
Neural Systems and Artificial Life Group,
Institute of Psychology,
National Research Council, Rome

Connessionismo evolutivo e origine della modularità

Raffaele Calabretta

Gennaio 2002

In: A. Borghi e T. Iachini (2002), *Scienze della Mente*, pp. 47-63, Il Mulino, Bologna.

Reparto Sistemi Neurali e Vita Artificiale,
Istituto di Psicologia, C.N.R., Viale Marx 15 - 00137 - Rome, Italy
voice: +39 6 86090233 fax: +39 6 824737
e-mail: rcalabretta@ip.rm.cnr.it
<http://gral.ip.rm.cnr.it/rcalabretta>

Connessionismo evolutivo e origine della modularità

Raffaele Calabretta
Istituto di Psicologia, Consiglio Nazionale delle Ricerche, Roma
rcalabretta@ip.rm.cnr.it
<http://gral.ip.rm.cnr.it/rcalabretta>

Introduzione

A chi apre un libro di anatomia o di fisiologia comparata appare subito evidente che lo studio della struttura e delle funzioni di molti organismi, soprattutto degli organismi più complessi, si basa implicitamente su un principio di organizzazione modulare.

In altre parole, si dà per acquisito il fatto che gli organismi complessi sono modulari, cioè composti da parti con forma e/o funzioni diverse. Il corpo umano, ad esempio, è formato da organi che sono specializzati a svolgere funzioni specifiche: il fegato ha la funzione di metabolizzare le sostanze che ingeriamo, il cuore è specializzato a pompare il sangue e a distribuirlo in tutto il corpo, ecc. A ben guardare, una organizzazione modulare si ritrova ai diversi livelli di organizzazione degli organismi (genetico, molecolare, neurale, comportamentale/cognitivo).

Dietro l'apparente ovvietà della modularità degli organismi complessi si nascondono però numerose questioni che non sono state ancora del tutto risolte, alcune delle quali, anzi, non sono state mai formulate esplicitamente.

Per elencarne solo alcune: quali sono i vantaggi di una organizzazione modulare rispetto a una organizzazione non modulare? E quali sono i meccanismi che favoriscono l'emergere evolutivo della modularità e quali sono i meccanismi che al contrario ne ritardano o addirittura ne impediscono l'insorgenza? E questi meccanismi: sono uguali per tutti i tipi di moduli oppure c'è un meccanismo che spiega per esempio l'origine della modularità del cervello e un altro meccanismo che spiega l'origine della modularità del resto del corpo?

Le domande possibili sono troppe per poterle elencare tutte, e sono proprio quelle che si cominciano a porre le discipline che sono interessate allo studio degli organismi, della loro organizzazione strutturale e funzionale, e del loro comportamento.

Nel loro rispettivo ambito e utilizzando un diverso apparato teorico e verbale, la biologia, le neuroscienze, le scienze cognitive - per elencare solo alcune discipline -, cercano di dare una risposta a tutte queste domande. Ma ci riescono con molta difficoltà!

Uno dei motivi che spiegano questa difficoltà, forse il più decisivo, è la mancanza di quell'approccio interdisciplinare che lo studio della modularità invece richiederebbe. La realtà è che

esiste una separazione non solo tra discipline relativamente lontane tra loro quali sono la biologia dello sviluppo e le scienze cognitive, ma la separazione c'è anche tra discipline vicine, quali la biologia evoluzionistica e la biologia dello sviluppo.

Prendiamo il caso della biologia: mentre la biologia evoluzionistica si interroga sulle origini evolutive (filogenetiche) della modularità degli organismi, si interroga cioè sui cosiddetti moduli evolutivi, la biologia dello sviluppo si interroga su che tipo di programmi di sviluppo sono codificati nel codice genetico, si interroga quindi sui cosiddetti moduli di sviluppo. Ma, volendo cogliere l'occasione per aggiungere un'altra domanda alla serie precedentemente interrotta, ci si potrebbe chiedere: che relazione c'è tra i due tipi di moduli? In altre parole: i moduli evolutivi e i moduli dello sviluppo sono o non sono la stessa cosa? La realtà è che questa domanda, insieme a molte altre ancora, fino a poco tempo fa non poteva neppure essere formulata nell'ambito di due discipline che non comunicavano tra loro.

Il tema della modularità è da lungo tempo anche al centro del dibattito nelle scienze cognitive. Da una parte i cognitivisti ritengono che la mente sia composta da una collezione di moduli specializzati nelle diverse abilità cognitive e nelle diverse aree di attività. Questi moduli sarebbero innati, cioè codificati nel genoma e presenti quindi alla nascita, e rimarrebbe solo da capire quali e quanti sono questi moduli e come sono fatti [Fodor 1983, trad. it. 1988]. Dalla parte opposta, gli empiristi sottolineano il ruolo del processo di sviluppo nel modellare la mente e nel fare ciò si arroccano in una posizione anti-innatista della mente [Karmiloff-Smith 2000]. In questo ultimo ambito il connessionismo [Rumelhart e McClelland 1986] utilizza le reti neurali come modello del cervello per dimostrare le elevate capacità di apprendimento di reti neurali non-modulari che partono da una condizione di totale mancanza di competenza nel compito testato.

Fortunatamente la situazione sta cambiando. Negli ultimi tempi si è andato infatti sviluppando sia in biologia che nelle scienze cognitive un approccio interdisciplinare allo studio dell'organizzazione modulare degli organismi.

In biologia ciò è potuto avvenire grazie all'incremento notevole di conoscenze verificatosi negli ultimi anni riguardo alle basi molecolari dello sviluppo, che permette una maggiore comprensione dell'organizzazione modulare dei sistemi viventi e che quindi favorisce la possibilità di sintesi tra biologia evoluzionistica e biologia dello sviluppo.

Tesi principale di questo capitolo è che un approccio interdisciplinare allo studio della modularità sta diventando possibile anche nell'ambito delle scienze cognitive grazie all'utilizzo dei nuovi strumenti simulativi di Vita Artificiale [Calabretta *et al.* 1998a], che utilizzano gli algoritmi genetici come modelli del processo evolutivo e le reti neurali come modelli del cervello e che permettono quindi di simulare l'evoluzione filogenetica della modularità della mente.

Contrariamente ai modelli cognitivisti e ai modelli connessionisti classici, l'utilizzo dei modelli simulativi di Vita Artificiale consente uno studio integrato dei diversi livelli organizzativi degli organismi (genetico, neurale e comportamentale) e quindi permette di capire meglio che tipo di correlazione esiste tra i moduli presenti ai diversi livelli, per esempio tra un modulo genetico e un modulo neurale. Questa caratteristica li rende molto adatti allo studio di argomenti spiccatamente interdisciplinari qual è lo studio della evoluzione della modularità. Non è quindi un caso se questo stesso strumento si stia rivelando efficace per verificare la validità delle teorie postulate sia in ambito biologico che in ambito cognitivo [Calabretta *et al.* in press].

Nella sezione 2 darò una visione panoramica dei termini dello scontro tra cognitivisti ed empiristi riguardo a modularità e innatismo della mente e comparerò i modelli a scatole e frecce dei cognitivisti con i modelli connessionisti degli empiristi. La descrizione di questo scenario mi porterà a presentare un nuovo approccio allo studio della modularità del corpo e della mente degli organismi, il connessionismo evolutivo [Calabretta e Parisi in stampa], che si basa sull'uso delle simulazioni di Vita Artificiale in cui gli algoritmi genetici, modelli dell'evoluzione, fanno evolvere l'architettura e/o i pesi delle reti neurali, modelli del sistema nervoso [Di Ferdinando, Calabretta e Parisi 2001]. Nella sezione 3 presenterò una simulazione dell'evoluzione della modularità basata sull'approccio del connessionismo evolutivo [Calabretta *et al.* 2000], che è già servito a formulare nuove ipotesi sull'origine della modularità [Wagner, Mezey e Calabretta in stampa], e che ha quindi dimostrato le notevoli potenzialità di questo approccio.

1. Le tre opzioni per lo studio della modularità della mente

Nelle scienze cognitive, il campo interdisciplinare di ricerca che studia la mente, il tema della modularità è molto controverso. Estremizzando un poco si potrebbe dire che le scienze cognitive sono due: il cognitivismo e il connessionismo. Il cognitivismo rappresenta il paradigma teorico di più lunga data, nato negli Stati Uniti alla fine degli anni cinquanta. Esso è basato sulla analogia tra mente e computer software e considera la mente come una manipolazione di simboli [Newell e Simon 1976]. Negli ultimi 20 anni è emersa una scienza cognitiva di tipo diverso, il connessionismo [Rumelhart e McClelland 1986], che rifiuta l'analogia mente/computer e interpreta il comportamento e le abilità cognitive utilizzando modelli teorici che sono direttamente ispirati alla struttura fisica e al modo di funzionare del sistema nervoso. Questi modelli sono le reti neurali. Alla stregua dei neuroni cerebrali, le reti neurali artificiali sono formate da unità che sono collegate tra loro attraverso connessioni che somigliano alle sinapsi cerebrali.

Per il connessionismo la mente non è manipolazione di simboli, ma è il risultato globale di moltissime interazioni che avvengono nella rete di neuroni che formano il sistema nervoso e

consiste esclusivamente in processi quantitativi in cui cause fisico-chimiche producono effetti fisico-chimici.

Il cognitivismo è fortemente modularista. La mente sarebbe infatti formata da moduli distinti che sono specializzati a elaborare tipi di informazione diversa [Chomsky 1980; Fodor 1983, trad. it. 1988]. I modelli utilizzati dai cognitivisti sono i cosiddetti diagrammi a scatole e frecce. In questo tipo di modelli ogni scatola rappresenta un modulo con una specifica funzione e le frecce che collegano le scatole stanno a indicare che l'informazione processata da un particolare modulo viene poi trasferita a un altro modulo per essere processata ulteriormente.

Al contrario, il connessionismo tende a essere anti-modularista. Nelle reti neurali infatti l'informazione è rappresentata da pattern di attivazione distribuiti in grandi gruppi di neuroni e il modo di funzionare delle reti neurali consiste nella trasformazione di pattern di attivazioni in altri pattern di attivazione che si realizza lungo le connessioni che collegano i diversi neuroni artificiali. Come conseguenza le reti neurali utilizzate come modelli del cervello sono in gran parte non-modulari se si eccettua la differenziazione in unità di input, unità nascoste e unità di output [si veda Di Ferdinando in questo volume].

Il punto di contrasto più importante tra le due diverse scienze cognitive verte su innatismo e anti-innatismo. I cognitivisti sono innatisti. Secondo il loro modo di vedere le informazioni sui moduli sono specificate nel genoma degli individui.

Gli psicologi evoluzionisti, che sono cognitivisti, sostengono che la struttura modulare della mente sia il risultato di pressioni selettive e sono convinti della possibilità di risalire alla specifica pressione selettiva che spiega l'esistenza dei diversi moduli. Gli psicologi evoluzionisti [Cosmides e Tooby 1994] quindi sostengono una forma di adattivismo molto spiccata. Infatti ritengono non solo che i moduli siano codificati nei geni ma anche che questa codifica sia il risultato di processi di adattamento. Questa forma di pan-adattivismo non è però condivisa da tutti i cognitivisti. Noam Chomsky, ad esempio, considera la mente come una manipolazione di simboli e ritiene che esista uno specifico modulo mentale per il linguaggio o per la sintassi, ma non crede che il linguaggio sia emerso negli umani in seguito a specifiche pressioni evolutive [Fodor 2000].

Anche i biologi evoluzionisti come Stephen J. Gould [1997] hanno ripetutamente sostenuto che quello che viene ereditato geneticamente non è necessariamente il risultato di pressioni selettive e non è necessariamente adattativo. Può essere infatti il risultato di fattori casuali, o può essere neutro dal punto di vista adattativo e accompagnare un altro tratto adattativo, o può essere un "esattamento", un tratto cioè che si è evoluto perché svolgeva una funzione ed è stato poi riutilizzato per una funzione nuova [Gould e Vrba 1982].

Più recentemente, il contrasto tra Steven Pinker e Jerry Fodor, che sono entrambi cognitivisti e innatisti, ha dimostrato come la natura adattativa dei tratti ereditari può dividere gli stessi cognitivisti con Pinker [1999] che è a favore di un forte forma di modularismo innato, mentre Fodor è per una forte forma di modularismo non innato.

Contrariamente ai cognitivisti, i connessionisti sono anti-innatisti. Il connessionismo è infatti solitamente associato con le posizioni empiriste che considerano la mente come frutto soprattutto dell'apprendimento e dell'esperienza durante la vita.

La materia si complica se si considera anche lo sviluppo. Lo sviluppo è il processo di formazione dell'organismo adulto (fenotipo) a partire dall'informazione genetica (genotipo) contenuta nell'uovo fecondato. Il processo di sviluppo di un organismo non è un processo istantaneo ma un processo temporale e infatti consiste di una successione temporale di forme fenotipiche. Poiché ammettere che lo sviluppo è un processo temporale porta indirettamente ad ammettere l'importanza dell'apprendimento e dell'esperienza sul fenotipo, i cognitivisti non solo assumono una posizione innatista ma sminuiscono anche il ruolo dello sviluppo. Gli psicologi dello sviluppo di stampo cognitivista [Spelke *et al.* 1992; Wynn 1992] sostengono infatti che i moduli sono lì nel fenotipo sin dai primi stadi dello sviluppo e che quello che avviene durante la vita non cambia di molto la situazione. Inoltre come innatisti pensano che, anche se qualcosa cambia durante lo sviluppo, questo è dovuto non all'apprendimento e all'esperienza bensì a programmi temporali di sviluppo codificati nel codice genetico. Un esempio di questo modo di vedere è la maturità sessuale che non è presente alla nascita ma che è programmata geneticamente a emergere successivamente a un certo punto durante la vita.

Gli psicologi dello sviluppo che sono, al contrario, vicini alle posizioni del connessionismo [Karmiloff-Smith 2000], ritengono che i moduli non sono presenti nel fenotipo alla nascita bensì si sviluppano durante la vita e inoltre che i moduli sono solo in piccola parte codificati nel genoma mentre sono il risultato di complesse interazioni tra informazione genetica, sviluppo ed esperienza.

In questo capitolo sostengo una forma di connessionismo che non è né anti-modularista né anti-innatista. Il connessionismo non è necessariamente anti-innatista. Anche se molti modelli a reti neurali usano algoritmi di apprendimento per trovare i pesi delle connessioni appropriati per svolgere un compito particolare, il connessionismo non è incompatibile con il riconoscimento che alcuni aspetti della rete neurale non sono il risultato dell'apprendimento ma sono ereditati geneticamente.

Ad esempio, poiché la maggior parte delle simulazioni connessioniste parte da una architettura neurale fissa, si potrebbe ipotizzare che l'architettura delle reti neurali è determinata geneticamente mentre il ruolo dell'apprendimento è limitato a trovare i pesi appropriati per quella

architettura. (L'architettura di una rete neurale definisce come sono collegate tra loro le varie unità che la compongono.)

Infatti, Elman *et al.* [1996] sostengono che le reti connessioniste permettono al ricercatore di andare oltre al cognitivismo - che si limita a dire semplicemente che questo o quello è innato -, e di esplorare in modo dettagliato che cosa può essere innato e che cosa può essere appreso mostrando come le caratteristiche fenotipiche possono essere il risultato dell'interazione tra quello che è innato e quello che è appreso. Questi autori distinguono tra diverse cose che potrebbero essere innate in una rete neurale: i pesi delle connessioni, i vincoli di architettura e i vincoli cronotopici (che determinano quando devono succedere le cose durante lo sviluppo).

Tuttavia, per ipotizzare che qualcosa è innato nella rete neurale, non è sufficiente che alcune delle proprietà della rete neurale siano implementate a priori nella rete neurale dal ricercatore ma è necessario simulare realmente il processo evolutivo da cui originano queste proprietà o questi vincoli ereditati geneticamente. Le simulazioni di Vita Artificiale differiscono dalle simulazioni connessioniste classiche in quanto la Vita Artificiale utilizza gli algoritmi genetici [Holland 1992; Di Ferdinando in questo volume] per simulare i processi evolutivi e per evolvere le proprietà delle reti neurali ereditate geneticamente.

Diversamente dalle classiche simulazioni connessioniste, le simulazioni di Vita Artificiale simulano non una singola rete neurale che apprende alcune particolari abilità in base alla propria esperienza individuale, ma simulano una intera popolazione di reti neurali che passa attraverso una successione di generazioni di individui, ognuno dei quali nasce con un genotipo ereditato dai genitori. L'uso dell'algoritmo genetico mostra come l'informazione codificata nei genotipi ereditati cambia lungo le generazioni che si susseguono, in quanto la riproduzione è selettiva e nuove varianti dei genotipi vengono aggiunti costantemente al pool genetico della popolazione per mezzo delle mutazioni genetiche e della ricombinazione sessuata. Alla fine della simulazione si vede che i genotipi ereditati codificano le proprietà desiderate della rete neurale che rappresentano vincoli innati allo sviluppo e al comportamento.

Questo tipo di connessionismo è il *connessionismo evolutivo*. Nella tabella 2.1 vengono presentate in sintesi le tre opzioni che sono attualmente disponibili per lo studio del comportamento e della mente degli organismi.

Il connessionismo evolutivo non solo permette di studiare come l'informazione ereditata geneticamente può emergere spontaneamente in popolazioni di reti neurali, invece che essere implementata arbitrariamente nella rete neurale dal ricercatore, ma anche di esplorare tutte le possibili interazioni tra l'evoluzione a livello di popolazione e l'apprendimento a livello del singolo individuo che determina il fenotipo reale.

2. Le reti neurali modulari come modelli allo stesso tempo del cervello e della mente

È importante tenere bene in mente che la nozione di modulo non è la stessa per i cognitivisti e per i connessionisti. Il modularismo cognitivista è diverso dal modularismo neurale.

Per i cognitivisti infatti i moduli sono componenti di teorie in base ai quali i fenomeni empirici sono interpretati e spiegati. La teoria o il modello di un particolare fenomeno ipotizza l'esistenza di moduli separati con struttura e/o funzione diverse il cui lavoro congiunto spiega il fenomeno di interesse. Perciò i moduli cognitivisti sono entità postulate più che osservate. Per esempio, in psicolinguistica il comportamento linguistico osservato negli adulti e nei bambini è interpretato come richiedente due moduli distinti, uno che supporta la capacità di produrre il passato remoto dei verbi regolari inglesi (ad esempio, *worked*) e l'altro quella di produrre il passato dei verbi irregolari (ad esempio, *brought*) [Pinker e Prince 1988; Cangelosi e Turner in questo volume]. Questa nozione di modulo puramente teorica è stata esplicitamente difesa e precisamente definita nel famoso libro di Fodor [1983] *The Modularity of mind*.

<u>COGNITIVISMO</u>	Mente come manipolazione di simboli che avviene in un sistema analogo al computer	INNATISTA	MODULARISTA
<u>CONNESSIONISMO</u>	Mente come risultato globale delle moltissime interazioni fisico- chimiche che avvengono in una rete di neuroni	ANTI- INNATISTA	ANTI- MODULARISTA
<u>CONNESSIONISMO EVOLUTIVO</u>	Mente come risultato globale delle moltissime interazioni fisico- chimiche che avvengono in una rete di neuroni	INTERAZIONE TRA EVOLUZIONE E APPRENDIMENTO	MODULARISTA

Tabella 2.1. Tre opzioni per lo studio del comportamento e della mente.

Lo stesso vale per la psicologia evoluzionista che, come abbiamo già sottolineato, ha un orientamento di tipo cognitivista. La concezione della mente come coltellino svizzero, cioè come un insieme di moduli adattivi specializzati ed ereditati geneticamente, si basa su una nozione di modulo secondo la quale i moduli sono entità teoriche la cui esistenza è suggerita dall'osservazione del comportamento umano.

Anche i neuroscienziati hanno una concezione modulare della mente [Restak 1995]. Tuttavia la concezione della mente che hanno i neuroscienziati si basa sull'osservazione empirica dell'anatomia e della fisiologia del cervello piuttosto che su una teoria. Il cervello è ovviamente diviso in una varietà di "moduli": aree corticali; strutture sottocorticali; sistemi interconnessi come la corteccia visiva retino-genicolata per la visione o il sistema gangli della base-corteccia frontale per l'attenzione. Questa ricca modularità del cervello, sia strutturale (anatomica e citoarchitettónica) sia funzionale (fisiologica), è evidenziata tramite l'osservazione strumentale diretta, tramite i dati sulla localizzazione delle lesioni nelle varie patologie comportamentali/mentali e sulle dissociazioni neuropsicologiche, e più recentemente e sempre di più, tramite i dati delle neuroimmagini.

Si può andare a cercare se vi è corrispondenza tra i due tipi di moduli, i moduli teorici del cognitivismo e i moduli del cervello osservati sperimentalmente. E in effetti questo è quanto si ritiene che i neuropsicologi debbano fare. Essi interpretano i deficit comportamentali dei pazienti usando i modelli teorici "scatole e frecce" della psicologia cognitiva – nei quali le scatole sono i moduli e le frecce indicano le relazioni tra moduli – e poi cercano di far corrispondere questa analisi modulare con le osservazioni e le misure sulla localizzazione delle lesioni e con gli altri dati del cervello dei pazienti. Tuttavia, non si può assumere che i modelli modulari teorici del cognitivismo corrispondano necessariamente alla struttura modulare osservata e al funzionamento del cervello.

I moduli cognitivi potrebbero non corrispondere alla struttura fisica neurale o ai moduli funzionali del cervello, e il cervello potrebbe essere organizzato in moduli che non corrispondono ai componenti dei modelli teorici in base ai quali psicologi e ricercatori cognitivi interpretano e spiegano i dati comportamentali.

È importante tenere in mente ciò quando ci si volge verso l'altro tipo di modelli teorici che possono essere usati per interpretare e spiegare i dati comportamentali e le capacità cognitive, cioè verso le reti neurali. Queste sono modelli teorici che, diversamente dai modelli teorici del cognitivismo, si ispirano direttamente alla struttura fisica e al funzionamento del cervello.

Quindi, *le reti neurali sono allo stesso tempo modelli del cervello e della mente*. Le reti neurali finora usate nella maggior parte delle simulazioni sono di tipo non-modulare. Sono reti omogenee di unità dotate di struttura molto semplice e costituite da un set di unità o modulo di input, un modulo di output e in mezzo, quasi sempre, un singolo modulo interno di unità nascoste.

Tuttavia, questo dovrebbe essere considerato un limite degli odierni modelli di reti neurali e non una proprietà intrinseca delle reti neurali. Se si vuole che le reti neurali siano ispirate alla struttura e al modo di funzionare del cervello, allora esse dovranno essere modulari perché il cervello è modulare. È importante notare, tuttavia, che i moduli delle reti neurali saranno più simili ai moduli del cervello che “le scatole” teoriche dei modelli “scatole e frecce” del cognitivismo.

Un modulo in una rete neurale è un modulo fisico (simulato), e non un costrutto teorico postulato. Un modulo neurale potrà essere costituito da alcune unità della rete che hanno un numero maggiore di connessioni interne che collegano le unità del modulo, rispetto al numero di connessioni esterne che collegano le unità del modulo con le unità esterne al modulo. Oppure, da un punto di vista più funzionale, un modulo neurale potrà essere l'attività correlata osservata di un dato sottogruppo di unità della rete neurale, senza che vi sia separazione anatomica di questo sottogruppo di unità.

Se la struttura modulare della rete neurale è implementata dal ricercatore, il ricercatore dovrà trarre ispirazione dalla reale struttura del cervello invece che da considerazioni teoriche basate sui modelli cognitivi. Se, più nello spirito del connessionismo, l'architettura neurale non sarà implementata dal ricercatore ma sarà il risultato di un processo di evoluzione, sviluppo, e/o apprendimento, il ricercatore sarà maggiormente interessato a capire se la struttura modulare che emerge corrisponde alla reale modularità del cervello.

Il contrasto vero tra modelli a reti neurali e modelli cognitivi non riguarda però la modularità in sé, ma piuttosto la natura dei moduli e la domanda su quali modelli teorici sono più appropriati per spiegare il comportamento e la cognizione.

Si consideri di nuovo, per esempio, l'ipotesi che gli inglesi di madre lingua producono il passato remoto usando due moduli distinti, uno per i verbi regolari e l'altro per i verbi irregolari. Alcune evidenze sperimentali suggeriscono che questi due moduli sono localizzati in zone diverse del cervello. Pazienti con lesioni nella parte anteriore del cervello in genere non riescono a formare le forme del passato regolari, mentre la loro abilità a produrre le forme del passato irregolari sembra essere preservata. Al contrario, pazienti con lesioni nella parte posteriore del cervello tendono a mostrare un pattern opposto. Gli riesce difficile produrre il passato di verbi irregolari mentre sono in grado di formare quelli regolari. Un fenomeno come questo potrebbe indicare che ci sono veramente due moduli neurali distinti coinvolti nella produzione del passato dei verbi. Questa spiegazione può essere completamente accettabile per un connessionista, il quale andrà poi a simulare la produzione del passato dei verbi mediante una rete neurale modulare con due moduli distinti, uno per i verbi regolari e l'altro per i verbi irregolari. (Questi due moduli potranno essere o strutturali o funzionali, nell'accezione dei due termini descritta sopra).

Quello che distingue l'approccio cognitivo e l'approccio neurale al trattamento del passato dei verbi è la natura dei moduli. I cognitivisti sostengono che il modulo del passato dei verbi regolari sarebbe un modulo che funziona a regole. Quando produce il passato del verbo *to work*, il cervello starebbe applicando la regola: "Aggiungi il suffisso *-ed* alla radice del verbo". Al contrario, il modulo del passato dei verbi irregolari sarebbe un modulo che funziona ad associazioni che contiene un elenco definito di radici verbali ognuno associato con la sua forma di passato di verbo irregolare. Quello che farebbe il cervello è consultare questo elenco di associazioni, trovare la radice verbale appropriata (per esempio, *bring*), e produrre la corrispondente forma del passato (*brought*).

Questa interpretazione teorica del comportamento sul passato dei verbi è rifiutata dai connessionisti perché gli strumenti teorici da loro utilizzati, cioè le reti neurali, non permettono una interpretazione del genere. I modelli basati su reti neurali sono ispirati al cervello e il cervello è un sistema fisico fatto di entità e processi fisici in cui tutto quello che può avvenire è la produzione di fenomeni fisico-chimici ad opera di cause fisico-chimiche. Per principio, quindi, una rete neurale non può servirsi di una regola come spiegazione di alcun tipo di comportamento o di abilità cognitiva.

Un connessionista può accettare che zone separate del cervello, e di una rete neurale che simula il cervello, possano essere responsabili della produzione delle forme regolari e irregolari del passato dei verbi. Tuttavia, questi moduli non potranno che funzionare nello stesso modo basilare: le unità sono attivate dalle eccitazioni e inibizioni che arrivano dalle altre unità connesse.

Questo non esclude la possibilità che si possano scoprire delle differenze nella organizzazione e nel modo di funzionare dei due diversi moduli neurali coinvolti nei verbi inglesi regolari e irregolari e ciò naturalmente richiede una spiegazione del perché il cervello ha trovato utile avere due moduli separati per il controllo del passato dei verbi invece di uno solo. Ciò pone il problema della origine dei moduli che tratteremo nella sezione seguente.

3. Meccanismi evolutivi all'origine della modularità biologica: il ruolo della duplicazione genetica

Come è stato già sottolineato nell'introduzione, uno degli obiettivi della biologia evuzionistica è capire come la selezione naturale può aver agito sul fenotipo per produrre moduli evolutivi. Una definizione preliminare di modulo evolutivo è: un insieme di caratteristiche fenotipiche che sono altamente integrate dagli effetti pleiotropici di alcuni geni sottostanti e sono relativamente isolate dagli altri insiemi di geni dalla scarsità di effetti pleiotropici [Wagner e

Altenberg 1996]. (La pleiotropia è l'influenza degli stessi geni su caratteri diversi [Futuyma 1998, p. 429].)

Un modello evolutivo che voglia spiegare l'origine dei moduli deve spiegare come la selezione naturale può produrre questa distribuzione di effetti genetici. Negli ultimi due anni sono stati proposti ben sette diversi modelli teorici per spiegare l'origine della modularità [cfr. Wagner, Mezey e Calabretta in stampa]. In tutti questi modelli tranne uno, la modularità è direttamente o indirettamente connessa con qualche vantaggio selettivo. In uno studio sull'evoluzione di modularità funzionale che faceva uso di modelli di Vita Artificiale, Calabretta *et al.* [1998a; 1998b; 2000] hanno invece scoperto un meccanismo che non può essere classificato come selezione diretta o indiretta della modularità *per se*. Questa modularità funzionale emerge dalla specializzazione di moduli strutturali duplicati senza alcun intrinseco beneficio in termini di performance o tasso evolutivo. La modularità, cioè, emerge completamente come effetto collaterale delle dinamiche evolutive.

Calabretta *et al.* hanno utilizzato un modello di Vita Artificiale in cui un algoritmo genetico permette l'evoluzione sia dell'architettura e sia dei pesi di una popolazione di reti neurali che controllano il comportamento di un robot mobile. Ogni robot vive in una arena recintata con muri e ha il compito di esplorare questo ambiente, di trovare gli oggetti che vi sono sparsi, di sollevarli con la pinza di cui è munito e di trasportarli fuori dall'arena. Il robot è munito di sensori all'infrarosso che lo informano della presenza di oggetti e muri nelle vicinanze, di due ruote che gli permettono di muoversi nell'ambiente, e di una pinza utile per sollevare un oggetto alla volta e trasportarlo fuori dall'arena. Il compito del robot consiste nell'esplorare l'arena alla ricerca di oggetti, di sollevare con la pinza l'oggetto trovato e di trasportarlo vicino a un muro, sul quale dovrà infine rilasciarlo. Per fare ciò il robot ha quattro motori che controllano rispettivamente le due ruote, la procedura di chiusura e di sollevamento della pinza, e la procedura di abbassamento e di chiusura della pinza.

Questo compito è molto difficile da imparare perché la rete neurale deve imparare a eseguire la corretta sequenza di azioni: esplorare l'ambiente, trovare l'oggetto distinguendolo dai muri, agganciare e sollevare l'oggetto con la pinza, trovare un muro, disporsi nella corretta posizione rispetto a questo ultimo e infine abbassare la pinza e rilasciare l'oggetto sul muro. Ci sono quindi un certo numero di compiti comportamentali che richiedono a livello neurale un controllo differenziato dell'output motorio. Essenzialmente ci sono da eseguire correttamente due tipi di comportamenti principali: cercare un nuovo oggetto nell'arena e rimuoverlo dall'arena. L'assenza di un oggetto nella pinza richiede il comportamento di ricerca e presa dell'oggetto, mentre la presenza di un oggetto nella pinza richiede il comportamento opposto di ricerca del muro e di rilascio dell'oggetto

su di esso. La domanda è se queste due sequenze comportamentali debbano essere rappresentate da substrati neurali differenti, cioè se vi debba essere o meno modularità funzionale.

In uno studio condotto da Nolfi [1997] si è dimostrato che la modularità funzionale non era necessaria per risolvere questo compito. All'inizio della simulazione, Nolfi aveva provvisto la rete neurale del robot con due elementi neurali per ciascun output motorio che competevano tra loro per il controllo del corrispondente output motorio. Prendiamo ad esempio la ruota di sinistra: fin dall'inizio della simulazione vi erano due unità neurali che per ogni movimento del robot competevano per il controllo della ruota; la decisione su quale delle due unità di controllo assumeva effettivamente il controllo per ogni ciclo temporale dipendeva dalla comparazione delle attività di due selettori, uno per ciascuna unità di controllo. Nolfi ha mostrato che l'algoritmo genetico risolveva il problema senza che i comportamenti fossero rappresentati dai diversi elementi neuronali. In pratica, le unità di controllo non si specializzavano nel controllo di determinati comportamenti e quindi non evolveva alcuna modularità funzionale. Questo risultato mostra che la modularità funzionale non è necessaria per risolvere un complesso problema adattativo che consiste in un certo numero di compiti differenti l'uno dall'altro.

Calabretta *et al.* hanno modificato il set simulativo. All'inizio della simulazione il robot ha una rete neurale in cui per ciascun output motorio c'è un solo elemento neurale. Durante il processo evolutivo, però, oltre alla mutazione dei valori dei pesi è permessa anche la duplicazione di questi elementi neurali. Mediante l'analisi del comportamento dei robot è stato possibile dimostrare che la rete neurale modulare per duplicazione genetica era caratterizzata da un alto grado di specializzazione. Infatti, alcuni moduli neurali erano specializzati per il controllo di alcuni compiti (cioè per il controllo dei movimenti del robot quando il robot è alla ricerca di oggetti), mentre altri moduli neurali erano specializzati per il controllo degli altri compiti (cioè per il controllo del robot quando sta trasportando un oggetto e va alla ricerca dei muri). È importante sottolineare che le popolazioni di robot che hanno sviluppato la specializzazione, cioè la modularità funzionale, raggiungono gli stessi livelli di performance delle popolazioni non specializzate. Inoltre le reti neurali specializzate non raggiungono la soluzione più velocemente di quelle non specializzate. In altre parole, la modularità funzionale non comporta alcun intrinseco beneficio adattativo. Ma allora quale è il meccanismo che in queste simulazioni dà origine alla modularità funzionale?

Diversi elementi suggeriscono lo scenario seguente: innanzi tutto vi è la duplicazione di una unità neuronale di controllo motorio. Nel modello questo passaggio è neutro da un punto di vista evolutivo, non dà cioè vantaggi in termini di fitness, in quanto le due unità duplicate sono identiche. Successivamente, si ha l'acquisizione di un cambiamento neutro nella regolazione delle unità duplicate che modifica il contesto ambientale in cui una delle due unità assume il controllo

dell'output motorio, per esempio durante la ricerca di oggetti piuttosto che nella ricerca dei muri. Segue, infine, l'accumulo di mutazioni che permettono l'adattamento dell'unità di controllo al contesto funzionale in cui essa viene utilizzata più frequentemente. Questo ultimo passaggio porta a un co-adattamento della parte regolatoria (i selettori) e funzionale (le unità di controllo motorio) della rete neurale che blocca il sistema in uno stato che è funzionalmente specializzato.

Da un punto di vista della genetica di popolazione, l'evoluzione della specializzazione funzionale in questo modello è causata dalle interazioni tra geni che influenzano in quale situazione un'unità di controllo è attiva e geni che controllano l'output motorio prodotto. Vi è una sorta di meccanismo irreversibile tra mutazioni che influenzano il contesto ambientale in cui l'unità di controllo motorio viene utilizzata e mutazioni che portano alla specializzazione dell'output al contesto ambientale in cui viene utilizzato più frequentemente.

Conclusioni

In questo capitolo ho presentato un nuovo approccio allo studio della modularità della mente che tiene in considerazione la storia filogenetica dei moduli cerebrali degli organismi. Questo approccio, basato sui metodi simulativi della Vita Artificiale ci permette di simulare nello stesso modello i livelli genetico, neurale e comportamentale degli organismi e ci aiuta a chiarire come i moduli di un livello sono correlati con i moduli degli altri livelli.

La modularità della mente è un argomento molto dibattuto nelle scienze cognitive. I cognitivisti concepiscono la mente come un insieme di moduli specializzati distinti e ritengono che questa ricca modularità sia essenzialmente innata, con gli psicologi evolucionisti che addirittura sostengono che ogni modulo sia adattativo e sia stato selezionato biologicamente come risultato di specifiche pressioni evolutive. (Ma altri cognitivisti come Chomsky e Fodor, credono che i moduli sono innati ma non sono necessariamente adattativi [Fodor 2000].)

Al contrario, i connessionisti ritengono che la mente è un sistema più omogeneo che eredita geneticamente solo una generica capacità a imparare dalle esperienze e che, se anche ci fossero dei moduli, questi sarebbero il risultato dei processi di sviluppo e di apprendimento invece che essere innati.

In questo capitolo ho sostenuto che il connessionismo non è necessariamente modularista e anti-innatista. Al contrario, poiché le reti neurali si ispirano al cervello, esse non possono che essere modulari (sebbene la maggior parte delle reti neurali usate nelle simulazioni connessioniste siano non-modulari) in quanto il cervello è una struttura ricca di moduli specializzati. Considerare le reti neurali in una prospettiva di Vita Artificiale, ci permette di sviluppare un connessionismo che è appropriatamente modularista e innatista, un connessionismo evolutivo. Le simulazioni di Vita

Artificiale simulano l'evoluzione nel tempo di popolazioni di organismi, i quali ereditano un genotipo dai loro genitori che insieme all'esperienza e all'apprendimento determina il fenotipo del singolo individuo. Risulta così semplice esplorare con le simulazioni che tipo di architetture, modulari o non-modulari, emergono in base ai compiti considerati e come evoluzione e apprendimento possono cooperare nel formare il singolo fenotipo.

Anche se il connessionismo può essere modularista, questo non vuol dire che ci sia per forza una corrispondenza tra moduli dei connessionisti e moduli dei cognitivisti.

I moduli dei cognitivisti sono entità teoriche che sono postulati nei modelli "scatole e frecce" utilizzati per spiegare i dati comportamentali. I moduli dei connessionisti sono parti specializzate del cervello che sono separate anatomicamente e/o funzionalmente. Ci può essere solo una parziale coincidenza tra i due termini e in ogni caso la ricerca sui moduli neurali è orientata in maniera molto diversa della ricerca sui moduli cognitivi e considera altri tipi di evidenza empirica.

Il connessionismo evolutivo condivide lo stesso obiettivo primario della psicologia evoluzionista, cioè di sviluppare una psicologia che tenga conto del fatto che l'architettura della mente umana è il prodotto di un processo evolutivo [Barkow *et al.* 1992], ma si distingue dalla psicologia evoluzionistica per tre motivi principali: (1) usa le reti neurali invece che i modelli cognitivi per interpretare il comportamento; (2) adotta le simulazioni al computer per testare gli scenari evolutivi; (3) ha una concezione meno pan-adattivistica dell'evoluzione ed è più interessato alla complessa interazione tra informazione ereditata geneticamente e informazione acquisita mediante l'esperienza.

La simulazione degli scenari evolutivi ci permette di considerare i fattori casuali e gli altri fattori evolutivi non adattativi e perciò ci evita il rischio di spiegare tutte le caratteristiche morfologiche e funzionali degli organismi in chiave esclusivamente adattiva [per un esempio si veda Cangelosi, Parisi e Nolfi 1994].

Abbiamo presentato un esempio di simulazione di Vita Artificiale, i cui risultati hanno consentito di formulare un nuovo modello teorico per l'evoluzione della modularità. Questo modello si discosta dagli altri modelli presenti in letteratura in quanto in esso la modularità emerge come effetto collaterale delle dinamiche evolutive, senza cioè alcun vantaggio in termini di fitness o di tasso evolutivo. Questi risultati fanno supporre l'importante ruolo esercitato dal meccanismo di duplicazione genetica per l'evoluzione della modularità del corpo e della mente degli organismi [Allman, 2000, p. 43].

In biologia evoluzionistica sono stati identificati finora sette possibili meccanismi che possono spiegare l'origine della modularità [Wagner, Mezey e Calabretta in stampa]. La maggior parte di essi è stata formulata negli ultimi due anni e quindi nessuno di questi meccanismi è stato

ancora studiato abbastanza per poterlo escludere come candidato. Una notevole mole di ricerca è richiesta solo per classificare queste diverse possibilità e forse per trovarne altre nuove.

È importante sottolineare che una possibilità reale è che alla fine si scopra che non esiste un meccanismo unico in grado di spiegare l'origine della modularità in tutte le circostanze. Al contrario potrebbe esserci una moltitudine di meccanismi che contribuiscono all'evoluzione della modularità in alcune situazioni e non in altre [Sovrano e Vallortigara in questo volume]. I modelli di Vita Artificiale si sono già rivelati molto adatti per formulare nuove ipotesi e promettono di farlo ancor di più nel prossimo futuro.

Riferimenti bibliografici

Allman, J. (2000), *Evolving brains*, New York, W. H. Freeman and Company.

Barkow, J., Cosmides, L., & Tooby, J. (1992), *The Adapted Mind: Evolutionary psychology and the generation of culture*, New York, Oxford University Press.

Calabretta, R. & Parisi, D. (in stampa), *Evolutionary connectionism and mind/brain modularity*, in *Modularity. Understanding the development and evolution of complex natural systems*, a cura di W. Callebaut e D. Rasskin-Gutman, Cambridge, The MIT Press.

Calabretta, R., Di Ferdinando, A., Wagner, G. P. & Parisi, D. (in press), What does it take to evolve behaviorally complex organisms? *Biosystems*.

Calabretta, R., Nolfi, S., Parisi, D. & Wagner, G. P. (1998a), *A case study of the evolution of modularity: towards a bridge between evolutionary biology, artificial life, neuro- and cognitive science*, in *Proceedings of the Sixth International Conference on Artificial Life*, a cura di C. Adami, R. Belew, H. Kitano & C. Taylor, Cambridge, The MIT Press.

Calabretta, R., Nolfi, S., Parisi, D., & Wagner, G. P. (1998b), *Emergence of functional modularity in robots*, in *From Animals to Animats 5*, a cura di R. Pfeifer, B. Blumberg, J.A. Meyer & S.W. Wilson, Cambridge, The MIT Press.

Calabretta, R., Nolfi, S., Parisi, D. & Wagner, G. P. (2000), *Duplication of modules facilitates the evolution of functional specialization*, in "Artificial Life", 2000, n. 6, pp. 69-84.

Cangelosi A., Parisi D., & Nolfi S. (1994), *Cell division and migration in a 'genotype' for neural networks*, in "Network: Computation in Neural Systems", 1994, n. 5, pp. 497-515.

Chomsky, N. (1980), *Rules and representations*, New York, Columbia University Press.

Cosmides, L. & Tooby, J. (1994), *The evolution of domain specificity: the evolution of functional organization*, in *Mapping the Mind: Domain Specificity in Cognition and Culture*, a cura di L.A. Hirschfeld & S.A. Gelman, Cambridge, The MIT Press.

Di Ferdinando, A., Calabretta, R. & Parisi, D. (2001), *Evolving modular architectures for neural networks*, in *Proceedings of the Sixth Neural Computation and Psychology Workshop Evolution, Learning, and Development*, a cura di R. French & J. Sougné, London, Springer Verlag.

- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996), *Rethinking innateness. A connectionist perspective on development*, Cambridge, The MIT Press.
- Fodor, J. (1983), *The modularity of mind*, Cambridge, The MIT Press.
- Fodor, J. (2000), *The mind doesn't work that way: The scope and limits of computational psychology*, Cambridge, The MIT Press.
- Gould, S. J. (1997), *Evolution: the pleasures of pluralism*, in "New York Review of Books", June 1997, n. 26.
- Gould, S. J. & Vrba, E. S. (1982), *Exaptation - a missing term in the science of form*, in "Paleobiology", n. 8, pp. 4-15.
- Futuyma, D. J. (1998), *Evolutionary Biology*, Sunderland, Sinauer.
- Holland, J. H. (1992), *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, Cambridge, The MIT Press.
- Karmiloff-Smith, A. (2000), *Why babies' brains are not Swiss army knives*, in *Alas, poor Darwin*, a cura di H. Rose and S. Rose, London, Jonathan Cape.
- Newell, A. & Simon, H. A. (1976), *Computer science as empirical inquiry: Symbols and search*, in "Communications of the Association for Computing Machinery", 1976, n. 19, pp. 113-126.
- Nolfi, S. (1997), *Using emergent modularity to develop control system for mobile robots*, in "Adaptive Behavior", 1997, n. 5, pp. 343-364.
- Pinker, S. (1999), *Words and Rules. The Ingredients of Language*, Weidenfeld and Nicolson, New York.
- Pinker, S. & Prince, A. (1988), *On language and connectionism: Analysis of a parallel distributed processing model of language acquisition*, in "Cognition", 1988, n. 28, pp. 73-193.
- Restak, R. M. (1995), *The modular brain*, New York, Simon & Schuster.
- Rumelhart, D. & McClelland, J. (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Cambridge, The MIT Press.
- Spelke, E. S., Breinlinger, K., Macombe, J. & Jacobson, K. (1992), *Origins of knowledge*, in "Psychological Review", 1992, n. 99, pp. 605-632.
- Wagner, G. P., Mezey, J. & Calabretta, R. (in stampa), *Natural Selection and the origin of modules*, in *Modularity. Understanding the development and evolution of complex natural systems*, a cura di W. Callebaut e D. Rasskin-Gutman, Cambridge, The MIT Press.
- Wagner, G. P. & Altenberg, L. (1996), *Complex adaptations and the evolution of evolvability*, in "Evolution", 1996, n. 50, pp. 967-976.
- Wynn, K. (1992), *Addition and subtraction by human infants*, in "Nature", 1992, August 27, pp. 749-750.