere che esso nacque,

in un certo senso, già *vecchio*! Ma la cosa interessante da notare è che sia Gödel che Turing si mostrarono due sostenitori – moderato e possibilista il primo, appassionato e provocatorio il secondo – dell’idea che le limitazioni dei sistemi assiomatici poco hanno a che fare con l’idea centrale dell’IA, ed entrambi ritenevano plausibile considerare la mente alla stregua di un *automa finito* con un numero di stati molto grande (e per Gödel, *divergente nel tempo*)³. Per comprendere le relazioni tra questi risultati sui fondamenti della matematica e trarne indicazioni concrete sulle caratteristiche dei modelli matematici dei processi cognitivi, è necessario esaminare con maggior dettaglio ciò che finora abbiamo utilizzato in modo intuitivo, ossia la nozione di *computazione*, cosa che richiede la nozione di *Turing-computabilità*.

Nel suo fondamentale articolo *On Computable Numbers* (1936), Turing propone una *definizione operativa* di *computabilità*, astraendo progressivamente il ragionamento a partire dall’analisi apparentemente *naive* delle operazioni che un essere umano effettivamente compie calcolando, ad esempio, il valore di una funzione. Questa capacità sorprendente di passare da un livello concreto, a volte ingegneristico, a uno più speculativo e matematico, era una caratteristica del genio di Alan Mathison Turing. Le caratteristiche generali del lavoro di Turing sono piuttosto note. Per i nostri scopi sarà qui utile limitarci soltanto ad alcuni aspetti che riguardano la natura *atomica* del singolo *passo computazionale*, e le limitazioni fisiche che questo tipo di definizione impone ad ogni sistema computante, sulla linea della generalizzazione proposta da R. Gandy, amico di A. Turing e continuatore della sua opera (Gandy, 1980).

Un *automa finito* è individuato da un certo numero finito di *stati di ingresso*, *stati interni* e *stati d’uscita*, e da una *funzione di transizione* che, fissato uno stato interno e una sequenza in ingresso, stabilisce univocamente lo stato d’uscita. Ogni successione di stati è dunque deterministica. E’ facile comprendere che gli automi finiti possono riconoscere soltanto sequenze regolari finite di simboli (*Teorema di Kleene*). La *Macchina di Turing* (MT) è un dispositivo che supera queste limitazioni grazie a un *serbatoio* illimitato

³ Anche se la storia scientifica li accomuna, è difficile immaginare due persone più diverse di Turing e Gödel. Quest’ultimo tendeva a un platonismo piuttosto radicale, temperato appena dal suo bisogno logico di seguire ogni possibilità offerta da un’ipotesi fino in fondo; Turing invece, sin da giovane, utilizzò spesso in modo irriverente un materialismo *naive*.

di memoria, che tradizionalmente viene descritto come un *nastro infinito* suddiviso in caselle sulle quali è presente un simbolo (o uno spazio vuoto). La Macchina di Turing può leggere un simbolo per volta, cancellarlo e scriverne un altro, spostare il nastro avanti e indietro di una o più caselle. Turing mostra che un dispositivo di questo genere è in grado di svolgere *ogni compito* che tradizionalmente rientra nel concetto di computazione. Infatti per una MT è possibile riconoscere e lavorare su sequenze di input non regolari e di qualsivoglia lunghezza, senza alcun limite per il tipo di operazione da svolgere; lo stato interno può modificarsi e svolgere sempre nuove elaborazioni suggerite ogni volta da una diversa successione di input. Si introduce la nozione di *Macchina di Turing Universale* (MTU) considerando che qualunque MT progettata per un compito specifico può essere descritta e usata come input per un'altra MT e così via, in modo da poter pensare alla MTU come a una *collezione infinita* di automi. Il prezzo da pagare per questo *automa computazionalmente universale* è piuttosto alto, poiché è facile rendersi conto che il processo di calcolo così definito fa astrazione da ogni aspetto che riguarda lo spazio e il tempo della computazione. L'unico *spazio* considerato nell'operare di una MT è quello delle caselle del nastro, e il *tempo* è semplicemente una successione discreta di operazioni elementari, ognuna delle quali definisce un *istante computazionale*. Questa sembrerebbe a prima vista una forte limitazione, soprattutto se si pensa che l'analisi di Turing debba poi potersi applicare a forme *concrete* di calcolo, ma un'analisi più attenta rivela che la limitazione è in realtà una forma di *semplificazione*, in modo da ottenere una *definizione estensionale* di computazione sufficientemente ampia da poter far rientrare in essa una gran varietà di casi particolari. Pur essendo una costruzione matematica, è possibile far riferimento alla fisica per trovare argomentazioni a sostegno delle caratteristiche della MT che ne giustificano l'universalità. In questo modo ci manteniamo fedeli allo *spirito di Turing*, poiché non poniamo alcun limite alle MT, e tantomeno ne forniamo alcun modello fisico, ma ci limitiamo ad asserire che un calcolatore universale, per quanto idealizzato, non deve violare le leggi della fisica. Le condizioni sulla *serialità* e la *località* del procedimento, ossia di poter *leggere/scrivere* un simbolo per volta, possono trovare un sostegno nel *limite di Bekenstein*, che afferma che un sistema fisico in un volume V può contenere un numero finito di *stati distinguibili energeticamente*, legati alla sua entropia. Un sistema fisico reale non potrà certo leggere e manipolare più informazione di quella che può contenere, e perciò l'assunzione di Turing appare più che ragionevole: il

punto essenziale è che l'intero nastro non è disponibile nella sua interezza alla macchina ad ogni singolo *istante computazionale*. Il *nastro infinito*, reso così famoso da innumerevoli esposizioni, va inteso come possibilità ideale, ossia come un modo per utilizzare il concetto di automa anche in un ambito assai più complesso di quello fissato dal Teorema di Kleene. Anche l'essenziale *discretizzazione* del procedimento può essere discussa facendo riferimento al teorema di Bekenstein, che è poi una *filiazione* del *Principio di Indeterminazione di Heisenberg*; questo fondamentale pilastro della fisica esprime le condizioni minime perché il concetto di *stato* non perda di senso in una suddivisione dello spazio delle fasi di un sistema fisico. Se questo avviene è impossibile ogni codifica dello stato in questione in forma simbolica, ed è dunque impossibile definire ed elaborare informazione relativa a quello stato fisico. Alla memoria *infinita* è strettamente connessa l'assenza di una qualunque valutazione del *tempo di computazione*, ma qui l'obiettivo di Turing è quello di suggerirci, *in linea di principio*, che anche un sistema parallelo e asincrono può essere computato tramite una MT. Consideriamo infatti un sistema computazionalmente distribuito nello spazio e nel tempo, come gli *automi cellulari* (AC). In questo caso molte unità si attiveranno o disattiveranno in modo complesso durante la computazione. Una MT equivalente deve poter fare la stessa cosa, e per poterla eseguire deve far scorrere il nastro avanti e indietro per fissare il valore di una unità spazialmente e temporalmente localizzata.

Per il *principio di relatività* dobbiamo vincolare questo procedimento ad un valore della velocità di scorrimento del nastro inferiore alla velocità della luce. Ma la stessa limitazione deve essere imposta all'automa cellulare, e dunque, a parte considerazioni di *efficienza di calcolo* che in questa sede non hanno rilevanza, una MT può eseguire lo stesso compito di un automa cellulare. In definitiva, l'analisi di Turing non ci dice *come* fare una computazione, ma fissa le caratteristiche generali sintattiche di *ogni* processo di calcolo. E' possibile dimostrare che altre definizioni di computazione (*Lambda-calcolo di Church*, *Sistemi di Produzione di Post*, *di Markov*, *Funzioni Ricorsive di Kleene*, *Grammatiche non ristrette di Chomsky*, etc.) sono equivalenti ad una MT. Si arriva così, in modo naturale, alla:

Tesi di Church-Turing: *ogni funzione effettivamente calcolabile mediante un algoritmo è calcolabile con una opportuna Macchina di Turing.*

Ovvero: *ogni algoritmo è Turing-computabile* (Tesi CT).

Per smentire questa tesi sarebbe necessario trovare un esempio di funzione calcolabile, ma non Turing-computabile. Per quanto esista un certo dibattito sulla questione, l'esistenza di una siffatta funzione è piuttosto controversa, e dunque accetteremo la tesi CT vera in tutti i casi che ci occorrerà prendere in considerazione.

L'idea veramente innovativa di Turing fu quella di mostrare come questa costruzione non soltanto permetteva una definizione costruttiva di calcolabilità, ma forniva un'analisi computazionale dei famosi teoremi di *incompletezza* e *indecidibilità* di Gödel. Per i nostri scopi sarà qui sufficiente ricordare l'aspetto informativo di questi risultati utilizzando la nozione di *compressione algoritmica*, secondo l'analisi proposta da R. Solomonoff e G. Chaitin. La richiesta di *completezza* e *coerenza* di un *sistema assiomatico*, centrale per il programma formalista di Hilbert, equivale ad affermare che una certa quantità di informazione – quella contenuta, ad esempio, nell'aritmetica naturale – può sempre essere ridotta a un numero finito di assiomi tra loro non-contraddittori, e da questi ricavata utilizzando un sistema di regole completamente specificate da un linguaggio formale. I teoremi di Gödel pongono forti limitazioni alla piena realizzazione di un simile programma, dimostrando che è sempre possibile derivare dal sistema di assiomi una stringa, perfettamente *in regola* dal punto di vista sintattico, ma *indecidibile*, ossia della quale è impossibile stabilire la *verità* o la *falsità* con gli strumenti del sistema; l'*incompletezza* consiste invece nel dimostrare che un sistema assiomatico non può produrre una quantità di informazione tale da poter dimostrare la propria coerenza (in modo, come suol dirsi, *canonico*, ossia definitivamente).

I due risultati considerati globalmente possono essere espressi dicendo che esiste sempre un *gap informativo* tra la produzione di un linguaggio formale e la sua compressione algoritmica in forma assiomatica. Per esaminare questi risultati dal punto di vista della Turing-computabilità, introduciamo alcuni concetti elementari della *teoria dei linguaggi formali*. Un *linguaggio ricorsivamente enumerabile* è un sistema di produzione formale nel quale ogni espressione è completamente specificabile con una procedura effettiva. Il concetto di *procedura effettiva* ci suggerisce subito che per i linguaggi ricorsivamente enumerabili è possibile costruire una MT in grado di fermarsi in uno stato particolare (*stato di accettazione*) per ogni espressione del linguaggio, mentre per espressioni non appartenenti al linguaggio in oggetto si ferma in uno *stato di rifiuto* oppure *non si ferma mai*. Un *linguaggio ricorsivo*, più semplicemente, è un linguaggio per cui

è possibile costruire una MT che si ferma in accettazione o in rifiuto, a seconda che la stringa in ingresso appartenga o meno al linguaggio studiato. Una MT di questo tipo è dunque in grado, con il suo comportamento, di *decidere* dell'appartenenza di una stringa al linguaggio analizzato o meno. E' possibile dimostrare il fondamentale:

Teorema sui linguaggi ricorsivi: *esistono linguaggi ricorsivamente enumerabili, ma non ricorsivi.*

In altre parole, i due tipi di linguaggi non sono equivalenti. Questo teorema permette di formulare il risultato di Gödel in termini di MT. Consideriamo infatti una MT particolare, che indicheremo con MT_0 , e una stringa S costruita secondo le regole di un linguaggio ricorsivamente enumerabile. Ci poniamo il problema di costruire una MT, che indicheremo con MT_1 , in grado di dirci se la coppia formata da (MT_0, S) si fermerà o meno. Turing dimostra che questo *problema della fermata* (*halting problem*) è *indecidibile*, e che non esiste alcuna MT_1 in grado di risolvere il problema. In pratica, questo fornisce una vivida rappresentazione del teorema di indecidibilità di Gödel, suggerendoci una MT_1 che computa incessantemente, senza fermarsi mai, una sequenza di stringhe a lungo termine imprevedibili, anche se ogni singolo passo può essere descritto partendo dalla descrizione dell'input e dello stato interno. In altre parole, una predicibilità locale e limitata dell'output non implica una predicibilità globale del risultato finale. E' possibile intravedere una sottile analogia tra la crisi del deterministico in fisica e lo scacco del programma di Hilbert. Dal problema della fermata si deriva anche il secondo teorema di Gödel sulla coerenza, che può essere inteso come una forma di indecidibilità *estesa* all'intero sistema assiomatico.

Ci siamo soffermati sui risultati di decidibilità e incompletezza perché sono stati al centro di un dibattito molto confuso e fuorviante, ma che pure ha avuto una certa rilevanza. Si trattava di valutare l'utilità della MT come modello per i sistemi cognitivi, e il ruolo che in questo giocano i limiti di Gödel, con poco riguardo per le cose che i due pensatori avevano già espresso in proposito. Innanzitutto il concetto stesso di *procedura effettiva* dovrebbe suggerire che l'utilità della MT risiede nell'analisi astratta dei procedimenti sintattici di computazione, e non ha nulla in comune con la conoscenza strutturata di cui fanno uso gli organismi biologici. Un sostenitore della tesi meccanicista potrebbe pur sempre ipotizzare una

gerarchia di MT in grado di offrire una descrizione di tutte le attività umane. Naturalmente si tratta di un'ipotesi vaga, che non considera, ad esempio, se questa gerarchia è finita o meno. Se consideriamo però una posizione come quella di F. Tipler (1986), che fa uso sia di una posizione meccanicista che di argomentazioni relative al livello fondamentale di descrizione della teoria quantistica dei campi, è evidente che ogni sistema spazialmente e temporalmente limitato può essere visto come una sorta di automa soggetto ad un numero enorme, ma finito, di transizioni tra stati quantici, selezionati nel corso dell'evoluzione (Turing sosteneva qualcosa di simile, in modo più diretto, quando asseriva che la memoria umana è *evidentemente* limitata). Questo però non ci dice nulla sull'origine e l'organizzazione funzionale di questi stati, e dunque una simile argomentazione ha pochissima utilità nella costruzione di una teoria cognitiva. J.

Von Neumann aveva studiato l'equivalenza tra reti nervose formali del tipo McCulloch-Pitts e gli automi finiti, e aveva concluso che se a livello dei singoli neuroni poteva essere sostenuta una *rozza* analogia tra sistema nervoso e sistemi digitali, questa cadeva proprio al livello della mente, intesa come gestione di risorse. I sistemi digitali sono efficienti e *rigidi*, capaci di andare in *crisi* anche per un singolo stato, mentre la mente umana è meno *precisa* ma più *affidabile*. Questo per von Neumann suggeriva che l'analogia sostenibile per gli aspetti della micro-computazione neurale, da lui chiamata *computazione di primo livello*, non era di alcun aiuto nella comprensione della *computazione di secondo livello*, che consiste proprio nell'attività della mente di organizzare e utilizzare i processi di primo livello. D'altra parte, il filosofo J. R. Lucas (1961) e più recentemente R. Penrose (1992), hanno sostenuto che i teoremi di indecidibilità pongono un limite ben preciso e invalicabile tra i sistemi naturali e l'IA, perché dove le MT si imbattono nell'*halting problem*, un sistema di intelligenza naturale può ricorrere ad una gestione euristica del problema, basandosi su una scelta *gordiana* (mettere da parte il problema o scegliere casualmente una strategia) o semplicemente ricorrendo all'esperienza. Altri hanno controbattuto alla tesi di Lucas-Penrose sostenendo non soltanto che ogni euristica può essere – al solito, *in linea di principio!* – opportunamente descritta da un algoritmo, ma che le situazioni di indecidibilità pongono le basi per una comparazione tra menti e macchine, perché anche la mente umana va in stati di stallo di grande complessità computazionale e imprevedibili. Bisogna osservare che, a posteriori e in linea di principio, un meccanicista può sempre considerare qualsiasi comportamento, una volta debitamente

osservato e registrato, come suscettibile di una simulazione algoritmica, il ch     ben diverso dal *produrlo*, comprenderne la genesi, e il rapporto dinamico con le risorse globali della mente. In ogni caso va ricordato che i teoremi di G  del riguardano le limitazioni dei sistemi assiomatici, e che l'indecidibilit   che la dimostrazione di G  del ricava, e che nelle MT   all'origine del problema della fermata,   qualcosa che poco o nulla ha in comune con l'attivit   della mente di porre problemi e cercare soluzioni! In definitiva, aveva ragione G  del quando, nell'ultima fase della sua vita (Hao Wang, 1984), riassunse l'intera questione in modo molto semplice e pregnante: ammettendo pure che la mente, in qualche *stato originario*, possa essere assimilata a un sistema assiomatico (idea che potrebbe conservare del fascino per certo neo-riduzionismo biologico basato sulla genetica), l'interazione con l'ambiente *complessificher  * il sistema di partenza e lo *modificher   strutturalmente* in modo imprevedibile, al punto da sfuggire alla nostra analisi e conservare comunque una sua peculiare *coerenza*. La capacit   della mente di intuizioni, strategie e *novit  * ha ben poco a che fare con la decidibilit   e la coerenza dei sistemi formali e le loro limitazioni. Anche l'ultimo Turing, durante il suo lavoro sulla morfogenesi⁴ (Turing, 1952), sembrava pi  orientato a considerare l'attivit   neurale come un sistema dinamico complesso e non-lineare. Fu in quel periodo che esprese a R. Gandy la sua convinzione che *le descrizioni debbono essere non-lineari, le previsioni debbono essere lineari*. Quest'idea, come vedremo,   ricca di fecondi sviluppi, e suggerisce che una comprensione dei fenomeni complessi implica necessariamente il ricorso alla non-linearit  , ma l'uso pratico dei modelli, compresa la simulazione, pu  richiedere e giustificare delle drastiche semplificazioni. Un livello, per , non va mai confuso con un altro.

Il vero problema consiste non tanto nell'indagare quale utilit   diretta ha il concetto di MT per lo studio della mente, questione evidentemente mal posta, ma studiare le relazioni tra il concetto di Turing-computabilit  , cos  *naturale* nei sistemi digitali, e il modello PDP pi  potente, le RN non-lineari. L'insieme di tutte le MT, infatti,   un sistema *discreto e numerabile*, mentre le RN sono descritte da sistemi di equazioni differenziali (*EqDiff*),

⁴ A. M. Turing, *The Chemical Basis of Morphogenesis*, in *Phil. Trans. of the Royal Soc.*, 237B, 1952. Si tratta di un lavoro *profetico* sui *processi di reazione-diffusione* studiati poi da I. Prigogine e H. Haken.   interessante notare, in relazione alle riflessioni che presenteremo nell'ultimo paragrafo, che questa potente classe di sistemi dinamici si pu  descrivere in termini del formalismo di seconda quantizzazione.

e utilizzano funzioni continue. Dobbiamo chiederci dunque se un concetto intrinsecamente discreto è utile pure quando si lavora con la *potenza del continuo*.

5 - Computazione Naturale nei Sistemi Fisici e Biologici

Nella sua accezione più generale, *una computazione è una relazione tra oggetti matematici* che associa ad un elemento appartenente a un dominio di input un elemento appartenente a un dominio di output. Questo concetto può essere formalizzato rigorosamente in vari modi, ma qui ci interessa l'idea essenziale. Sia il dominio che il co-dominio possono essere *continui* o *discreti*, o appartenere a casi intermedi. Turing mise a punto la sua definizione di computabilità studiando quello che un operatore umano fa concretamente quando calcola qualcosa. Abbiamo poi visto che le astrazioni richieste da Turing per definire il suo modello di automa universale non soltanto trovano un certo sostegno nella fisica (nel senso che non violano alcun principio noto), ma ci guidano nella costruzione di particolari sistemi fisici che computano, i computer digitali. A questo punto possiamo però domandarci se in natura si possono trovare modelli di computazione diversi, equivalenti in modo non-banale o persino più potenti della MT. Ci sono diverse motivazioni per questa ricerca. Da una parte abbiamo già notato che il concetto di Turing-computabilità (*T-comp*) non ci dice molto sulle *modalità effettive* con cui il nostro cervello *computa*, ed è possibile soltanto proporre un'ingenua analogia tra i processi sinaptici e le transizioni di stato in una MT. In altre parole, la computabilità secondo Turing è una nozione poco rilevante nello studio dei sistemi biologici e cognitivi. Ci sono poi una serie di casi che suggeriscono l'esistenza di numeri non-Turing computabili, ma come vedremo, computabili comunque in un senso più *ampio* e inaspettato. Sono ben noti, ad esempio, i risultati di M. B. Pour-El (1989) e I. Richards (1989) a proposito di un'equazione d'onda tipo-Maxwell, soggetta a evoluzione deterministica, e dunque fissata dalla struttura del campo e dalle condizioni iniziali. Per un particolare insieme di *valori iniziali*, computabili, il campo può assumere *valori non-computabili nel senso di Turing*. La portata del risultato è stata poi *ridimensionata* da una serie di osservazioni sulla scarsa plausibilità fisica del campo utilizzato, e soprattutto sulle condizioni relative ai dati iniziali scelti, molto poco *naturali* nell'ambito di uno *spazio-tempo continuo* come quello in cui si costruisce la teoria di

Maxwell. Un problema analogo è stato individuato in cosmologia quantistica da R. Geroch (1986) e J. B. Hartle (1986), dove la crisi della T-comp nasce da un osservabile dato dalla somma integrale di valori calcolabili su una superficie legata alle condizioni iniziali, però il *set* delle superfici, nel suo complesso, risulta T-incomputabile. Il lavoro di Geroch e Hartle, proprio come quello di Pour-El e Richards, è stato considerato non troppo *drastico*, poiché il problema potrebbe risiedere nel metodo scelto per definire le grandezze significative del sistema, oppure più a fondo nei metodi oggi usati per applicare la teoria quantistica all'universo. Però, anche accettando queste critiche, è possibile individuare nei lavori citati due suggerimenti preziosi, uno di carattere fisico ed uno di tipo epistemologico. Poniamo, ad esempio, che esista in natura un campo come quello descritto da Pour-El e Richards. In ogni punto dello spazio e ad ogni istante del tempo, i valori del campo sono definiti, anche se non T-comp. Questo vuol dire che comunque il campo scambia continuamente informazione con l'ambiente esterno, e dunque *computa*, seppure in un modo che non rientra nella definizione di Turing! Al di là del caso *patologico* analizzato da Pour-El e Richards, nulla ci vieta di fare questo ragionamento per ogni tipo di sistema fisico, anche per quelli di cui è difficile dare una descrizione matematica semplice; ogni sistema fisico, secondo l'analisi sistemico-informazionale, rientra in una delle tre categorie fondamentali, e questo fatto è ben più generale della specifica definizione di computazione adottata. Non c'è alcun motivo per considerare *computante* soltanto un sistema fisico che possiamo analizzare in termini di T-comp, come i calcolatori digitali, ma dobbiamo piuttosto considerare *ogni sistema fisico un elaboratore di informazione*. Studiando i vari sistemi fisici in modo critico, possiamo cercare di rispondere alla domanda se *tutta* l'informazione elaborata in natura può essere soddisfacentemente descritta dall'attuale teoria della computazione. In definitiva l'informazione è una grandezza che può essere messa in relazione diretta con i concetti fisici più radicati, come i livelli energetici, mentre la computazione riguarda semplicemente *un modo* per analizzare sintatticamente l'informazione.

Nel caso del campo di Pour-El e Richards è possibile comunque fare alcune critiche al modello in base all'esperienza. Nel lavoro di Geroch e Hartle, invece, se si vuole evitare la *crisi* della T-comp, è necessario rifiutare il loro modello *in blocco*, in favore di un altro che, dato l'argomento, non può trovare nella dimensione sperimentale maggior conforto, ma può soltanto avere il vantaggio di conservare una definizione *tradizionale* di computabilità. Tutto ciò appare notevolmente arbitrario, perché in

definitiva le teorie fisiche e le loro formulazioni matematiche sono libere costruzioni della mente umana, ed è ipotizzabile proprio in base ad esempi tipo Geroch-Hartle, che la T-comp risulti un vincolo *superfluo* per l'eleganza e la fecondità delle nostre descrizioni della natura. I due casi citati ci suggeriscono *la possibilità che la computabilità (o meno) di un ente dipenda in modo drastico dalle assunzioni della teoria*, ed è dunque molto *naturale* adottare l'idea che una nozione di computazione utile in un certo ambito può esserlo molto meno in un altro. Possiamo anche considerare l'intera faccenda sempre da un punto di vista sistemico, ricordando che ogni descrizione di un sistema fisico sottintende una serie di scelte selettive che definiscono un *osservatore*. Una MT può essere considerata parte della *descrizione formale di un osservatore*, quella appunto che si occupa della numerabilità e calcolabilità degli enti della teoria attraverso un processo ben definito sintatticamente e basato sulla discretizzazione, ma indefinito dal punto di vista della complessità temporale e che in generale nulla dice sui comportamenti spaziali dell'informazione. Cercare e utilizzare una nozione diversa di computazione significa semplicemente definire nuovi punti di vista sullo studio dei sistemi fisici e sulle caratteristiche formali dei nostri modelli.

Ogni processo fisico può essere analizzato a vari livelli di dettaglio, e sotto molteplici prospettive, e non c'è alcuna ragione per ritenere una descrizione continua più o meno fondamentale di una discreta. Proprio nel campo delle neuroscienze, gran parte dei processi misurati ben si prestano ad essere rappresentati con grandezze continue. E' vero che una connessione sinaptica ha degli aspetti discreti, se non proprio *booleani*, ma è strettamente connessa ad altri processi, come la crescita del potenziale d'azione e il processo di rilascio assonico, che sono continui. Lo stesso *fattore di crescita nervoso* può essere visto come un *processo di reazione-diffusione*. In generale, il modello matematico di un sistema fisico definisce una struttura di insiemi scelta in virtù degli obiettivi modellistici centrati sulle caratteristiche osservabili del sistema. In molti casi una *Equadiff* ordinaria o alle *derivate parziali* (*PEquadiff*) può essere l'approssimazione di un discreto più o meno *fine*, tipico di un *modello a differenze finite* (EDF), in altri casi è più conveniente il contrario. Ad esempio, uno dei primi famosi modelli di RN del periodo d'oro, quello di J. Hopfield, era composto da *neuroni binari* che si attivavano in modo sincrono su passi temporali discreti (Hopfield, 1982). Più tardi lo stesso Hopfield con D.Tank (Hopfield, 1984; Hopfield-Tank, 1985, 1986) rese il modello

più *flessibile* utilizzando *neuroni analogici* con output variabili. Questo significava sostituire la funzione di trasferimento *a gradino* del primo modello con altre continue e differenziabili. Ma è possibile dimostrare che se si sceglie la funzione di trasferimento sufficientemente *ripida*, i due modelli sono equivalenti sotto molti punti di vista. Una ragione profonda per questo tipo di scelte può consistere non tanto in qualche aspetto *fondamentale* del sistema fisico, quanto nelle possibilità di semplificare le tecniche matematiche; ad esempio, per quanto a prima vista possa apparire strano, le funzioni continue permettono – per via della loro *regolarità* – di applicare delle considerazioni di simmetria che rendono l'analisi dei sistemi continui molto più facile di quella dei sistemi discreti, dove i vincoli imposti dalla continuità lasciano il posto a configurazioni assai più complicate. In termini leggermente più precisi, *il gruppo delle trasformazioni continue è di un ordine di cardinalità inferiore al numero totale delle trasformazioni definite su insiemi discreti*. Una computazione è una corrispondenza tra il modello e l'insieme delle successioni sintattiche che produce a diversi livelli d'analisi. Data l'arbitrarietà dei livelli in gioco e la varietà modellistica, non c'è alcuna ragione di principio per assegnare ad ogni sistema fisico lo stesso concetto di computabilità.

Queste considerazioni hanno spinto un gran numero di ricercatori a occuparsi delle caratteristiche computazionali dei modelli continui, con l'obiettivo di studiarne le proprietà e le relazioni con il concetto di T-comp. Poiché questo tipo di lavoro si svolge nell'ambito dei *sistemi formali continui* che trova il suo campo d'elezione nei sistemi dinamici e nelle RN, si parla anche di *computazione analogica* o *naturale*, in contrapposizione con l'*artificialità discreta* delle MT.

Negli anni '50 Kleene aveva dimostrato un gruppo di teoremi che pongono le basi per mostrare un'equivalenza tra una qualunque MT e una rete formale di McCulloch-Pitts con *feed-back*. Le RN, dunque, nascono già *molto potenti*, ma una serie recente di risultati rimette in discussione l'intera questione da una diversa prospettiva, mostrando che le RN possono mostrare *capacità computazionali diverse e più potenti* rispetto alle MT. Lavori rilevanti sono quelli di S. Franklin (1990) e M. Garzon (1990), i quali hanno dimostrato che una rete neurale con un numero infinito numerabile di componenti può risolvere il problema della fermata, ossia in qualche modo *decidere* sulla computazione. M. Stannett e altri (Stannett, 1990; Bournez e Cosnard, 1995) hanno esteso questo risultato a classi generali di sistemi dinamici. Altri risultati decisivi sono stati ottenuti da

H.T. Siegelman e E. Sontag (1994; 1995; 1999), che hanno concentrato la loro attenzione alla classe delle *RN ricorrenti* (RNR), reti con un numero finito di neuroni connessi tra loro da un *grafo ciclico*, in modo da assicurare un *gioco di feed-back* tra varie unità, e modificazione *sincrona* di ogni neurone tramite una funzione sigmoidale. I due studiosi hanno esteso in modo originale i risultati classici tipo Kleene, mostrando che se i coefficienti di connessione, ovvero i pesi, della rete sono numeri razionali allora è possibile stabilire un'equivalenza tra una MT ed una RNR. Se i pesi sono invece rappresentati da numeri reali, e le funzioni di attivazione sono non-lineari e continue, è possibile dimostrare che una RNR non soltanto equivale, su tempi di computazione polinomiali, a una MT, ma può essere molto più potente. Non solo, infatti, si dimostra equivalente a Macchine di Turing che consultano i cosiddetti *oracoli* (procedure di decisione esterne aggiunte come vincolo) su tempi polinomiali, ma è anche capace di riconoscere linguaggi non T-comp su tempi esponenziali! Notiamo che in questa analisi il confronto tra RNR e MT non viene svolto soltanto sul piano *sintattico* (capacità di riconoscere o meno linguaggi), ma anche sui *tempi di elaborazione*, com'è naturale aspettarsi da un concetto di computazione che aspiri ad avere caratteristiche di maggiore fedeltà nella descrizione dei processi fisici e biologici. D'altra parte alcuni di questi autori (Maass-Sontag, 1999) hanno anche indagato le difficoltà per una rete neurale soggetta a rumore di riconoscere semplici sequenze di linguaggi regolari.

Tutto ciò porta a concludere che la computazione analogica può avere, a seconda dei casi, capacità sub-Turing o super-Turing, ma comunque qualitativamente diverse. B. MacLennan ha fatto un'analisi generale delle caratteristiche tipiche della computazione naturale (MacLennan, 2001), definendo i *campi computazionali* come campi continui o costituiti da un numero di unità sufficientemente denso da poter essere trattati come continui. Un campo di questo tipo può essere espresso tramite uno *Spazio di Hilbert*, ben noto in fisica quantistica. I punti essenziali sono:

1. *Tutte le quantità coinvolte sono continue.* Devono essere contemplati sia i casi *normali*, dove la differenziabilità non viene mai meno, che i casi *patologici*, tipo Pour-El e Richards, ed è possibile comunque il passaggio al discreto, necessario per correlare, ad esempio, i processi simbolici tipo SST con le attività sub-simboliche delle RN, cosa che è stato possibile realizzare per alcuni modelli di produzione linguistica. MacLennan ha fornito inoltre un'interessante versione continua del celebre automa cellulare *Life*.

2. *L'informazione è trattata come un tutto continuo nello spazio e nel tempo.* Nella computazione naturale l'informazione viene trattata per pattern continui detti *immagini*, come ad esempio il suono o la visione; in questo modo è più facile modellare quel tipo di computazione che in un sistema naturale corrisponde alla *gradazione continua* e alla variazione per *sfumature*.
3. *Rumore e incertezza sono sempre presenti.* Questa è una richiesta ovvia, poiché abbiamo visto che le fluttuazioni giocano un ruolo essenziale negli amplificatori di informazione; nei sistemi naturali, diversamente che nelle MT dove anche l'imprecisione di un singolo simbolo può compromettere la computazione, l'indeterminazione è una *risorsa*. In questo modo gli aspetti *fuzzy* diventano parte integrante dell'elaborazione di informazione, e le varie generalizzazioni del principio di Heisenberg proposte per i fenomeni biologici e cognitivi trovano un adeguato contesto teorico.
4. *Il processo non termina, ed è adattativo.* Questo è un punto di fondamentale importanza, e coglie differenze essenziali tra la T-comp e la computazione analogica. I sistemi fisici e biologici processano continuamente informazione, e non soltanto in relazione a un compito definito formalmente. Dunque la terminazione o meno del processo non ha la stessa rilevanza, e neppure lo stesso significato (fine del sistema per dissipazione). Si può dire che *i sistemi di computazione naturale non risolvono problemi*, in senso formale (simboli e regole fisse), *ma elaborano strategie di adattamento e soluzioni di risposta*. Questo corrisponde a quanto abbiamo evidenziato nel paragrafo 2, la necessità nello studio dei processi cognitivi di considerare l'attività computazionale *estesa*, distribuita tra ambiente ed organismo.
5. *Le immagini non rappresentano simboli o regole.* Una computazione analogica per immagini non rappresenta nulla, ma corrisponde piuttosto *al comportamento e agli obiettivi dell'agente*, e dunque *non ha bisogno di interpretazione*. E se può comunque introdurre una forma di interpretazione, questa varia continuamente con le immagini, e si hanno perciò *gradazioni di significato* contro i *significati fissati da regole* tipici della T-comp.

Del resto, lo sviluppo di questo tipo di computazione richiede la necessità di specificare meglio in senso matematico alcune caratteristiche, come la *flessibilità*. L'idea della computazione come qualcosa connesso alla scelta dell'osservatore permette di chiarire ulteriormente questi aspetti.

Adottare una computazione analogica significa adottare il punto di vista dell'agente cognitivo. Lasciando da parte per il momento gli aspetti intrinseci di potenza computazionale, è evidente che la differenza qualitativa tra T-comp e computazione naturale è analoga alla differenza che c'è tra un processo *in vitro* e *in vivo*! Infatti la T-comp richiede per poter essere eseguita una completa specificazione degli stati del sistema fisico in senso formale, cosa possibile solo per i sistemi conservatori di informazione, mentre la computazione naturale descrive lo scambio di informazione-energia tra il sistema e l'ambiente, e se abbiamo a che fare con un amplificatore di informazione non sarà possibile, se non in un senso globale, una caratterizzazione definita degli stati e delle *regole* in gioco, poiché queste sono soggette continuamente a cambiamento.

L'interesse per la simulazione dei sistemi fisici ha aumentato in questi anni l'attenzione per gli *automi cellulari* (AC), il cui studio fu iniziato da J. von Neumann negli anni '40. Tra l'altro va detto che uno dei limiti della T-comp è quello di non fornire alcun aiuto nello studio di sistemi di computazione *vocazionalmente* diversi; ad esempio è possibile dimostrare che un AC è equivalente a una MT, e sappiamo che anche le RNR sono (come minimo) equivalenti a una MT, però quest'equivalenza nulla ci dice sulle modalità spazio-temporali della computazione tra i due sistemi. Un automa cellulare è un *reticolo n-dimensionale* dove ogni cella può assumere un numero finito di valori; il comportamento di una cella dipende da quello delle celle contigue tramite un *set di regole di transizione* che vengono applicate simultaneamente a tutte le celle su un *tempo discreto* che definisce il *passo* computazionale. Ad ogni istante lo stato del sistema è identificato da un *vettore di stato* le cui componenti sono i valori di ogni cella.

Esiste una letteratura piuttosto ampia sulle capacità degli automi cellulari di simulare i processi fisici, con particolare riguardo alle configurazioni spaziali. Qui ci limiteremo a ricordare che, negli anni '80, S. Wolfram e C. Langton identificarono quattro classi fondamentali di AC, tra loro correlate al variare di un *parametro d'ordine* λ , che può essere identificato come l'equivalente computazionale della *temperatura* del sistema fisico. Le 4 classi di Wolfram-Langton sono identificate dal tipo di regole di transizione, e il loro sviluppo è indipendente dalle particolari condizioni iniziali assegnate:

Classe 1 : E' una classe di AC che evolve piuttosto velocemente verso *stati omogenei*; nel linguaggio dei sistemi dinamici questo equivale ad avere nello *spazio delle fasi* un *attrattore a punto fisso*, ossia una configurazione ben definita in un tempo finito.

Classe II : Evoluzione verso *configurazioni periodiche o quasi-periodiche*, ossia verso *cicli limite*. Si tratta di comportamenti equivalenti, nel discreto, ai sistemi dissipativi e dunque sono *amplificatori polinomiali di informazione*.

Classe III : Evolve verso *configurazioni aperiodiche molto complesse*, che nel linguaggio del continuo equivalgono agli *attrattori strani*. Si tratta di *amplificatori esponenziali di informazione*.

Classe IV : E' quella che manifesta una maggior *ricchezza* di comportamenti. E' possibile infatti osservare configurazioni variamente complesse, come quelle delle classi II e III, ma anche *strutture dinamiche* regolari, una sorta di *frozen patterns* che ben si prestano a descrivere strutture spaziali definite che si evolvono nel tempo in modalità diverse, da quelle *like-particles* a fenomeni di *accrescimento regolare* con contorni variabili. Un esempio classico sono le forme di *vita artificiale* di *Life*.

L'analisi di C. Langton mostra questo tipo di correlazione al variare del parametro d'ordine:

$$I \rightarrow II \rightarrow IV \rightarrow III$$

cosa che può essere vista come una successione di *transizioni verso il caos* analoga a quella descritta nei sistemi dinamici da uno scenario come quello del *raddoppiamento del periodo* di Feigenbaum.

Dal punto di vista della T-comp è interessante notare che la classe IV può prestarsi ad alcune analogie con il problema dell'indecidibilità di Gödel e degli algoritmi impredicibili di Turing, poiché, fissate le regole e date le condizioni iniziali, è impossibile sapere quali configurazioni verranno fuori. Anche in questo caso, il determinismo locale non implica quello globale. A un'analisi più attenta, però, quest'analogia mostra molti limiti, connessi non soltanto all'imprecisione con cui viene troppo spesso espressa, ma soprattutto con le caratteristiche proprie degli AC. Infatti, nel lavoro di Gödel l'indecidibilità è *puramente* la dimostrazione della incompletezza del sistema, e nella teoria della MT la non-terminazione deve essere correttamente interpretata come un limite dell'analisi computazionale. In altre parole, si tratta di risultati *negativi*, nel senso che fissano dei limiti alle capacità dei linguaggi formali e dei sistemi assiomatici. Negli AC, invece, l'impredicibilità ha un altro significato, più vicino a quelli correntemente attribuiti a situazioni analoghe nella teoria dei sistemi dinamici. Nel caso della classe IV *le regole sono equivalenti a*

forme di non-linearità tipiche degli amplificatori di informazione, e infatti il comportamento degli AC di questa classe è tipico delle ben note situazioni di *criticità auto-organizzata*, situazioni dinamiche in cui il bilanciamento del sistema tra creazione di nuova informazione e dissipazione, tra ordine e caos, si mantiene all'interno di un *ristretto range di valori critici* ed è molto sensibile alle condizioni al contorno (Bak-Tang-Wiesenfeld, 1987). Gli AC di questa classe sono utili per un vasto tipo di simulazioni di sistemi fisici, perché diversamente dalle classi I e II, che conservano l'informazione computata senza elaborarla ulteriormente, o dalla classe III, che non conserva invece *memoria* alcuna delle strutture prodotte, nella classe IV si ha la possibilità di simulare strutture in interazione dinamica. E' questo aspetto che ha stimolato le analogie con la MT universale, ma il punto da sottolineare qui, ancora una volta, è che ci troviamo davanti a un tipo di computazione *qualitativamente diversa*, in grado di descrivere la distribuzione spaziotemporale delle configurazioni prodotte!

La classificazione di Wolfram-Langton ci invita a riflettere su alcune questioni relative alla portata del risultato e sull'attuale *status* della sua formulazione matematica. E' davvero *universale*? In altri termini, è possibile dimostrare che per un automa cellulare a n -dimensioni, con n intero qualunque, non esistono altri comportamenti se non quelli definiti dalle 4 classi esaminate? Bisogna dire che c'è ancora del lavoro da fare in queste direzioni, ma una serie di indicazioni ci autorizzano a rispondere sì a questa domanda. Le ragioni vanno ricercate *nel modo* in cui S. Wolfram ha dedotto l'esistenza delle 4 classi, ponendole a confronto con la teoria generale degli automi finiti ed infiniti, e in corrispondenza con la *gerarchia dei linguaggi formali* di N. Chomsky; inoltre anche l'analisi di C. Langton, basata sulla variazione del parametro λ , porta alle stesse conclusioni, che sono piuttosto saldamente basate su metodi di fisica statistica.

Un'altra domanda riguarda la validità di questa classificazione per quanto riguarda il mondo dei sistemi continui. Anche in questo caso, ci sono diverse considerazioni a favore dell'universalità di questi comportamenti anche per i sistemi continui, motivo per cui è possibile considerare la classificazione di Wolfram-Langton significativa per i sistemi dinamici discreti e continui. Innanzitutto, a livello intuitivo, è possibile pensare le dimensioni di una cella piccole a piacere, in modo da rientrare nel caso dei sistemi continui nel senso di MacLennan, che ha fornito la *versione continua* di *Life*. Del resto, quando in una simulazione al computer si fa corrispondere ad ogni cella un *pixel* dello schermo, si

attua una sorta di *compromesso* tra continuo e discreto, e sperimentalmente noi osserviamo sempre *porzioni* di un fenomeno continuo. Argomentazioni più forti vengono dal lavoro citato di Bak-Tang-Wiesenfeld, che hanno identificato i comportamenti di criticità auto-organizzata lavorando sui sistemi continui, come continui sono i sistemi dissipativi e quelli caotici. Il quadro delle corrispondenze è dunque soddisfacente.

Abbiamo detto che le RN e gli AC sono sistemi di computazione *vocazionalmente* differenti. Infatti, anche se dal punto di vista della T-comp è possibile fare con un AC tutto quello che è possibile fare con una RN e viceversa, una rete neurale è *più adatta* a riconoscere ed apprendere grazie alla sua *plasticità* mediante una serie di input, mentre un AC fornisce un ottimo strumento di simulazione quando si conoscono le regole generali e si vuole indagare l'effetto della variazione delle condizioni iniziali. In pratica, gli AC devono essere progettati in funzione dell'obiettivo modellistico che ci si prefigge.

Il modo in cui la natura elabora l'informazione ha ispirato un gran numero di altri sistemi di computazione. Tra questi bisogna qui ricordare almeno i *sistemi a classificatore*, come gli *algoritmi genetici* (AG), introdotti da J. Holland. Si tratta di algoritmi agenti in uno spazio dei problemi progettati in modo da avere un meccanismo di *mutazione* e *cross-over*, e una *pressione selettiva esterna*, in modo da favorire quelli con il maggior grado di adattamento. Questa ispirazione evolutiva è all'origine del loro nome. In pratica si tratta di sistemi computazionali molto potenti che possono essere analizzati in termini di T-comp, ma anche qui è il *modo* di lavorare che è diverso. Essenzialmente un sistema a classificatore è composto da :

- (a) Un insieme di *rivelatori*, che forniscono gli input alla struttura interna dell'algoritmo *leggendola* dall'ambiente;
- (b) Un insieme di *effettori*, che danno in uscita l'output;
- (c) Un set di *regole*, consistenti in una *condizione* ed un'*azione* del tipo '*se* → *allora*';
- (d) Una set di *messaggi attivi*, che *guidano* l'algoritmo verso l'attivazione della regola *giusta*.

Lo schema di attività degli AG può essere descritto in modo semplice, dicendo che una *popolazione* di AG viene progettata per uno spazio dei problemi, e viene lanciata alla ricerca di soluzioni; ad ogni ciclo la popolazione si modifica grazie a procedure di selezione di ispirazione biologica che cambiano il *genoma* dell'algoritmo e che permettono di eliminare quelli meno adatti, cioè meno efficaci nel riconoscere messaggi

ed elaborare *risposte adattive*. In questo modo si migliora il *paesaggio di fitness*, costituito dagli algoritmi più *forti*. Si utilizzano nella mutazione anche *regole stocastiche*, che equivalgono, in un sistema fisico e biologico, a una temperatura diversa da zero che può testare l'efficienza degli algoritmi senza mai inibirli completamente. Infine, si ottiene una popolazione altamente selezionata, in grado di offrire l'ottimizzazione del paesaggio di fitness. Infatti lo sviluppo impetuoso degli AG è avvenuto in ambito informatico in relazione a complessi problemi di ottimizzazione. Anche da questa sommaria descrizione appare evidente che le cose possono essere molto complicate in dettaglio, ma è importante qui notare che gli AG costituiscono un interessante modello computazionale *misto*. Da una parte infatti possono essere paragonati a una RN multistrato che amplifica l'informazione modificando la sua struttura nell'interazione con l'ambiente esterno, dall'altra la struttura algoritmica del tipo 'se → allora' e la codifica dello spazio dei problemi rimandano ai SST. Gli AG sono infatti, assieme alle RN, tra gli *armamentari* tipici della *Vita Artificiale* (AL, *Artificial Life*) e dell'*Intelligenza Artificiale Distribuita* (DAI), basata sulla programmazione ad agenti che si muovono come entità autonome in un ambiente cooperando e complessificandosi in vari modi. Una procedura tipica è quella di *guidare* una RN dotandola di un *progetto genotipico* tramite un AG.

L'obiettivo ambizioso della AL è quello di individuare degli aspetti formali universali per una comprensione dei processi del vivente, cosa che lascia piuttosto scettici i biologi. Più in generale, l'insieme di questi metodi di computazione viene anche denominato collettivamente *soft-computing*, poiché la quantità di informazione iniziale immessa nel sistema tende a essere minima, lasciando poi all'evoluzione il resto. Tutti questi metodi computazionali rispondono a una duplice esigenza, che ci rimanda alla distinzione iniziale tra SST e PDP: la possibilità di apprendere e modificarsi con l'esperienza, e allo stesso tempo la necessità di gestire l'organizzazione di quest'esperienza secondo una quantità minima indispensabile di conoscenza *a priori*. Si ripropone, in forma matematica, la classica *opposizione/complementarietà* tra *empirismo* e *razionalismo*!

Possiamo ordinare allora i modelli visti in una sorta di gerarchia, che va dai sistemi che apprendono senza regole e sono completamente definiti dall'architettura e dall'interazione con l'ambiente, a quelli dove invece c'è una forte prevalenza delle regole e richiedono una codifica completa dell'ambiente. Entrambi possono dare luogo a fenomeni di

complessificazione ed emergenza, e tutti sono caratterizzati dal non essere sufficientemente descritti in termini di T-comp. I sistemi basati puramente sull'esperienza danno luogo a fenomeni emergenti notevoli, ma scarsamente strutturati, mentre la produzione di nuova informazione è molto vincolata nei sistemi che dipendono fortemente dalle regole, anche se sono in grado di esibire una gestione dell'informazione molto articolata. L'analisi di gerarchie di questo tipo necessita di una discussione più approfondita della nozione di *processo emergente*.

6 - Apertura Logica ed Emergenza

L'interesse per i sistemi non-lineari amplificatori d'informazione e i modelli di computazione centrati sulle relazioni sistema-ambiente ha contribuito allo sviluppo di un intenso dibattito critico sulle definizioni di *emergenza*. Questo concetto ha una lunga *preistoria* scientifica, legata soprattutto all'insoddisfazione dei biologi e degli psicologi per il *riduzionismo meccanicista e determinista della fisica classica*.

E' doveroso qui ricordare l'intensa attività dei neo-realisti di Cambridge negli anni '20, e i contributi di studiosi come S. Alexander, J. Laird, C. D. Broad, C. L. Morgan, e A. N. Whitehead che, diversamente dai loro colleghi del Circolo di Vienna, avevano avvertito i limiti esplicativi profondi della fisica classica in relazione ai problemi fondazionali delle scienze della vita e della cognizione. In discussione era l'aspetto *ontologico* del riduzionismo, che dietro l'apparente *naturalezza* delle posizioni nascondeva sottili problemi concettuali. Infatti dire che un sistema è costituito da certe unità elementari, e che dunque la conoscenza del sistema passa attraverso un'analisi attenta dei componenti, è *ingannevole* almeno sotto due punti di vista: (1) si dimentica il ruolo fondamentale dell'osservatore nella *definizione del sistema* e nella scelta delle variabili e dei parametri da osservare, rilevare, descrivere, e questo significa trascurare l'importanza della scelta modellistica nel definire cosa è e cosa non è un' *unità elementare*; (2) ammesso che non ci siano ambiguità di sorta per ciò che concerne il punto (1), un approccio di questo tipo di per sé non chiarisce perché le cose appaiono organizzate in un modo oppure in un altro, e non è dunque adatto a spiegare quei comportamenti collettivi che non dipendono soltanto dalle caratteristiche dei costituenti, ma del loro modo di interagire su piccola e larga scala, manifestando complesse strutture gerarchiche e dinamiche.

Il problema di quello che oggi viene indicato come *emergentismo classico* riguarda l'incapacità di opporre al riduzionismo uno scenario teorico rigoroso per rispondere effettivamente alle questioni sollevate. In particolare, c'è una tendenza radicalmente opposta al riduzionismo, l'*emergentismo ontologico*, che sostiene l'esistenza di una *struttura multi-livello del mondo oggettivamente intesa*. Ad ogni livello corrispondono principi e leggi diverse che emergono dal livello precedente tramite processi di *auto-organizzazione* del livello inferiore. Questa forma di emergentismo, tipica ad esempio della *prima cibernetica* (*first-order cybernetics*), ha avuto il limite storico di sottovalutare anch'essa il punto relativo alle scelte dell'osservatore, e di non fornire adeguati modelli di auto-organizzazione; contiene inoltre qualche incertezza epistemologica di fondo. Infatti se il mondo è oggettivamente strutturato secondo una gerarchia multi-livello, date la caratteristiche del livello inferiore e le leggi di auto-organizzazione, ogni livello superiore viene ad essere completamente specificato dagli altri.

Questo tipo di emergentismo trasferisce dunque lo stesso concetto di riduzionismo della fisica classica dalle entità elementari ai livelli costituenti. Bisogna dunque rivolgersi alla seconda cibernetica (*second-order cybernetics*) nell'accezione di H. von Foerster per una definizione dell'emergenza che tenga conto delle relazioni tra sistemi, processi ed osservatori, e alla ricca varietà degli strumenti concettuali che sono oggi disponibili per realizzare quest'obiettivo: rami della matematica pura ed applicata (la *teoria qualitativa delle equazioni differenziali*, la teoria delle *singolarità lagrangiane* e delle *biforcazioni*, lo sviluppo della *matematica non-lineare*, le tecniche di *soft-computing*), una gran quantità di fisica (la *termodinamica dei sistemi aperti* e delle *strutture dissipative*, la *sinergetica*, la *fisica del caos*, le relazioni tra *statistiche classiche*, *quantistiche* e *teoria dell'informazione*), i *modelli connessionisti* in psicologia, lo studio delle *correlazioni long-range nelle neuroscienze cognitive*, il dibattito su SST e PDP in IA, e gli approcci 'strutturalisti' e 'sistemici' in biologia e nelle scienze sociali.

L'approccio *sistemico-cibernetico*, sviluppato da studiosi del calibro di N. Wiener, L. von Bertalanffy, R. Ashby e H. von Foerster, si è mostrato un contesto assai naturale e fecondo per affrontare lo studio dei sistemi complessi, perché l'idea stessa di *sistema come insieme di elementi in interrelazione tra di loro*, chiama in gioco esplicitamente una serie di richieste particolarmente rilevanti per la nostra analisi teorica:

- (A) La definizione di *proprietà sistemiche*, che caratterizzano *globalmente* il sistema studiato e non possono essere dunque *banalmente* ricondotte alle proprietà degli elementi costituenti.
- (B) La necessità di considerare, all'interno delle proprietà sistemiche, *comportamenti e vincoli* che possano agire da *regolatori* dei micro-comportamenti del sistema, in modo da poter rendere conto, ad esempio, di quella che il biologo teorico R. Rosen ha chiamato *causalità verso il basso*.
- (C) La possibilità di *molteplici livelli di descrizione* dello stesso sistema, strettamente legati alle scelte del quadro osservativo-sperimentale che è, in ultima analisi, sempre l'espressione di un particolare obiettivo modellistico. Possiamo parlare a questo proposito di una posizione *costruttivista*, che valuta il grado di validità del modello in relazione alla sua fecondità, alla stregua di uno strumento d'osservazione che in sé non è né giusto né sbagliato in *assoluto*, ma viene costruito e utilizzato per certi scopi ben definiti. Bisogna a questo punto eliminare il sospetto di solipsismo, poiché in ogni caso l'osservatore non *crea* il mondo, ma piuttosto si *confronta con esso* tramite i suoi strumenti su un *range* ben definito. Anche la dibattuta questione sull'interpretazione delle teorie, vedi ad esempio quella sui fondamenti della meccanica quantistica, non viene *elusa* da una posizione di questo tipo; è perfettamente legittimo cercare *teorie e interpretazioni diverse*, è necessario però *metterle alla prova* su un preciso *territorio* sperimentale, definendo con chiarezza gli obiettivi che ci si propone e che fanno la differenza tra diversi modelli o diverse interpretazioni della stessa teoria.
- (D) La necessità di introdurre, accanto alle tradizionali descrizioni in termini di materia-energia, descrizioni complementari per rilevare modalità e gradi di organizzazione di queste all'interno di un sistema su varie scale; per questo si sono dimostrati preziosi i metodi della fisica statistica con l'attenzione alla distribuzione dei livelli energetici, e le connessioni di questi con la *teoria dinamica dell'informazione di Sinai-Kolmogorov*.

Durante gli anni '90 si sono sviluppate varie teorie dell'emergenza, tutte centrate sul tipo di strumenti modellistici usati come *misura* del tipo di emergenza. Un'intensa attività di ricerca nel campo dei sistemi dinamici e del soft-computing, è alla radice del lavoro di J. Crutchfield (1994), che propose una distinzione dei vari tipi di emergenza in rapporto alla

complessità computazionale osservata. Il problema diventa allora quello di ampliare il repertorio modellistico attraverso una procedura denominata *Hierarchical Machine Reconstruction*. Si tratta in sintesi di *ottimizzare il rapporto informazione/rumore* quando si cerca di individuare il modello *corretto* per descrivere un certo livello di complessità misurato; questo è individuato dalla capacità di inquadrare in una serie di *relazioni causali*, opportunamente definite, le sequenze prodotte dal sistema, minimizzando il *rumore* e la *casualità*. La modificazione del modello è costituita dunque da una *gerarchia induttiva* che parte dall'obiettivo di *comprimere algoritmicamente* la maggior quantità di informazione possibile. Nella stessa direzione si è mosso F. Heylighen (1991) con la sua *Dinamica delle Distinzioni*, un *meta-modello* che permette di classificare i vari modelli particolari in relazione alle variabili di stato e le funzioni di transizione che lo caratterizzano e di orientare la ricerca di nuovi modelli.

Un processo emergente opera in un sistema una variazione all'interno della classe delle variabili e delle regole, e modifica così l'input e l'output del sistema e le relazioni con l'ambiente. Il problema di trovare un modello adeguato diventa allora quello di operare nuove distinzioni nel meta-modello, grazie a due procedimenti fondamentali di ispirazione darwiniana, la *variazione* e la *selezione*. Con la variazione ci si assicura la possibilità di generare un *set* di nuovi stati spazio-temporali; con la selezione si opera invece in senso opposto, rimuovendo stati e relazioni che non supportano comportamenti stabili, e dunque misurabili. Poiché il problema delle relazioni tra sistema e ambiente non può essere fissato una volta per tutte in modo netto, Heylighen ha adottato il concetto di *chiusura relazionale*, utilizzato anche da Maturana e Varela nella loro *teoria dell'autopoiesi*. In pratica si riconduce a una rete di relazioni interne al sistema lo status del rapporto con l'ambiente, sottolineando così non soltanto l'*autonomia del sistema*, ma assicurandosi sempre la possibilità di descrivere l'influenza dell'ambiente (e sull'ambiente) attraverso una quantità finita di informazione relativa alla struttura interna del sistema e ai vincoli sulla configurazione della rete. La variazione e la selezione applicate alla rete permettono così di ottenere il modello cercato in funzione del tipo di emergenza che si vuole descrivere.

Un terzo approccio è quello sviluppato da N. A. Baas e C. Emmeche (1997), ed è basato sul concetto di *Iper-Struttura*, una struttura di insiemi del tipo (OBS_1, S_1, INT_1) , dove OBS è un quadro osservativo-sperimentale, S un insieme di sistemi, e INT indica le relazioni tra i

sistemi di S . Le peculiarità di S e delle INT individuano una nuova struttura $S_2 = F(OBS_1, S_1, INT_1)$. Una proprietà P è detta emergente se appartiene alla *struttura di secondo livello* S_2 , ma non è osservabile in S_1 , e può essere studiata solo all'interno di un nuovo quadro OBS_2 , e così via. In pratica, le strutture emergenti sono definite attraverso una gerarchia multi-livello che è descritta nella sua forma più generale dall'Iper-Struttura e dalle sue modificazioni.

Anche da questa descrizione sommaria si può notare che i tre approcci sono tra loro molto simili. Quello di Crutchfield è tipicamente connesso al mondo concettuale della meccanica computazionale e della teoria dell'informazione, mentre gli altri due sono più tradizionalmente basati sul linguaggio formale della cibernetica e della teoria dei sistemi. Li accomuna l'idea che una certa sequenza di dati osservativi *inaspettati* all'interno di un modello ne richiede una modificazione strutturale che in qualche modo rispecchia la riorganizzazione del sistema e le nuove proprietà emergenti. E' facile comprendere che la compressione algoritmica di Crutchfield ha lo stesso obiettivo dei vincoli di variazione e selezione nella dinamica delle distinzioni di Heylighen e degli *assestamenti* gerarchici della Iper-Struttura di Baas-Emmeche. Questi ultimi hanno sviluppato anche una distinzione molto utile tra *emergenza computazionale* ed *emergenza osservativa* che costituisce una distinzione implicita fra i 3 approcci. La prima, come suggerisce la terminologia, è relativa a configurazioni che emergono da un processo computazionale e sono dunque, in linea di principio, individuate da un set di regole e da una stringa di condizioni. In altre parole, esiste sempre un processo computazionale D tale che $P = D(OBS_1, S_1, INT_1)$. Nel caso dell'emergenza osservativa invece non esiste alcun processo algoritmico che porta dal sistema di primo livello alla proprietà emergente del secondo livello. Adottare una posizione di questo tipo vuol dire in pratica mettere in discussione a un livello radicale ogni teoria che pretende di definire emergenza effettiva come integrazione ricorsiva di relazioni, come accade, con modalità diverse, sia nella teoria di Crutchfield che in quella di Heylighen. Per Baas-Emmeche i sistemi non-lineari continui e discreti e gli AC *non costituiscono* processi emergenti *genuini*, perché essendo algoritmicamente comprimibili sono contenuti *in nuce* nel formalismo di primo livello.

Il problema è molto sottile, e riguarda la stessa natura della scienza della complessità. A questo punto bisogna chiedersi quali sono i sistemi che esibiscono emergenza osservativa, e in che senso si tratta di un'emergenza

non alitmicamente comprimibile. Un esempio sono gli agenti autonomi della DAI che si complessificano nell'interazione con l'ambiente e, alla lunga, mostrano comportamenti imprevedibili sulla base della descrizione iniziale del progetto. Però in questo caso si potrebbe obiettare che comunque il tipo di emergenza osservata negli agenti si può ricondurre a una ristrutturazione interna che rientra nel *modello autopoietico* di Heylighen. E se estendiamo la descrizione alitmica a una struttura del tipo *sistema+ambiente*, passiamo di nuovo a un caso di emergenza computazionale! Tutto ciò ci porta ad interrogarci con più attenzione sulle modalità del rapporto tra un sistema e il suo ambiente nell'ambito di una scelta modellistica. Una analisi di questo tipo è stata proposta dal teorico dei sistemi P. Cariani (1992), e può essere illustrata facendo riferimento al rapporto tra T-comp e *computazione naturale*.

Per definizione, la T-comp implica una completa codifica del flusso informativo, che è parte integrante essenziale della dinamica computazionale. Ora consideriamo un sistema naturale, o un sistema artificiale molto raffinato, e situiamolo in un ambiente *reale*, ossia descrivibile in termini di fisica e biologia *ordinarie*. L'apparato *sensoriale* del sistema rileva, e tra l'altro non in modo asettico ma *modificando irreversibilmente l'ambiente*, un flusso informativo, continuo o discreto. Per poter elaborare e utilizzare il sistema, deve convertirlo in serie multiple di *codici interni*, ossia deve *operare un'appropriazione semantica dell'informazione*. Benché qui non ci addentriamo in definizioni formali, va sottolineata l'importanza fondamentale del *concetto di codice*, inteso come capacità del sistema di immagazzinare e gestire l'informazione in relazione alla sua struttura interna, alla storia dinamica e alle relazioni con l'ambiente. Una volta elaborato l'output, questo per agire sul mondo esterno deve essere suscettibile di riconvertirsi nel codice ambientale, che in ultima analisi è quello della fisica classica o quantistica, e per il quale valgono dunque le relazioni tra quanti, livelli energetici e informazione. Possiamo concludere che esiste un'*emergenza sintattica*, dove la computazione agisce a livello dello stesso codice, e che è virtualmente riconducibile a una compressione alitmica, o comunque ad una descrizione unitaria tramite le specifiche sintattiche in un super-modello, mentre l'*emergenza semantica* è quella in cui rispetto all'osservatore si creano nuovi comportamenti del sistema, legati anche a nuove modalità di misura e controllo. Questo può anche essere visto come un modo del sistema di costruire un *modello del mondo* e quindi, se vogliamo comprendere e fare previsioni, dobbiamo

adattare la nostra modellistica in modo da poter descrivere non soltanto le modificazioni strutturali del sistema, a livello sintattico, ma anche la sua *rappresentazione interna del mondo*, indagandone la *semantica*!

La possibilità di costruire questo *modello del modello*, secondo lo schema di Baas-Emmeche, è dunque fondata sul gioco della pluralità dei codici esistenti in natura. A questo proposito, è opportuno qui ricordare la *teoria semantica dell'evoluzione* sviluppata da Marcello Barbieri (1985; 2000). Adesso supponiamo per un attimo di poter conoscere la varietà di codici che regolano l'attività dei vari sistemi. L'idea è in sé assai vaga, perché implica che non ne emergano più di nuovi, il ché è proprio l'opposto di ciò che osserviamo. Conoscendoli, sarebbe in teoria possibile costruire un *super-codice* che operi come un *decodificatore universale*. In realtà, ammesso che ciò fosse possibile, si ricadrebbe in una nuova forma di *riduzionismo sintattico*, e *non si potrebbe più rilevare emergenza*, poiché l'essenza dei processi emergenti consiste proprio nella grande varietà di livelli sintattici e semantici dei processi osservati (noi compresi, naturalmente). Sembra dunque naturale associare l'*emergenza intrinseca*, la forma più *forte*, con l'*emergenza semantica*, tipica dei sistemi biologici e cognitivi, la cui capacità peculiare è proprio quella di interagire in modo selettivo e mirato con l'ambiente, tramite strategie adattative che richiedono proprio l'emergenza di nuove rappresentazioni del mondo centrate sull'agente.

Ritornando alle distinzioni tra rappresentazioni simboliche e algoritmiche dei SST e dinamiche sub-simboliche, è possibile ricavare un elemento chiarificatore dalla nostra analisi sull'emergenza. Una volta che un processo cognitivo emerge e viene rilevato, opportunamente fissato nel tempo e in relazione ad una precisa situazione ambientale, allora è possibile descriverlo come procedura algoritmica, che è appunto una forma di riduzionismo sintattico, ma è impossibile presumere che una descrizione simbolico-formale possa fare altro che descrivere un insieme di atti cognitivi in un ambiente-schema, perché per sua natura non può accedere al livello di produzione semantica costituito proprio dai processi emergenti. Si ripresenta la dicotomia *in vivo/in vitro*; l'IA *classica* studia la cognizione *in vitro*, il connessionismo *in vivo*. Ma a ben guardare questa dicotomia è anche una forma indispensabile di complementarità. Infatti il livello sub-simbolico è quello dei processi non-lineari che amplificano informazione, ed è dunque il *luogo* dove *accadono* le emergenze semantiche. Per sua natura, però, un'emergenza intrinseca non può essere descritta e tantomeno prevista *prima* che si produca, ma soltanto *dopo*, una volta fissato quello che

Baas ed Emmeche chiamano il secondo livello. La giustificazione profonda della *trasparenza semantica* dei sistemi simbolici consiste proprio nel fatto che essi descrivono un processo cognitivo *lontano dalle zone d'emergenza* che l'hanno generato, ed è dunque fissata una volta per tutte l'informazione necessaria all'osservatore per una completa codifica sintattica, come ad esempio mettere in forma verbale un pensiero. Ed è per questo motivo che non sarebbe neppure possibile una *scienza della mente* senza accettare un *rappresentazionalismo simbolico minimo*, la cui ambizione non è quella di descrivere come si sviluppa un processo cognitivo, che è qualcosa di intrinsecamente emergente, ma piuttosto di *fissarne* i risultati.⁵ E' per motivi analoghi che la termodinamica all'equilibrio è ben diversa da quella dei regimi di non-equilibrio.

L'analogia tra la *produzione di un processo cognitivo*, come ad esempio la soluzione di un problema, e le *transizioni di fase* sono state sviluppate nella *Harmony-Theory* (HT) di P. Smolensky (1986). Abbiamo già fatto notare che ogni forma ingenua di riduzionismo neurale *bottom-up* che pretendesse di poter dedurre il livello SST dal livello neuronale sarebbe destinata al fallimento, e le considerazioni fin qui svolte sull'emergenza ci confortano in questo senso. Quello che Smolensky ha fatto è di tipo diverso, ed anzi compatibile con quest'analisi. Dato un compito cognitivo descritto da uno schema, l'apparato formale della teoria prevede che il valore di una particolare quantità, detta *armonia*, viene ottimizzato con l'attivazione dei pattern neurali che completano l'informazione fornita dallo schema. Ad uno stesso schema corrisponderanno pattern diversi, ma tutti soddisferanno la condizione sull'armonia, in modo simile alla varietà dei livelli energetici disponibili ad una temperatura fissata. Si tratta di un procedimento *top-down* interessante non soltanto per le rilevanti applicazioni nella progettazione di reti neurali, ma anche sotto l'aspetto teorico. Se consideriamo infatti la reale complessità del sistema nervoso, la teoria di Smolensky ci fa comprendere proprio come processi neurali diversi ad un livello possono apparire attività cognitive confrontabili e descrivibili ad un altro livello, ma questo non autorizza minimamente

⁵ Nello *stile di Wittgenstein*, potremmo dire che un pensiero descritto è un pensiero già pensato! A ben guardare, questa è anche una chiave di lettura delle due *fasi estreme* dell'opera di L. Wittgenstein: *Tractatus* (chiusura logica) e *giochi linguistici* (apertura logica).

alcuna idea di corrispondenza banale, poiché uno schema in ingresso, in sintesi, *stimola ma non identifica univocamente* i processi di emergenza che innescano. Queste idee possono essere inquadrare nell'ambito della *teoria dei sistemi logicamente aperti* (H. von Foerster, 1987, 1996; G. Minati, M. P. Penna, E. Pessa, 1998), in relazione alla triade inscindibile *osservatore-sistema-modello*.

La teoria parte dalla posizione fondamentale del pensiero sistemico, ossia che è necessario considerare *sistemi* e *modelli* come dimensioni concettuali *in continua inter-relazione*. Infatti la costruzione di una teoria scientifica comincia sempre con la scelta di una *porzione del mondo* da studiare, e fissa dunque delle relazioni tra sistema ed ambiente. Questo equivale a selezionare una quantità *finita* d'informazione a partire da un'indefinita varietà di scelte possibili. Un'operazione di questo tipo implica l'uso di quello che potremmo chiamare un *modello implicito*, una serie di argomentazioni a sostegno della nostra identificazione del sistema in oggetto. L'attività osservativa-sperimentale produrrà dei dati che saranno inquadrati all'interno di un *modello formale*, ideato per fornire una comprensione strutturale delle interazioni in gioco e produrre delle previsioni quantitative sul comportamento del sistema. Il confronto tra dati e modello potrà portare a delle modificazioni nella costruzione teorica, e questo potrà riflettersi a sua volta in una *ri-definizione del sistema* di partenza, in un progressivo raffinamento dei modelli formali e dei quadri sperimentali. Naturalmente questa è una descrizione semplificata, poiché in campi dove si dispone già di una grossa quantità di conoscenza acquisita, la definizione di nuovi strumenti modellistici e di nuovi sistemi nasce con il *vincolo* di essere *almeno compatibile* con le teorie consolidate, in un processo naturale di *stabilizzazione* della conoscenza (processo che rischia spesso di essere *ipostatizzato* tramite un'identificazione banale tra modelli e sistemi, che vengono così intesi *erroneamente* come *oggetti ontologici* costituenti *oggettivamente* il mondo).

E' possibile stabilire una gerarchia di modelli formali sulla base delle relazioni sistema-ambiente fissate da un osservatore. Con il termine *osservatore* intendiamo l'insieme delle procedure che permettono di ricavare una serie di valori dall'analisi del sistema, valori codificati all'interno di un modello formale ricorsivo. E' possibile, seguendo le idee di H. von Foerster, dare una definizione più rigorosa di osservatore attraverso un'*equazione agli autovalori* analoga a quelle usate in meccanica quantistica, ma per i nostri scopi sarà sufficiente l'idea intuitiva di *insieme di procedure*.

Diremo che abbiamo un *modello logicamente chiuso* quando è possibile dare una descrizione completa ed esplicita del sistema e delle sue interazioni con l'ambiente, e dunque della sua evoluzione e delle caratteristiche strutturali. Un esempio è la *fisica classica*, dove l'*equazione di Newton*, assieme ai valori di posizione q e momento p ad un istante t , *fissano* l'intera storia dinamica del sistema. L'introduzione della non-linearità modifica in modo drastico le aspettative deterministiche del sistema. Storicamente, lo studio del famoso problema dei 3 corpi in mutua interazione gravitazionale è all'origine della crisi del determinismo classico e ha fornito uno spunto decisivo per la nascita della *teoria qualitativa delle EqDiff*, sviluppata tra gli altri da Poincaré e Ljapunov. Questa teoria offre la possibilità di studiare un sistema in modo *globale*, indipendentemente dal fatto che non ogni stato del sistema è determinabile a causa dell'amplificazione di informazione. Anche in questo caso abbiamo a che fare con un sistema logicamente chiuso perché siamo comunque in grado di fare previsioni strutturali sugli *stati asintotici* del sistema studiando la forma dell'attrattore, e l'informazione che resta *indeterminata* rispetto al caso lineare è comunque racchiusa in una regione *finita* dello spazio delle fasi. Questo vuol dire che *qualitativamente* non differisce dall'informazione in nostro possesso, ed è possibile fare delle valutazioni *quantitative* della nostra *incertezza*. In altre parole, sia nel caso lineare deterministico che nei casi non-lineari, il ruolo dell'osservatore non cambia, poiché le relazioni tra sistema e ambiente sono sempre definite in modo netto, le modalità tramite cui l'informazione varia sono ben stabilite, e non ci sono livelli di codifica nuovi.

Definiamo come *modello ad apertura logica zero* ogni *modello logicamente chiuso*. Man mano che ci allontaniamo dal range dei fenomeni della fisica classica, lungo una direzione ideale di crescente complessità che attraversa la fisica statistica classica e quantistica fino ai fenomeni della vita e della cognizione, ci rendiamo conto che i sistemi logicamente chiusi sono l'eccezione più che la regola, ad esempio nei fenomeni di emergenza intrinseca. Si tratta di casi in cui è utile distinguere tra *modelli ad apertura logica di livello uno* e di *livello due*. Nel primo caso abbiamo a che fare con *sistemi meta-stabili*, in cui le relazioni con l'ambiente non sono esplicitabili poiché dipendono in larga misura dallo stato interno del sistema, come nei casi delle *transizioni di fase* e dei *problemi a frontiera variabile* tipici della matematica integro-differenziale. In tutti questi casi è possibile definire il tipo di dipendenza, sotto forma di una legge o di una serie di vincoli, e dunque la distinzione tra sistema e ambiente è parzialmente conservata, ma

parte dell'informazione in gioco non è accessibile all'osservatore e risulta *indeterminata*. Nel caso di modelli ad apertura logica di secondo livello, la *forma* stessa delle relazioni sistema-ambiente dipende dal sistema. La cosa può essere vista come una distinzione nel modo di gestire l'informazione da parte del sistema rispetto al modello: il sistema genera dei comportamenti imprevedibili sulla base dell'informazione in possesso dell'osservatore. Si tratta dunque del regno dei processi di *emergenza semantica*, tipico degli organismi biologici e dei processi cognitivi. Questo indica che *il sistema è sufficientemente complesso da giocare un gioco diverso da quello individuato dalle regole del modello*, creando nuovi codici che non possono essere analizzati dall'osservatore tramite una computazione *interna*, poiché l'insieme delle procedure formali disponibili non è chiuso, ossia non è ricorsivo (ad esempio, i dati ottenuti tramite misure e osservazioni dal sistema non possono essere ricavate dalle formule del modello). Arriviamo così a stabilire il fondamentale risultato che *i sistemi ad apertura logica di secondo ordine esibiscono processi di emergenza semantica*.

Introdurre un'apertura logica per livelli successivi – livello tre, livello quattro, etc. – significherebbe poter classificare ogni tipo possibile di emergenza semantica. Ma questo significherebbe postulare proprio un osservatore onnisciente capace di costruire un super-modello in grado di definire tutti gli stati interni e quindi ogni possibile relazione sistema-ambiente, cosa che le *basi stesse* della teoria individuano come un *obiettivo mal posto*. Si può procedere però in altro modo, assumendo che *i principi generali della fisica*, e di quella che abbiamo chiamato computazione naturale, *sono validi per ogni tipo di sistema*. Bisogna ricordare, anche senza addentrarsi in analisi dettagliate, che *più un sistema si complessifica diversificando la sua struttura interna, più aumenta la dissipazione*. Se un sistema non viene distrutto dalla dissipazione, allora vuol dire che una serie di *vincoli* lo impedisce, ed è abbastanza intuitivo che il numero di questi vincoli cresce con la complessità del sistema. I vincoli individuano in quale misura il sistema mantiene la sua struttura e la sua dinamica tramite un *compromesso* termodinamico con l'ambiente, e nei sistemi almeno di livello due, il vincolo non può essere esplicitato nel modello poiché dipende fortemente dalla struttura del sistema.

Attraverso la nozione di vincolo, dunque, possiamo stabilire una gerarchia di modelli di apertura logica che corrispondono alla capacità dell'osservatore di gestire almeno in parte la complessità di un sistema. Definiamo allora un *modello logicamente aperto di grado n* se esso può

essere caratterizzato almeno da un numero n di vincoli, con n finito. Da un punto di vista della chiusura operativa che fa da sfondo alla teoria di Maturana e Varela questo vuol dire che il tipo di chiusura del sistema può essere esemplificato da un grafo con n vertici. Poniamo al *top* di questa gerarchia un *modello logicamente aperto di grado n con $n \rightarrow \infty$* . Una definizione di questo tipo significa ipotizzare un modello in grado di rendere conto di ogni tipo di vincolo, e dunque di fornire una descrizione completa di tutte le possibili relazioni sistema-ambiente. Abbiamo già detto che questo non può essere inteso come la possibilità di costruire un super-osservatore, per via dell'arbitrarietà delle scelte modellistiche, ma piuttosto come la necessità di una teoria che sia in grado di *trattare in modo unitario sistema ed osservatore*.

Si può dimostrare che (1) è impossibile descrivere un modello logicamente aperto tramite *un solo* modello formale, e (2) ogni descrizione di un modello logicamente aperto tramite un modello logicamente aperto di grado n ha un dominio limitato di validità, per ogni valore finito di n . E' abbastanza facile comprendere che questi risultati limitativi sono connessi, attraverso un'opportuna scelta sintattica, ai teoremi di indecidibilità di Gödel e Turing sull'incomprimibilità algoritmica. Nel caso dei sistemi formali questi teoremi mostrano che *la matematica è un sistema aperto*, in cui ogni produzione non è sempre e necessariamente riconducibile a una struttura assiomatica nel senso di Hilbert. In senso più generale questo vuol dire che ogni particolare *sistema osservatore* non può, tramite le sue rappresentazioni, descrivere e prevedere ogni possibile *sistema osservato*, se non a prezzo di una certa approssimazione. Vediamo così che, lungi dall'essere una caratteristica di una teoria *esotica* come la meccanica quantistica, i *principi di indeterminazione* hanno un'ampia *cittadinanza* nello studio dei fenomeni naturali e artificiali di ogni tipo, e sono legati all'inevitabile *divario* di complessità tra osservatore e osservato, riconducibile sempre a un'arbitraria partizione del mondo.

Nel caso dello studio dei processi cognitivi si può affermare che *l'osservatore SST si pone a un livello di chiusura logica che riesce a formalizzare tutti quei processi dove il numero dei vincoli n è basso, e non dipende né dal tempo né dagli stati del sistema*. L'osservatore PDP o DAI, invece, *adotta una prospettiva di apertura logica su una serie di modelli con elevato numero n , dipendente dal tempo e dagli stati del sistema*.

I modelli connessionisti e distribuiti sono in grado, in qualche misura, di cogliere l'aspetto essenziale della complessificazione emergente in relazione

all'attività relazionale con il mondo. L'amplificazione di informazione prende in questi casi la forma non soltanto di un accrescimento della conoscenza inizialmente immessa come *seme iniziale*, ma la capacità del sistema di *riorganizzarla continuamente* per nuovi compiti. La stessa teoria ci suggerisce non soltanto la possibilità di costruire un ponte tra SST e PDP, ma se consideriamo la reale complessità strutturata e modulare del cervello, ossia il numero enorme di n , è evidente che queste due classi di modelli sono assai più vicine tra loro di quanto non si ritenga comunemente, almeno per quanto riguarda la *plausibilità* biologica.

Tutte le teorie dell'emergenza qui presentate hanno il grosso limite di fornire strumenti più o meno efficaci per cogliere e descrivere emergenza, ma non sono in grado di connetterla ad alcuna teoria generale del mondo fisico. La teoria dei sistemi logicamente aperti ci suggerisce che *occorre una teoria dove l'osservatore sia esplicitamente preso in considerazione* e trattato alla stessa stregua degli altri sistemi. I sistemi della teoria devono essere in grado non soltanto di *esibire un numero elevato di stati*, ma anche di *poterli sufficientemente strutturare*, in modo da poter disporre di un *framework* teorico per l'emergenza e per i *comportamenti collettivi*. L'indeterminazione, inoltre, deve fare la sua comparsa in modo naturale, non come limitazione pratica o incertezza metodologica, ma piuttosto come una *componente fondamentale* dei processi naturali.

Poiché sappiamo che l'informazione naturale è legata allo scambio di energia tra sistemi fisici e che può avere aspetti sia continui che discreti, sappiamo già che dobbiamo rivolgerci alla fisica, ma non alla fisica classica. E' possibile dimostrare infatti che in un mondo di oggetti *puramente classici*, regolato dalla *statistica di Maxwell-Boltzmann*, è possibile rendere conto della dissipazione, ma non della formazione di strutture, a causa della relativa *povertà* dei comportamenti in gioco. I sistemi classici non-lineari permettono di considerare fenomeni di amplificazione di informazione, ma se cerchiamo un modello di sistema cognitivo dobbiamo restringere la nostra ricerca ai comportamenti di criticità auto-organizzata, perché ci interessa non soltanto la genesi delle strutture, ma anche la capacità di *conservare e riutilizzare* l'informazione acquisita. Ma la teoria della criticità auto-organizzata, come del resto la teoria dei processi caotici, ha mostrato dei forti limiti nella costruzione di modelli biologici e cognitivi proprio per il suo *sfondo concettuale classico*, che riduce l'importanza e la varietà della nozione di *transizione di fase*, che è in definitiva legata al tipo di statistica considerata. Del resto, ogni comportamento classico, sotto

opportune condizioni, può essere visto come limite di una teoria che a tutt'oggi appare l'unica candidata per la costruzione di una teoria unitaria dei fenomeni naturali, la teoria quantistica.

7 - Teoria Quantistica e Processi Cognitivi

Nel seguito ci limiteremo a delineare gli aspetti essenziali del formalismo quantistico, mettendo l'accento su quelle caratteristiche della teoria che sono necessarie per i nostri scopi. Contrariamente a quel che si sostiene, i concetti di base della *teoria quantistica* sono molto semplici, e descrivono i sistemi fisici nel modo più generale. In particolare, non è vero che il range della *meccanica quantistica* (MQ) è limitato ai fenomeni su scala microscopica (scala della *costante di Planck*, h). Questa terminologia *imprecisa* dimentica che se è storicamente vero che lo sviluppo della fisica dei quanti è legato allo sviluppo della microfisica, l'evoluzione della teoria ha poi mostrato che la descrizione quantistica è quella più naturale e profonda. Il *mondo è essenzialmente quantistico*, e non in senso banalmente riduzionista, perché composto da 'particelle' dal comportamento quantistico, ma perché la MQ fornisce pure le *ragioni* della fisica classica. Inoltre è sempre più rilevante anche dal punto di vista sperimentale l'evidenza di *effetti quantistici macroscopici*, legati proprio al tipo di *processi collettivi* che ci interessano.

L'oggetto fondamentale in MQ è il *vettore di stato* $|\Psi\rangle$ definito nello *Spazio di Hilbert* $\mathbf{H}$, uno spazio a infinite dimensioni. Il vettore di stato contiene tutte le informazioni necessarie per determinare ogni grandezza misurabile del sistema, e può essere espresso come *combinazione lineare* di *autovettori* $|a_n\rangle$. Ad ogni misura corrisponde un *operatore hermitiano* a valori reali che indicheremo genericamente con $\hat{O}$. In questo modo l'osservatore entra direttamente e in *modo radicale* nell'apparato della teoria. Una misura modifica il vettore di stato *irreversibilmente*, producendo un determinato autovettore $|a_n\rangle$ che produrrà gli *autovalori* A_n (*collasso del vettore di stato*). In sintesi, il processo di misura è descritto da un'equazione agli autovalori del tipo:

$$(7.1) \quad \hat{O} |a_n\rangle = a_n |a_n\rangle$$

Gli autovalori possono essere un insieme di *valori discreti* o *continui*, fatto che dipende dalle condizioni al contorno del problema (es., i valori discreti

dell'energia nello studio delle interazioni). E' importante sottolineare che ogni oggetto fisico sperimentalmente rilevante appartiene alla successione degli A_n , e può essere considerato il risultato dell'operazione di misura sul vettore di stato. In questo senso non esistono *mattoni del mondo*, poiché il vettore di stato fornisce una *descrizione completa di ogni sistema fisico*. Si può affermare al più che il vettore di stato individua un insieme sovrapposto di stati, espressi attraverso la combinazione degli autovettori, che vengono poi *selezionati e fissati* dall'operazione di misura, producendo serie di autovalori la cui probabilità è connessa alla situazione fisica studiata. Quest'aspetto della teoria è al centro di un interessante dibattito che è ben lontano dall'essersi concluso, ma che *non* ha mai messo in crisi la *coerenza* della teoria e la sua *fecondità* in un range sperimentale vastissimo. Sarà qui sufficiente accennare che ulteriori chiarificazioni di questi problemi indicano una profonda connessione tra teoria quantistica dei campi (dinamica) e struttura *fine* dello spazio-tempo (topologia) (Licata, 2003).

La legge del moto più semplice in MQ è la ben nota *equazione di Schrödinger*, che descrive l'evoluzione del vettore di stato di una particella libera. La completezza della descrizione quantistica può essere specificata dal punto di vista matematico dicendo che vale per gli autostati una *relazione di chiusura*. Fisicamente vuol dire che sono *invarianti nel tempo* e contengono *tutta l'informazione disponibile* per la descrizione di un sistema fisico. Questa relazione ha due limitazioni. Innanzitutto l'insieme degli autostati non è sempre specificabile nella sua interezza, poiché si tratta comunque di una piccola frazione degli stati possibili. E' possibile soltanto in un ristretto numero di situazioni fisiche. Inoltre, il processo del collasso del vettore di stato non permette, *per la natura stessa della relazione quantistica tra sistema e misura*, di individuare *prima* di un'osservazione gli autovalori. Questo vuol dire che *il formalismo quantistico gode di una chiusura logica soltanto parziale*: la ricchezza degli stati del sistema non è riconducibile a una manipolazione algoritmica, come invece è possibile fare, almeno in linea di principio, nella fisica classica. La MQ gode quindi di alcune caratteristiche necessarie per costruire una teoria dell'apertura logica.

Un risultato essenziale sui sistemi quantistici è il *principio di indeterminazione di Heisenberg*, che fissa le condizioni sugli osservabili ed esprime *in nuce* le differenze tra la *risoluzione continua* dei sistemi classici e la descrizione quantistica *discreta*. Il principio di Heisenberg è la *via regia* per la costruzione della *Teoria Quantistica dei Campi* (TQC),

che si sviluppa dalle condizioni imposte alla descrizione dei campi continui classici dai principi della MQ. Ad ogni punto dello spazio-tempo è associato un *oscillatore armonico* la cui attività vibrazionale fornisce una successione di valori discreti dell'energia (*quanti del campo*). La quantizzazione di un campo classico equivale alla quantizzazione dei suoi oscillatori. Per il principio di Heisenberg questi oscillatori non sono mai in uno stato di riposo analogo a quello che si potrebbe trovare in fisica classica. Il *vuoto quantistico* è un'entità dinamica che produce effetti virtuali, tramite le fluttuazioni di punto zero (ZPF, *Zero Point Fluctuations*), ed effetti osservabili, descrivibili tramite opportuni *operatori di creazione e di distruzione* di tipo *fermionico* e *bosonico*, in relazione alle due principali famiglie di oggetti quantistici. Questi mostrano comportamenti collettivi molto diversi. I fermioni possono occupare uno stesso stato quantico, secondo il principio di esclusione di Pauli, e obbediscono alla *Statistica di Fermi-Dirac*. Ad un singolo stato quantistico può corrispondere invece un numero arbitrario di bosoni, nella modalità descritta dalla *Statistica di Bose-Einstein*. La statistica classica è un caso-limite di quelle quantistiche, con la differenza, tra l'altro, che *in TQC il numero di particelle non è costante*, ed è un'espressione della variazione dei modi vibrazionali del campo.

Per comprendere le possibilità espressive offerte dalla teoria bisogna rimarcare la differenza tra MQ e TQC. La prima è *poco più che una cinematica* dei sistemi quantistici, mentre la complessa attività del vuoto quantistico ci permette di descrivere a un livello fondamentale la *dinamica sottile* delle interazioni. Ci sono due teoremi significativi a questo proposito. Il *teorema di von Neumann* (1939) afferma che un sistema quantistico finito, come nella MQ, ammette un unico stato di minima energia, mentre un sistema quantistico dotato di infiniti gradi di libertà, come avviene nella TQC, ha *infiniti stati di vuoto tra loro non-equivalenti*. Dal punto di vista modellistico questo equivale a dire che il numero di rappresentazioni di un sistema quanto-meccanico è assai limitato, ed è definito univocamente dalle condizioni al contorno, mentre in TQC abbiamo a che fare con un numero infinito di rappresentazioni fisicamente non-equivalenti, legate all'organizzazione energetica dei sistemi all'interno degli stati di vuoto multipli. E' facile comprendere che il teorema di von Neumann è l'analogo in campo quantistico dei teoremi classici della fisica statistica, con la differenza che nessun oscillatore quantistico conosce uno stato di minima energia definito e fisso. In questo modo abbiamo la possibilità di dare una descrizione di sistemi complessi dove intervengono più livelli

energetici in modo strutturalmente stabile. Nella stessa direzione va il *teorema delle fluttuazioni dissipative*, dovuto a Callen e Welton (1951). Semplificando, il teorema illustra un aspetto importante della TQC, ossia che è impossibile ogni distinzione tra sistema e ambiente, e *l'informazione quantistica è sempre distribuita tra i vari modi del campo*.

In un sistema fisico aperto qualunque eccesso di energia viene dissipato dalle fluttuazioni e, aspettando un tempo sufficiente, l'energia del sistema ritorna allo stato base (GS, *Ground State*). Le strutture osservabili nascono tutte dal GS come *stabilizzazione* dei livelli energetici del vuoto. Il mondo fisico nella TQC è descritto da un complesso insieme di oscillatori, e l'analisi delle interazioni tra campi può comportare delle notevoli complicazioni di calcolo che si possono in parte evitare ricorrendo a *soluzioni perturbative*, ossia approssimate ed essenzialmente lineari.

La ricerca crescente nel campo dei *metodi non perturbativi e non-lineari* ha permesso in questi ultimi anni un grosso passo in avanti nello studio della *ricchezza delle strutture quantistiche*, permettendo di estendere il range della teoria e di verificarne l'universalità. Per affrontare la questione da un punto di vista generale, rivediamo un modo di scrivere le *relazioni di Heisenberg* che è più adatto alla TQC:

$$(7.2) \quad \Delta n \Delta \varphi \geq \hbar$$

dove n è il numero di quanti del campo in un certo modo vibrazionale, φ è la fase dell'oscillatore ed $\hbar$ la costante di Planck. Quando l'incertezza in fase-ampiezza è molto grande, ad esempio nel caso di un numero finito di quanti scarsamente interagenti, otteniamo *stati incoerenti*, tra loro *slegati* e descrivibili attraverso vettori di stato individuali. Nella TQC possono aversi invece situazioni dove le condizioni al contorno del problema permettono a un sistema con un alto valore di Δn di avere una *finestra di fase* $\Delta \varphi$ molto *stretta*. In questo caso possono manifestarsi fenomeni di *coerenza di fase* che giocano un ruolo fondamentale in molti fenomeni tipici della materia condensata. In altre parole, c'è la possibilità di un *rinforzo* che permette fenomeni di *correlazione long-range* tra oscillatori (*phase-linking*), informazione distribuita che non trasporta energia ma *sintonizza* il sistema su un *processo cooperativo* che è essenzialmente il modo in cui nella TQC si manifesta la *proprietà peculiare dei sistemi quantistici, la non-località*. Agli stati coerenti si oppone continuamente il *gioco* delle fluttuazioni e del rumore termico, l'entropia e il disordine, ed è possibile dimostrare che questa forma di correlazione quantistica è piuttosto

fragile, e va incontro alla *decoerenza*. Gli *esperimenti EPR-Bell*, sui quali si è molto discusso a proposito dei problemi interpretativi e delle *presunte stranezze* della fisica quantistica, appaiono dunque situazioni estremamente semplificate, in cui una descrizione che utilizza la MQ è sufficiente, ed è basata sulla possibilità di estendere la coerenza sull'ordine dei domini ottici degli apparati sperimentali, scommessa ardua sulla quale si basa la realizzazione dei *computer quantistici*. Ma la fondamentale *unità quantistica dell'universo* si manifesta in tutta la sua ricchezza in ben altri domini, ed è compito della futura fisica dello stato solido indagare il ruolo della non-località nei processi sub-nucleari, nucleari, atomici, molecolari, fino alle possibilità ipotizzate per i *processi viventi*. In generale il rumore termico distrugge la coerenza a livello macroscopico, ed è il motivo per cui il mondo classico ci appare *locale* e *separabile*. Fenomeni di correlazione quantistica caratterizzano i fenomeni di *superfluidità*, *super-conduttività*, *super-radianza*, e la *condensazione di Bose-Einstein*, processi dove un gran numero di particelle mostra un *fitto* comportamento cooperativo descritto da un unico vettore di stato. In generale, l'intera fisica della materia condensata è ricca di fenomeni di questo tipo, dove si assiste alla formazione di *domini di coerenza*.

La domanda fondamentale sulla morfogenesi equivale allora a chiedersi sotto quali condizioni la *disponibilità energetica* del vuoto quantistico permette la formazione di strutture. E' possibile dimostrare che se vivessimo in un mondo di oggetti puramente classici, dove vale soltanto la *statistica di Maxwell-Boltzmann*, sarebbe *impossibile* ottenere delle strutture. Questo è abbastanza intuitivo, se pensiamo al *punto materiale newtoniano*, e dunque non sarà necessario entrare in questioni di fisica statistica e di limiti termodinamici. Ma abbiamo ricordato che la statistica classica è un limite delle *statistiche di Bose-Einstein* e di *Fermi-Dirac*, e queste sono legate alla *struttura topologica a stati di vuoto multipli* della teoria quantistica dei campi. Tutto ciò rende possibile un meccanismo per la formazione di strutture noto come *rottura spontanea di simmetria* (SSB, *Spontaneous Symmetry Breaking*). Questo è un processo di *transizione di fase del secondo ordine* che al variare di un *parametro critico* fa passare *tutto* il sistema da uno stato a un altro in modo estremamente veloce, con una radicale *riorganizzazione energetica* che permette la comparsa di un ordine completamente nuovo, dotato di *minore simmetria* e dunque di *maggior organizzazione*, descritto dal cosiddetto *gruppo di contrazione di Inou-Wigner*. Alcuni esempi di questi processi sono l'espansione

inflazionaria con liberazione di energia del vuoto in cosmologia, la condensazione di Bose-Einstein, il passaggio da stato paramagnetico a ferromagnetico e la comparsa di onde di spin, la transizione da conduttore a super-conduttore, le linee vorticali nei superfluidi, etc. In tutti questi casi si verifica un passaggio da una fase *meno strutturata e omogenea* ad un'altra *più strutturata e meno omogenea*, e il sistema compensa i vincoli del nuovo ordinamento energetico sotto forma di *correlazioni long-range* legate in qualche misura alla nuova *rigidità* del sistema. In generale si parla di *eccitazioni collettive* che vengono denominate *bosoni di Goldstone*. Dal punto di vista formale questi sono l'espressione delle relazioni tra dinamica e simmetria, e fisicamente possono essere considerati come forme *quasi-particellari (like-particles)* in cui *il disordine residuo si localizza* in strutture stabili che sono manifestazioni peculiari della nuova struttura e della sua dinamica. L'uso della terminologia 'quasi-particelle' si riferisce proprio al fatto che questi modi vibrazionali long-range sono tutt'uno con la nuova struttura del sistema, e *viceversa* l'organizzazione energetica del sistema si esprime proprio attraverso l'emergenza di questo tipo di strutture che, essendo quantizzate, assumono comportamenti particellari. Un esempio classico sono le *eccitazioni fononiche* nei cristalli. Un fonone non è deducibile dalla struttura *statica* del cristallo inteso come reticolo di atomi *impacchettati uno ad uno*, ma è l'espressione della *dinamica globale* caratteristica del reticolo!

Queste poche osservazioni bastano a far comprendere in che senso *i sistemi con SSB sono amplificatori esponenziali di informazione*, e più in generale che i *Quantum Fields* permettono di costruire un mondo ricco di strutture complesse grazie alla struttura del vuoto. La TQC appare dunque come il *regno naturale* dei processi di emergenza intrinseca, e c'è un'intensa attività per *sviluppare la TQC come linguaggio generale per la teoria dei sistemi e dell'emergenza*.

La TQC offre dunque un *framework* molto generale per lo studio dei sistemi logicamente aperti, in grado di esibire processi di emergenza intrinseca come transizioni di fase SSB. Inoltre lo spazio di Hilbert ben si presta a trattare sistemi che presentano sia aspetti continui che discreti, tipico di ciò che abbiamo definito computazione naturale. Non bisogna dimenticare che l'informazione in natura è sempre connessa a processi di scambio di energia, e questi *sono* processi quantistici. Queste caratteristiche possono suggerire una relazione non-banale tra TQC e processi cognitivi, che è in effetti al centro di gran parte della ricerca attuale. Con 'non-

banale' intendiamo qui riferirci alle caratteristiche cosiddette *bio-inspired*. Se supponiamo, senza perdere appunto l'ispirazione biologica, che la relazione tra mente e cervello è in qualche modo analoga a quella esistente tra la gamma delle emergenze fenotipiche e il progetto genotipico, siamo indotti a utilizzare una vasta gamma di modelli il cui valore dipende dalla questione studiata. In alcuni casi sarà utile considerare le attività cognitive in stretta relazione con la struttura del cervello in una relazione gerarchica (cosa che appare indispensabile nello studio delle patologie e delle lesioni), ma in altri casi può essere più utile una concezione eterarchica, che metta l'accento sulle differenze tra i due sistemi e sulle emergenze tipiche delle procedure cognitive. Del resto abbiamo visto che i modelli PDP ed i SST sono ugualmente lontani dalla reale architettura del cervello, e il vantaggio principale dei modelli connessionistici non consiste tanto, come si ritiene, in una sorta di isomorfismo geometrico, ma piuttosto nella capacità di cogliere gli aspetti di apertura da cui dipende la possibilità dei fenomeni di emergenza. Anzi, l'*abilità* delle reti neurali nella formazione di pattern come minimi parziali del paesaggio energetico che si struttura in relazione a una determinata sequenza di input può essere vista come una *pallida imitazione* della capacità dei sistemi TQC di esibire una grande varietà di strutture legata all'esistenza di un vuoto multiplo! Infine, da un punto di vista generale, il vantaggio di utilizzare la fisica quantistica nel modellare un sistema cognitivo sarebbe quello, non indifferente, di tenere conto *ab inizio*, dei vincoli e delle possibilità insite nella descrizione della *physis* più potente e generale disponibile.

Storicamente è possibile individuare diversi tipi di approccio al problema mente/cervello in relazione alla fisica quantistica. Alcuni dei fondatori della teoria hanno ipotizzato che l'*osservatore cosciente* può avere un ruolo nello spezzare la *catena di von Neumann*, ossia nel fissare il collasso del vettore di stato, distruggendo la coerenza degli stati sovrapposti ed ottenendo così le caratteristiche locali del mondo classico (dove *i gatti sono o vivi o morti*, per riferirci al famoso argomento di Schrödinger). Eugene Wigner è sicuramente il rappresentante più famoso di questa posizione. E' probabilmente un effetto del prestigio del grande teorico se questa *ambiziosa* teoria dell'osservatore ha avuto un qualche credito. Infatti è *assolutamente arbitrario* nonchè *estremamente vago* fissare il limite tra non-locale e locale in questo modo. Anche se in forme diverse, ogni interpretazione della teoria quantistica provvede a stabilire questo confine tramite strumenti essenzialmente termodinamici, legati alle relazioni tra

microfisica e macrofisica. Esempi in questa direzione sono i lavori di Loinger-Prosperi, la teoria di Ghirardi-Rimini-Weber, la teoria dinamica di Bohm-Bub, le posizioni dell'ultimo Bohm relative ai range energia-tempo di collasso; tutte queste teorie forniscono quello che la posizione di Wigner non può dare, ossia la possibilità di test sperimentali. Più recentemente, H. Stapp ha ripreso la posizione di Wigner capovolgendo però il rapporto tra osservatore e collasso. Per Stapp è *il collasso del vettore di stato che produce la coscienza*, e quest'ultima è l'immagine isomorfa del collasso a livello degli eventi cerebrali. Cambiando l'ordine dei fattori il risultato epistemologico non cambia, e per quanto Stapp abbia fornito un apparato matematico interessante per argomentare la sua tesi, resta il fatto che nei suoi lavori la coscienza è definita in modo piuttosto confuso. Le posizioni del tipo Wigner-Stapp, è stato detto, pretendono di risolvere un mistero con un'altro mistero. Guardate in una prospettiva storica esse riflettono la difficoltà di accettare un postulato così anti-intuitivo come quello del collasso, maturate in un periodo in cui si era ben lontani anche solo dal poter immaginare di studiare in laboratorio effetti quantistici macroscopici (*mettere sotto osservazione i gatti di Schrödinger*). La teoria di Stapp ha però qualche merito nell'aver riportato l'attenzione sulla possibilità di identificare nei processi cerebrali degli effetti quantistici genuini. In effetti, come ha scritto E. H. Walker, sarebbe piuttosto strano che *non* si trovasse traccia di processi quantistici nel cervello, e più in generale nelle strutture viventi, considerato il loro peculiare range bio-chimico! Uno dei primi tentativi in questo senso è dovuto al padre della seconda cibernetica, Heinz von Foerster, che in un saggio del 1943, pubblicato in lingua tedesca, riuscì a correlare il *parametro di dimenticanza*, che si stabilisce in base agli studi sperimentali sulla memoria, con la *costante di decadimento* delle macromolecole organiche durante una transizione quantica. Si stabilisce così una relazione interessante tra teoria quantistica e informazione, seppure non legata direttamente a processi di sovrapposizione di stati.

La classe di modelli più fecondi nasce con l'ormai classico lavoro di L. M. Ricciardi e Hiroo Umezawa sul *modello quantistico del cervello costruito in analogia con la fisica dei sistemi condensati* (Ricciardi-Umezawa, 1967). Ricciardi e Umezawa partono proprio dalla constatazione che le reti neurali sono modelli molto astratti e potenti per studiare alcuni specifici aspetti dei processi di riconoscimento e apprendimento, e sicuramente in grado di indicare una via *affidabile* per la costruzione di *dispositivi intelligenti*, ma il confronto con il cervello appariva già allora a dir poco

impari. Meglio cercare di concentrare la propria attenzione alla capacità della mente di costruire strutture stabili o semi-stabili intermezze da regimi transitori senza un diretto riferimento alla complessità cerebrale. La teoria si situa dunque a un livello intermedio tra i sistemi rappresentazionali, perchè non pretende di avere alcun isomorfismo *ingenuo* con il cervello e con la ricca varietà di fenomeni bio-chimici che ne regola l'attività, e quelli connessionistici, perchè si propone un'analisi della mente in termini dinamici, tenendo conto dell'interazione con l'ambiente. A questo scopo il modello fa riferimento alla TQC come unico strumento formale in grado di offrire un ricco scenario di comportamenti che si possono osservare in un sistema quantistico al variare di un certo parametro d'ordine. Bisogna qui evidenziare che l'analisi di Ricciardi-Umezawa (*RU model*) si basa sull'utilizzo dei metodi formali della TQC nello studio degli stati condensati, dunque in un *volume finito*. Alcune valutazioni numeriche possono suggerire la plausibilità della scelta dei due studiosi. E' noto che il numero dei neuroni è stimato intorno a 10^{10} unità, e ogni neurone sta almeno in un rapporto di 1 a 10 con le cellule gliali, che sono strettamente connesse con l'attività dei neuroni. Utilizzando il limite di Bekenstein per l'analisi di un numero di stati quantici in un volume finito, anche utilizzando approssimazioni grossolane, otteniamo per il numero di stati in cui può trovarsi un essere umano la cifra enorme di $10^{10 \exp 45}$, e un numero di transizioni al secondo dell'ordine di 10^{53} . Queste cifre mostrano che considerare un volume finito nulla toglie alla *ricchezza* dell'analisi di RU, che erano concentrati a mostrare in modo esemplare come la variazione di un numero *finito* di parametri d'ordine che corrispondono all'interazione con l'ambiente, porta a una serie di processi SSB in grado di garantire la manifestazione di un gran numero di modi vibrazionali di tipo bosonico in un sistema finito.

L'introduzione dei parametri d'ordine equivale a semplificare le relazioni tra sistema e ambiente. In particolare, vuol dire mettere l'accento sull'influenza dell'ambiente sul sistema piuttosto che il contrario. E' bene evidenziare questi aspetti del modello RU per comprendere meglio gli sviluppi più recenti. Le variabili dinamiche macroscopiche usate vengono chiamate *corticoni*, e non hanno alcuna relazione diretta con i costituenti del cervello. I processi SSB producono degli *slow-mode* denominati *simmetroni*, le cui caratteristiche si prestano alla costruzione di un modello generale del funzionamento della memoria. Il *nuovo livello energetico emerso è di per sé stesso una forma fisica di codifica dell'informazione* poichè individua l'effetto

che ha avuto la variazione del parametro d'ordine sul sistema, inducendo una rottura di simmetria connessa con la densità bosonica dello stato. La stabilità dello stato e la sua coerenza su larga scala assicura le caratteristiche fondamentali della *memoria a lungo termine*, ossia la persistenza nel tempo e le sue caratteristiche globali, in accordo con il *modello ologrammatico* di K. Pribram. La *memoria a breve termine* è invece associata ai *fast-mode* che si producono durante la transizione di fase; questo tipo di ricordi sono dunque connessi a stati meta-stabili del sistema. RU suggeriscono una relazione tra il richiamo selettivo e sequenziale dei ricordi basata sul fatto che ogni livello è, in sostanza, un'informazione codificata, e dunque ogni livello esclude l'altro, cosa che corrisponde all'impossibilità di più sequenze associative allo stesso tempo. Questo fatto ha conseguenze notevoli che qui possiamo soltanto accennare. Uno *slow-mode* è per sua natura uno *stato coerente con caratteristiche continue*, il che rispecchia in modo naturale il trattamento *globale* delle immagini che la nostra mente è in grado di fare grazie alle risorse retinotopiche e cocleotopiche. D'altra parte il processo di richiamo ed elaborazione è intrinsecamente sequenziale, cosa che offre la possibilità di indagare una relazione tra le associazioni della memoria visiva e l'elaborazione sequenziale della produzione linguistica. Un punto interessante nel modello di RU consiste invece nella *difficoltà di descrivere i processi di decadimento della memoria*, se non come *pura distruzione* di uno stato provocato dall'instaurarsi di un nuovo livello. C'è insomma la possibilità di *sovrascrivere* i ricordi ma non di una loro *gradazione* e *sovrapposizione*. Questa è una *conseguenza dell'aver limitato l'apertura del modello*, cosa che porta a un'eccessiva *rigidità* dei modi eccitati. Una serie di sviluppi recenti hanno risolto il problema adottando pienamente la *ricchezza* descrittiva della TQC e considerando la mente un *sistema aperto tout-court*.

Nel seguito adotteremo l'analisi di G. Vitiello (1995). In questo modello il gioco complesso dei processi SSB dovuti allo scambio continuo sistema-ambiente, non riducibile a un'analisi semplificata in termini di parametri d'ordine, porta all'immagine di un sistema che *vive* tramite una serie continua di transizioni di fase e dunque di nuovi livelli emergenti. L'introduzione della dissipazione in modo esplicito equivale a considerare attività cerebrale e ambiente come *un unico sistema*, dando un senso fisico nuovo e più radicale all'espressione *embodied mind*. In questo modo la stessa coerenza matematica del formalismo rende possibili infiniti stati coesistenti e non-distruttivi per codificare l'informazione, L'arrivo di

nuova informazione non produce necessariamente la distruzione di quella precedentemente immagazzinata, ma piuttosto, com'è naturale aspettarsi da un sistema dissipativo, una sua *riorganizzazione*. Tra l'altro questo introduce in modo naturale una freccia del tempo legata all'attività mentale. C'è un aspetto suggestivo del formalismo canonico della dissipazione quantistica che è interessante qui accennare brevemente. Nella trattazione matematica giocano un ruolo importante due *set* di modi che denoteremo genericamente con A_k e con $\tilde{A}_k$, con k grado di libertà del campo. Queste due strutture sono profondamente interconnesse, e in particolare $\tilde{A}$ può essere considerato il *time-reversed mirror image* del modo A , ossia l'inestricabile *riflesso* sulla struttura di A dell'ambiente. Il sistema $\tilde{A}$ è stato denominato da Vitiello il *Doppio* di A , o il suo *Sosia*, e ha suggerito che possa svolgere un ruolo significativo nello spiegare i processi coscienti come una speciale proprietà di *auto-interazione* del sistema con sè stesso (Vitiello, 2001). Detto in altro modo, è possibile dire, in accordo con Maturana e Varela, che l'attività mentale è una continua *produzione del mondo* che origina dalla natura irreversibile e dissipativa delle nostre interazioni con l'ambiente. Uno sviluppo recente di questi modelli è centrato sulla costruzione di analogie con le reti neurali (Pessa-Vitiello, 1999).

Bisogna notare esplicitamente che il modello RU ha fatto uso di processi quantistici in relazione alla possibilità di descrivere la dinamica del sistema. Gli stati coerenti e la non-località sottese a questa classe di modelli *non suggerisce necessariamente* la possibilità di trovare processi quantistici nel cervello *reale*. D'altra parte, lo sviluppo recente di questa modellistica ha portato ad alcune valutazioni quantitative che ripropongono la domanda in modo da incontrarsi con le attuali possibilità osservative nelle neuroscienze. Ad esempio, nel modello di Del Giudice e collaboratori (Del Giudice et al., 1985, 1986) un ruolo chiave è svolto dalla *rottura della simmetria rotazionale del dipolo*, e poichè gran parte delle macromolecole significative sono dipolari, è possibile *ipotizzare una ricca fenomenologia genuinamente quantistica interna al cervello*. I bosoni di Goldstone prodotti da questo tipo di SSB hanno infatti una piccolissima massa (*Higgs-Kibble mechanism*), possono condensare a temperature biologicamente rilevanti producendo stati coerenti (*evanescent photons*) e portano a una peculiare *filamentazione* del campo elettrico. Altri stati coerenti derivano dal fenomeno della *super-radianza*, che è analogo a un'emissione tipo laser, ma che non ha bisogno nè di cavità nè di pompaggio, ma soltanto di temperature e/o densità critiche. In questo caso un gruppo di molecole entra in un dominio di coerenza

generando un forte campo elettromagnetico piuttosto stabile, e in grado di comunicare con domini dello stesso tipo su distanze molto grandi. Jibu e Yasue (1997) hanno studiato la possibilità che domini di coerenza di questo tipo dentro e fuori la membrana neurale nelle reti dendritiche possano formare una *giunzione Josephson* in grado di produrre il caratteristico *effetto tunnel*. Lo stesso significato delle misure EEG potrebbe essere rimesso in discussione. L'attività oscillatoria misurata viene solitamente intesa come una (quasi) media statistica dell'attività degli *spikes* neuronali. In uno dei suoi ultimi lavori Varela (Varela et al., 2001) ha osservato l'esistenza di forti connessioni polisinaptiche tra regioni lontane della corteccia quando i potenziali EEG si trovano in coerenza di fase. E' possibile pensare che *siano proprio queste forme di coerenza* a stimolare la modalità di scarica tra gruppi neuronali. Non è nostra intenzione soffermarci su queste proposte. I modelli proposti sono molti, e piuttosto complessi; inoltre la possibilità di verifica sperimentale appare oggi *plausibile* ma *difficilmente realizzabile*.

In un lavoro recente Penrose et al. (Marshall, W., Simon, C., Penrose, R. & Bouwmeester, D., 2003), hanno proposto un apparato sperimentale di interferometria per osservare effetti quantistici macroscopici relativi a un oggetto della grandezza all'incirca di un globulo rosso, ossia dell'ordine di 10^{14} atomi, o 10^{-12} Kg. Lo stesso R. Penrose, in collaborazione con S. Hameroff, ha concentrato il suo interesse sui *microtubuli*, formazioni proteiche di struttura cilindrica costituite da fibre longitudinali che costituiscono il citoscheletro delle cellule, come possibili luoghi di accadimenti quantistici nell'organismo, dando il via a una ricca serie di proposte teoriche e sperimentali. E' evidente che se questi fenomeni dovessero trovare conferma la nostra conoscenza non soltanto della mente ma di tutti i processi viventi acquisterebbe un nuovo livello di dettaglio, e si potrebbe ipotizzare la nascita di una biologia quantistica o di una neuroscienza quantistica. Quello che possiamo fare qui è considerare questa possibilità e chiederci se all'interno del quadro già piuttosto complesso della computazione naturale dobbiamo considerare anche forme di computazione quantistica.

E' noto da tempo che gli organismi sono sensibili all'azione di brevi sequenze di fotoni, come appare dagli ormai classici esperimenti sulla fotosensibilità della retina. E' stato dimostrato che nell'essere umano questo meccanismo è molto selettivo e ha una soglia più alta che in altri animali, cosa che impedisce di perdere l'immagine in una sorta di *rumore visivo*. La domanda sulla computazione quantistica non riguarda dunque

l'effetto di singoli eventi quantici sul sistema computante, perchè in natura è piuttosto facile individuare soglie biochimiche sensibili a questi effetti, ma piuttosto la modalità di questa computazione in relazione al fenomeno dell'*entanglement*, ossia dell'esistenza di *stati sovrapposti* e di *correlazioni non-locali*. Per indicare questa forma di immagazzinamento ed elaborazione dell'informazione introduciamo *l'unità elementare dell'informazione quantistica, il qbit*, che si definisce semplicemente come *lo stato di sovrapposizione di due bit classici 0 ed 1*, e si scrive:

$$1\text{qbit} = |0\rangle + |1\rangle$$

Per qualche misteriosa ragione, gli studiosi del *quantum-computing* hanno sempre sostenuto che la potenza computazionale di un dispositivo quantistico è *al più pari* alla Turing-computabilità, a parte probabilmente una maggiore efficienza. Un argomento generale a favore di questa tesi è che un insieme *stocastico* (la probabilità quantistica ben si presta ad essere interpretata in questo senso) può sempre essere considerata come una famiglia di *insiemi deterministici*. Questo vuol dire che un computer quantistico equivale a una coppia di computer classici che operano in una sorta di parallelismo. In realtà esistono anche argomenti a favore di una *potenza super-Turing* del quantum computing, basate proprio sulle analogie formali con le reti neurali. Questi lavori mostrano infatti che una rete neurale deve avere una serie di vincoli piuttosto complessi per poter fornire prestazioni paragonabili al *quantum brain* con dissipazione in relazione alla capacità di memoria. Sotto parecchi punti di vista, la struttura matematica della fisica quantistica appare assai più ricca di quella della maggior parte dei sistemi dinamici, che appaiono come limiti classici della TQC, e dunque argomentazioni tipo Siegelman-Sontag suggeriscono che la potenza del quantum computing si situi *almeno* a un livello super-Turing. In ogni caso, come mostra l'esempio delle frequenze EEG, ancora una volta non è tanto la *mera* potenza di calcolo in gioco, ma le *modalità specifiche* di questo, e la sua funzione nella relazione sistema-ambiente. Includere nella computazione naturale anche il livello quantistico significa dunque considerare un nuovo range di *codici naturali emergenti*.

Osservatorio, Cracovia
Settembre 2003

Ringraziamenti

Sono molte le persone a cui devo la mia (relativa) apertura logica. Voglio qui fare un ringraziamento particolare a Marcello Barbieri, per avermi suggerito l'importanza del concetto di codice nella costruzione di una teoria generale dei sistemi, e per l'amicizia e la gentilezza dimostrate durante la preparazione di questo lavoro.

BIBLIOGRAFIA

Paragrafo 1

- A. Newell, *Unified Theories of Cognition*, Harvard Univ. Press, Cambridge, MA, 1990.
- J. R. Anderson, *Rules of the Mind*, Erlbaum, NJ, 1993.
- Z. Pylyshyn, *Computation and Cognition*, Cambridge, MIT Press, 1986.
- J. Fodor & Z. Pylyshyn, *Connectionism and Cognitive Architecture*, in *Cognition*, 28(3), 1988.
- S. J. Russel & P. Norvig, *Intelligenza artificiale. Un approccio moderno*, UTET, Torino, 1998.
- N.J. Nilsson, *Intelligenza artificiale*, Apogeo, Milano, 2002.
- V. Somenzi & R. Cordeschi (a cura di), *La filosofia degli automi. Origini dell'intelligenza artificiale*, Bollati-Boringhieri, Torino, 1994.
- P. McCorduck, *Storia dell'intelligenza artificiale*, Muzzio., Padova, 1987.
- H. Gardner, *La nuova scienza della mente. Storia della rivoluzione cognitiva*. Feltrinelli, Milano, 1988.
- N. A. Chomsky, *A Minimalist Program for Linguistic Theory*, in *MIT Working Papers in Linguistics*, Cambridge, MA, 1992.
- G. A. Miller, E. Galanter, K.H. Pribram, *Piani e struttura del comportamento*, Franco Angeli, Milano, 1993.
- J.L. McClelland & D.E. Rumelhart (eds.), *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, 2 voll., MIT Press, Cambridge, MA, 1986.
- J.A. Anderson, E. Rosenfeld, A. Pellionisz (eds.), *Neurocomputing: Direction for Research*, Vol.1, 1991; vol.2, 1993, MIT Press.
- I. Aleksander & H. Morton, *Introduction to Neural Computing*, Thompson Int. Press, London, 1995.

- J. Anderson & E. Rosenfeld (eds.), *Talking Nets. An Oral History of Neural Network*, MIT Press, Cambridge, MA, 1998.
- T. Kohonen, *Self-Organization and Associative Memory*, Springer, Berlin, 1984.
- S. Grossberg (ed.), *Neural Networks and Natural Intelligence*, MIT Press, Cambridge MA, 1988.
- D. S. Levine, *Introduction to Neural and Cognitive Modeling*, Erlbaum, NJ, 2000.
- E. Kandel & J. Schwarz, *Principi di neuroscienze*, CEA, Milano, 1994.
- E. Pessa, *Reti neurali e processi cognitivi*, Di Renzo, Roma, 1993.
- D. Floreano & C. Mattiussi, *Manuale sulle reti neurali*, Il Mulino, Bologna, 2002.
- D. Marconi, *Filosofia e scienza cognitiva*, Laterza, Bari, 2001.
- M. Di Francesco, *Introduzione alla filosofia della mente*, Carocci, Roma, 1996.
- M. Di Francesco, *La coscienza*, Laterza, Bari, 2000.
- A. Clark, *Microcognizione. Filosofia, scienza cognitiva e reti neurali*, Il Mulino, Bologna, 1994.
- A. Clark, *Dare corpo alla mente*, McGraw-Hill, Milano, 1999.
- J. A. Fodor, *La mente modulare. Saggio di psicologia delle facoltà*, Il Mulino, Bologna, 1999.
- J. A. Fodor, *Concetti. Dove sbaglia la scienza cognitiva*, McGraw-Hill, Milano, 1999.
- P. M. Churchland, *La natura della mente e la struttura della scienza. Una prospettiva neurocomputazionale*, Il Mulino, Bologna, 1992.
- P. M. Churchland & T. J. Sejnowski, *Il cervello computazionale*, Il Mulino, Bologna, 1995.
- A. Karmiloff-Smith, *Oltre la mente modulare. Una prospettiva evolutiva sulla scienza cognitiva*, Il Mulino, Bologna, 1995.
- A. M. Borghi & T. Iachini (a cura di), *Scienze della mente*, Il Mulino, Bologna, 2002.
- D. Parisi, *Mente. I nuovi modelli della vita artificiale*, Il Mulino, Bologna, 1999.
- D. Parisi, *Simulazioni. La realtà rifatta nel computer*, Il Mulino, Bologna, 2001.
- M. Minsky, *La società della mente*, Adelphi, Milano, 1989.
- E. Pessa & M. P. Penna, *Manuale di scienza cognitiva. Intelligenza artificiale classica e psicologia cognitiva*, Laterza, Bari, 2000.
- E. Pessa & M. P. Penna, *Introduzione alla psicologia connessionistica*, Di Renzo, Roma, 1993.

- H. Putnam, *Mente, linguaggio, realtà*, Adelphi, 1987.
- L. Valente Torre, *L'evoluzione dell'intelligenza in Jean Piaget. Aspetti strutturali e funzionali*, Bollati-Boringhieri, Torino, 1993.
- E. Reed & R. Jones (eds.), *Selected Essays of James J. Gibson*, Erlbaum, NJ, 1982.
- G.M. Edelman, *Sulla materia della mente*, Adelphi, Milano, 1993.
- A.R. Damasio, *L'errore di Cartesio*, Adelphi, Milano, 1995.
- J. Ledoux, *Il Sé sinaptico*, Raffaello Cortina, Milano, 2002.
- H. Maturana & F. Varela, *L'albero della conoscenza*, Garzanti, Milano, 1992.
- F. Varela, E. Thompson, E. Rosch, *La via di mezzo della conoscenza*, Feltrinelli, Milano, 1992.

Paragrafo 2.

- C.E. Shannon & W. Weaver, *La teoria matematica delle comunicazioni*, Etas Kompass, Milano, 1972.
- P. Smolensky, *Connectionist AI, and the Brain*, in *Behavioural and Brain Sciences*, 11(1), 1988.
- P. Smolensky, *Il connessionismo tra simboli e neuroni*, Marietti, Genova, 1992.
- R. Port & T. van Gelder, *Mind as Motion: Explorations in the Dynamics of Cognition*, MIT Press, Cambridge MA, 1995.
- E. Thelen & L. Smith, *A Dynamic System Approach to the Development of Cognition and Action*, MIT Press, Cambridge MA, 1994.
- R. Beer & J. Gallagher, *Evolving Dynamical Neural Networks for Adaptive Behaviour*, in *Adaptive Behaviour*, 1(91) 1992.
- S. Kelso, *Dynamic Patterns*, MIT Press, Cambridge MA, 1995.
- T. van Gelder, *Connectionism and Dynamical Explanation*, in *13th Annual Conference of the Cognitive Science Society*, Chicago, 1991.
- K. Matsumoto & I. Tsuda, *Noise-Induced Order-Complexity. Theoretical Digression*, in Y. Kuramoto (ed.), *Chaos and Statistical Methods*, Springer, 1984.
- I. Tsuda, *Dynamic Link of Memory-Chaotic Memory Map in NonEquilibrium Neural Networks*, in *Neural Networks*, 5(313), 1992.
- W.J. Freeman, *Neurodynamics. An Exploration of Mesoscopic Brain Dynamics*, Springer, Berlin, 2000.
- W.J. Freeman, *How Brains Make Up Their Minds*, Columbia Univ. Press., NY, 2001.

Paragrafo 3

- D. Marr, *Artificial Intelligence: A Personal View*, in *AI*, 9(37), 1977.
- D. Marr, *Vision*, Freeman, San Francisco, CA, 1982.
- Ignazio Licata (a cura di), *Informazione e Complessità*, Andromeda, Bologna, 1998.
- G. Gestri, *Teoria dell'informazione*, ETS, Pisa, 1991.
- R. Shaw, *Strange Attractors, Chaotic Behaviour and Information Flow*, in *Zeit. Für Naturf.*, 36A(80), 1981.
- Hao Bai-Lin (ed.) , *Chaos*, World Scientific, Singapore, 1984.
- A. Scott, *Non-Linear Science: Emergence and Dynamics of Coherent Structures*, Oxford Univ. Press, Oxford, 1999.
- H. Haken, *Synergetics*, Springer, 1977.
- M. Bushev, *Synergetics: Chaos, Order, Self-Organization*, World Scientific, London, 1994.
- G. Nicolis & I. Prigogine, *Self-Organization in NonEquilibrium Systems from Dissipative Structures to Order Through Fluctuations*, Wiley, NY, 1977.
- I. Prigogine & I. Stengers, *La nuova alleanza. Metamorfosi della scienza*, Einaudi, Torino, 1999.
- R. Thom, *Stabilità Strutturale e Morfogenesi*, Einaudi, Torino, 1981.

Paragrafo 4

- E. Casari, *Introduzione alla logica*, UTET, Torino, 1997.
- G. Lolli, *Introduzione alla logica formale*, Il Mulino, Bologna, 1991.
- G. Lolli, *Filosofia della matematica. L'eredità del Novecento*, Il Mulino, Bologna, 2002.
- B. Russell, *I principi della matematica*, Newton-Compton, Roma, 2000.
- M. Davis, *Computability and Unsolvability*, Dover, NY, 1982.
- M. Davis, *Il Calcolatore Universale*, Adelphi, Milano, 2003.
- M. Minsky, *Computation. Finite and Infinite Machines*, Prentice-Hall, NY, 1972.
- M. Arbib, *Brains, Machines and Mathematics*, Springer, 1987.
- G. Tamburrini, *I matematici e le macchine intelligenti*, Bruno Mondadori, Milano, 2002.
- J.W. Dawson Jr, *Dilemmi logici - La vita e l'opera di Kurt Gödel*, Bollati-Boringhieri, Torino, 2001.
- S.G. Shanker (a cura di), *Il teorema di Gödel*, Franco Muzzio Editore, Padova, 1991.
- Hao Wang, *Dalla Matematica alla Filosofia*, Bollati-Boringhieri, Torino, 1984.

- A. Hodges, *Storia di un enigma. Vita di Alan Turing 1912-1954*, Bollati-Boringhieri, Torino, 1991.
- A.M. Turing, *On Computable Numbers with an Application to the Entscheidungsproblem*, in *Proc. London Math. Soc.*, 42(230), 1936.
- R.O. Gandy, *Church's Thesis and Principles for Mechanisms*, in J. Berwise et al (eds), *The Kleene Symposium*, North-Holland, Amsterdam, 1980.
- G. Chaitin, *Information, Randomness and Incompleteness*, World Scientific, Singapore, 1990.
- J.D. Bekenstein, *Energy Cost of Information Transfer*, in *Phys. Rev. Lett.* 46(623), 1981.
- F.J. Tipler, *Cosmological Limits on Computation*, in *Int. Jour. of Theor. Phys.*, 25(617), 1986.
- J.R. Lucas, *Mind, Machines and Gödel*, in *Philosophy*, 36, 1961.
- R. Penrose, *La mente nuova dell'imperatore*, Rizzoli, Milano, 1992.

Paragrafo 5

- M.B. Pour-El & I. Richards, *Computability in Analysis and Physics*, Springer, 1989.
- R. Geroch & J.B. Hartle, *Computability and Physical Theories*, in *Found. of Phys.*, 16(533), 1986.
- S. Franklin & M. Garzon, *Neural Computability*, in O.M. Omidvar(ed), *Progress in Neural Networks*, Vol.1, Norwood, NJ, 1990.
- S. Franklin & M. Garzon, *Computation on Graphs*, in O.M. Omidvar(ed), *Progress in Neural Networks*, vol. 2, Norwood, NJ, 1990.
- M. Stannett, *X-Machines and The Halting Problem: Building a Super-Turing Machine*, in *Formal Aspects of Computing*, 2(331), 1990.
- O. Bournez & M. Cosnard, *On The Computational Power and Super-Turing Capabilities of Dynamical Systems*, in *Tech. Rep. 95-30*, Ecole Normale Sup. de Lyon, Lab. de l'Informatique du Parallelisme, 1995.
- H.T. Siegelman & E.D. Sontag, *Analog Computation via Neural Networks*, in *Theor. Comp. Sciences*, 131(331), 1994.
- H.T. Siegelman & E.D. Sontag, *On The Computational Power of Neural Networks*, in *J. of Comp. and Syst. Science*, 50(132), 1995.
- H.T. Siegelman, *The Simple Dynamics of Super-Turing Theories*, *Tech. Rep. 94-NN-1*, Technion, Haifa, Israel.
- H.T. Siegelman, *Computation Beyond the Turing Limit*, in *Science*, 268(545), 1995.

- H.T. Siegelman, *Neural Networks and Analog Computation. Beyond the Turing Limit*, Birkhäuser, Boston, 1999.
- W. Maass & E.D. Sontag, *Analog Neural Nets with Gaussian or Other Common Noise Distribution Cannot Recognize Arbitrary Regular Languages*, in *Neural Computation*, 11(771), 1999.
- B. MacLennan, *Transcending Turing Computability*, Tech.Rep.UT-CS-01-473, Dep. of Comp. Science, Univ. of Tennessee, Knoxville, 2001.
- M.V. Volkenshtein, *Complementarity, Physics and Biology*, in *Soviet Phys. Usp.* 31(140), 1988.
- B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical System Approach to Machine Intelligence*, Prentice Hall, Englewood Cliffs, 1992.
- G. Spezzano & D. Talia, *Calcolo parallelo, automi cellulari e modelli per sistemi complessi*, Franco Angeli, Milano, 1999.
- R. Serra & G. Canarini, *Sistemi complessi e processi cognitivi*, Calderini, Bologna, 1994.
- C. Emmeche, *Il giardino nella macchina. La nuova scienza della vita artificiale*, Bollati-Boringhieri, 1996.
- S. Wolfram, *Theory and Applications of Cellular Automata*, World Scientific, Singapore, 1986.
- S. Wolfram, *A New Kind of Science*, WolframMedia, 2002.
- S. Forrest(ed) ,*Emergent Computation*, North-Holland, Amsterdam, 1990.
- J.P. Crutchfield, *The Calculi of Emergence: Computation, Dynamics and Induction*, in *Physica D*75(11), 1994.
- P. Bak, C. Tang, K. Wiesenfeld, *Self-Organized Criticality*, in *Phys.Rev.*A38(364)1988.
- P. Bak, *How Nature Works. The Science of Self-Organized Criticality*, Copernicus, NY & Oxford Univ. Press, Oxford, 1997.
- M. Mitchell, *Introduzione agli algoritmi genetici*, Apogeo, Milano, 1998.

Paragrafo 6

- N. Wiener, *Cybernetics: or Control and Communication in the Animal and the Machine*, MIT Press, 1961.
- L. von Bertalanffy, *General System Theory*, Braziller, NY, 1968.
- R. Ashby, *Introduzione alla cibernetica*, Einaudi, Torino, 1970.
- H. von Foerster, *Sistemi che osservano*, Astrolabio, Roma, 1987.
- H. von Foerster, *Attraverso gli occhi dell'altro*, Guerini, Milano, 1996.
- H. Pattee, *Hierarchical Control in Living Matter*, in *Math.Biol.*1, 1973.

- R. Rosen, *Anticipatory Systems*, Pergamon Press, 1986.
- F. Heylighen, *Modelling Emergence*, in *World Futures*, 31(89), 1991.
- N.A. Baas & C. Emmeche, *On Emergence and Explanation*, in *SFI Working Paper*, 97-02-008, Santa Fe Inst., 1997.
- P. Cariani, *Adaptivity and Emergence in Organism and Devices*, in *World Futures*, 32(111)1991.
- M. Barbieri, *La teoria semantica dell'evoluzione*, Boringhieri, Torino, 1985.
- M. Barbieri, *I codici organici. La nascita della biologia semantica*, PeQuod, Ancona, 2000.
- G. Minati, M.P. Penna, E. Pessa, *Thermodynamical and Logical Openness in General Systems*, in *Syst. Res.* 15(131)1998.

Paragrafo 7

- P.W. Anderson & D.L. Stein, *Broken Symmetry, Emergent Properties, Dissipative Structures, Life. Are They Related?*, in F.E. Yates(ed), *Self-Organizing Systems: The Emergence of Order*, Plenum Press, NY, 1985.
- I. Licata, *Osservando la sfinge. La realtà virtuale della fisica quantistica*, Di Renzo, Roma, 2003.
- A.S. Davydov, *Meccanica quantistica*, MIR, Mosca, 1981.
- A.S. Davydov, *Biology and quantum mechanics*, Pergamon Press, Oxford, 1982.
- H. Froehlich, *Theoretical Physics and Biology*. In H. Froehlich (Ed.), *Biological Coherence and Response to External Stimuli*, Springer, 1988.
- L.M. Ricciardi & H. Umezawa, *Brain and Physics of Many-Body Problems*, in *Kybernetik*, 4(44), 1967.
- K.H. Pribram, *Brain and Perception: Holonomy and Structure in Figural Processing*. Lawrence Erlbaum, NJ, 1991.
- G. Vitiello, *Dissipation and memory capacity in the quantum brain model*, in *Int. J. Mod. Phys.* 9(973), 1995.
- G. Vitiello, *My Double Unveiled*, John Benjamin Publ., 2001.
- E. Pessa & G. Vitiello, *Quantum Dissipation and Neural Net Dynamics*, in *Bioelectrochemistry and Bioenergetics*, 48(339), 1999.
- E. Del Giudice, S. Doglia, M. Milani G. Vitiello, *A Quantum Field Theoretical Approach to the Collective Behavior of Biological Systems*, in *Nucl. Phys.* B251(375), 1985.
- E. Del Giudice, S. Doglia, M. Milani, G. Vitiello, *Electromagnetic Field and Spontaneous Symmetry Breakdown in Biological Matter*. In *Nucl. Phys.* B275(185), 1986.

M. Jibu & K. Yasue, *What is Mind? Quantum Field Theory of Evanescent Photons in Brain as Quantum Theory of Consciousness*, in *Informatica*, 21(471), 1997.

F. Varela, J.P. Lacheaux, E. Rodriguez, J. Martinerie, *The BrainWeb: Phase Synchronization and Large-Scale Integration*, in *Nature Reviews/Neuroscience* 2(229), 2001.

W. Marshall, C. Simon, R. Penrose, D. Bouwmeester, *Towards Quantum Superpositions of a Mirror*, in *Physical Review Letters*, 91, 2003.

M. Jibu, S. Hagan, S. R. Hameroff, K.H. Pribram, K. Yasue, *Quantum Optical Coherence in Cytoskeletal Microtubules: Implication for Brain Function*, in *Biosystems* 32(195), 1994.

Hoi-Kwong Lo, S. Popescu, T. Spiller, *Introduction to Quantum Computation and Information*, World Scientific, Singapore, 2002.

LOGOS, IL CREATORE DI IMPERI*

Cesare Marchetti

IIASA - International Institute for Applied Systems Analysis

Schlossplatz 1, 2361 Laxenburg, Austria

e-mail marchetti@iiasa.ac.at

Sommario

Partendo dall'ipotesi che il DNA e il linguaggio rappresentino due stadi dello stesso cammino evolutivo per manipolare l'informazione, esamineremo le analogie quantitative del loro comportamento. Come ho dimostrato con numerosissimi esempi le equazioni di Volterra, così efficienti per descrivere la dinamica di sistemi biologici in competizione, sono di fatto altrettanto efficienti nel descrivere il comportamento sociale. L'emergenza del linguaggio sintattico prima e poi della scienza, interpretata come un metalinguaggio, sono attribuiti ad una "iperciclizzazione" di strutture di base in competizione, in analogia con la iperciclizzazione di quasi-specie di molecole replicanti nella teoria di Manfred Eigen sullo sviluppo e l'evoluzione del DNA. Questa continuità permette di trasferire a valle gli schemi dei processi evolutivi, mostrando ad esempio che l'ingegneria genetica fa parte naturale di un processo più generale di presa di potere del livello gerarchico più elevato, il logos, su quello che lo ha preceduto, in questo caso il DNA, e che si potrebbe pensare ad una possibile iperciclizzazione delle culture che permetterebbe di preservarle interattivamente.

Guardando agli effetti che la scienza e la tecnologia hanno avuto sull'evoluzione della nostra società negli ultimi trecento anni c'è di che meravigliarsi e spaventarsi. Non ci stiamo dirigendo con velocità

* Rielaborato da: "On the role of science in the postindustrial society"