

UNA PANORAMICA DEI SISTEMI COGNITIVI ARTIFICIALI

DAVID VERNON¹, GIORGIO METTA^{1,2}, GIULIO SANDINI²

¹LIRA-Lab, DIST, Università di Genova

²Istituto Italiano di Tecnologia, RBCS, Genova

La nostra indagine sui sistemi artificiali si propone di illustrare i vari approcci al problema, partendo da considerazioni generali sulla possibilità dello sviluppo autonomo di certe capacità cognitive, spaziando dal paradigma cognitivista a quelli cosiddetti emergenti. Aspetti come la capacità di raggiungere obiettivi generici, l'adattività al contesto e alle diverse situazioni, l'anticipazione degli eventi futuri o la predizione delle conseguenze delle azioni sono tutti elementi di questa immagine dei sistemi cognitivi.

Successivamente, tratteremo alcuni aspetti relativi alle architetture cognitive derivate dai diversi paradigmi. In ognuna di queste aree metteremo in evidenza le implicazioni e i problemi che si incontrano se si considera un approccio basato sullo sviluppo, che sembra naturale quando il modello di riferimento è quello dei sistemi biologici. Nel nostro caso lo sviluppo è sinonimo di sviluppo postnatale o ontogenetico. Infine, concludiamo con un riassunto delle caratteristiche chiave delle architetture cognitive nel caso in cui queste debbano soddisfare le condizioni al contorno necessarie per poter sviluppare delle capacità mentali in maniera autonoma.

Se sento dimentico, se vedo ricordo, se faccio capisco [Confucio].

1 Introduzione

La scienza e l'ingegneria dei sistemi artificiali con capacità cosiddette mentali ha una storia ormai lunga, le cui origini risalgono agli anni 50 del secolo passato. Per capacità mentali intendiamo tutti gli aspetti di un comportamento stabile e resistente alle perturbazioni dell'ambiente, ma anche gli aspetti relativi alla percezione, all'azione, alle capacità di deliberazione e alle motivazioni. Nel seguito usiamo il termine cognizione come sinonimo di queste capacità (M. L. Anderson, 2003).

La cognizione implica l'abilità di comprendere come potrebbero essere le cose, non nel presente ma piuttosto in un certo istante di tempo futuro; comporta inoltre la capacità di prendere in considerazione queste ipotesi nel momento di decidere come agire. Il fatto di ricordare ciò che è accaduto in una determinata circostanza nel passato può aiutare ad anticipare eventi futuri. La memorizzazione del passato non è fine a se stessa, bensì si pensa che possa avere proprio lo scopo di anticipare le conseguenze delle proprie azioni o le risposte dell'ambiente. Ne consegue che anche la memoria è un elemento essenziale: utilizzare il passato per predire il futuro (Berthoz, 2000) e, successivamente, assimilare quello che effettivamente accade al fine di adattare e migliorare questa capacità di anticipazione, in un ciclo virtuoso intrinseco nel processo di azione e percezione. La

cognizione si *libera del presente*, in un modo che permette al sistema di agire in modo efficace, di adattarsi e di migliorare.

Tuttavia, cosa fa sì che un'azione sia quella corretta in una data situazione? Quale tipo di comportamento rende possibile la cognizione? Queste domande affrontano un'altra dimensione del problema: che cosa motiva la cognizione? Come viene guidata la percezione? Come vengono selezionate le azioni? Che cosa rende possibile la cognizione? Le abilità cognitive possono migliorare attraverso l'apprendimento, ma che cosa è necessario per iniziare questo processo? Cosa guida il processo di sviluppo? In altre parole, oltre alla percezione, l'autonomia, l'azione, l'anticipazione, l'assimilazione e l'adattamento, ci sono motivazioni intrinseche che devono probabilmente essere considerate. Tali motivazioni guidano l'attenzione percettiva, la selezione delle azioni, lo sviluppo del sistema, dando origine a comportamenti che sono efficienti a lungo termine; il che, poi, è in fondo quello che ci aspettiamo da un sistema cognitivo.

Partendo da questo punto di vista, un sistema è detto cognitivo quando produce comportamenti efficaci, attraverso la percezione, l'azione, la deliberazione, la comunicazione e per mezzo dell'interazione diretta o mediata attraverso comportamenti sociali con l'ambiente esterno. L'impronta della cognizione si manifesta tramite la possibilità, per il sistema in questione, di funzionare efficacemente in circostanze non previste in modo esplicito al momento della progettazione del medesimo. In altre parole, perchè si possa parlare di un sistema cognitivo, è richiesto un qualche grado di plasticità e allo stesso tempo che questo sia resistente al presentarsi di eventi inaspettati (Vernon, 2006b).

Alcuni autori, nel discutere lo sviluppo dei sistemi cognitivi, vanno anche oltre tale affermazione. Per esempio Brachman (Brachman, 2002) definisce un sistema cognitivo computazionale come quello che – oltre a essere capace di ragionare, di imparare dall'esperienza, di migliorare le proprie prestazioni con il passare del tempo e di rispondere in modo intelligente a situazioni mai sperimentate in precedenza – sarebbe anche in grado di spiegare cosa stia facendo e perchè lo stia facendo. Tale abilità gli permetterebbe di identificare i problemi potenziali nel seguire una certa strada nel portare a termine un determinato compito, oppure di sapere quando si rendono necessarie delle nuove informazioni per completarlo. Hollnagel (Hollnagel & Woods, 1999) suggerisce che i sistemi cognitivi debbano essere capaci di esaminare un problema da più di un punto di vista ed essere in grado di utilizzare la conoscenza di se stessi e dell'ambiente circostante per pianificare e modificare le proprie azioni sulla base di tale conoscenza. Quindi, per alcuni studiosi la cognizione implica anche la capacità di riflettere su se stessi, oltre a quelle di adattamento e anticipazione.

La cognizione può essere considerata come il processo attraverso il quale il sistema ottiene comportamenti adattivi, anticipatori e autonomi, che includono chiaramente la percezione mediata dal proprio corpo e la generazione delle azioni. Tale punto di vista si scontra con quello di chi considera la cognizione come un componente distinto o un sotto-sistema del cervello – un modulo della mente – che si occupa solo della pianificazione razionale e del ragionamento e agisce sulle rappresentazioni prodotte dall'apparato percettivo, decidendo quindi quali azioni dovrebbero essere intraprese

subito dopo. Il punto di vista adattivo, anticipatorio e autonomo rispecchia le posizioni di Freeman e Núñez, i quali, in *Reclaiming Cognition* (Freeman & Nunez, 1999), ribadiscono il primato dell'azione, dell'intenzione e dell'emozione nel processo cognitivo. Nel passato, come vedremo più avanti, la cognizione è stata considerata da molti studiosi come disincarnata¹ in principio e, in pratica, un'aggiunta all'elaborazione di simboli relativi alla percezione e azione. Tuttavia, questa visione è in fase di mutamento e perfino i fautori degli approcci più computazionalisti e matematici adesso vedono una relazione molto più stretta tra la percezione, l'azione e la cognizione. Per esempio, consideriamo Anderson *et al.* che affermano: “Ci sono motivi per supporre che la natura della cognizione sia fortemente determinata dal sistema percettivo-motorio, come i fautori della cognizione incarnata e situata hanno sempre sostenuto” (J. R. Anderson et al., 2004), e Langley, che dichiara che “gli stati mentali si basano sempre su stati fisici reali o immaginari e gli operatori spaziali² si riferiscono sempre a capacità primitive che sono alla fine azioni che possono quindi essere eseguite” (Langley, 2005). Il nostro obiettivo, in questo articolo, è quello di fornire al lettore una panoramica completa dei diversi approcci dell'ingegneria dei sistemi cognitivi artificiali, con una particolare attenzione rivolta alle tematiche dello sviluppo attraverso l'interazione tra il sistema e l'ambiente.

Inizieremo con una presentazione dei vari paradigmi della cognizione, mettendo in evidenza le loro differenze ma anche i loro punti in comune. Successivamente, passeremo in rassegna diverse architetture cognitive elaborate a partire da tali paradigmi e presenteremo un'analisi comparativa, in termini di quelle che pensiamo essere le caratteristiche chiave quali, per esempio, la necessità di avere un corpo, la percezione, l'azione, l'anticipazione, l'adattamento, le motivazioni e l'autonomia. Identificheremo quindi una serie di aspetti di base condivisi dai diversi paradigmi della cognizione. Concluderemo poi con un riassunto delle caratteristiche chiave che dovrebbero possedere i sistemi cognitivi caratterizzati da uno sviluppo autonomo delle capacità mentali.

2 Paradigmi nello studio della cognizione

Esistono approcci diversi nel tentativo di comprendere i sistemi cognitivi; ognuno elabora ipotesi significativamente diverse sulla natura della cognizione, il suo obiettivo e il modo secondo il quale il sistema cognitivo dovrebbe essere analizzato e sintetizzato. Tra di essi possiamo distinguere due grandi gruppi principali: l'approccio *cognitivista*³, che si fonda sul processamento simbolico e rappresentazionalista; l'approccio dei sistemi *emergenti*, che include i sistemi connessionisti, quelli dinamici e quelli enattivi, basati in modo più o meno pregnante sui principi dell'auto-organizzazione (A. Clark, 2001; F. J. Varela, 1992).

¹ Disembodied

² Problem-space operators

³ Cognitivism: corrente di studi filosofici, psicologici, etologici che si occupa di attività conoscitive, come la percezione, il linguaggio, ecc. cercando di elaborarne teorie esplicative che riconducono i comportamenti osservabili a piani e programmi mentali profondi (DeMauro, Dizionario della lingua italiana).

Gli approcci cognitivisti corrispondono al classico e ancora comune punto di vista secondo il quale “la cognizione è un tipo di computazione” definita da rappresentazioni simboliche e i sistemi cognitivisti “istanzano tali rappresentazioni fisicamente come simboli e [...] il loro comportamento è una conseguenza causale di operazioni effettuate su tali simboli (Pylyshyn, 1984)”. I sistemi connessionisti, dinamici e quelli enattivi, raggruppati insieme sotto la denominazione generale di sistemi emergenti, sostengono un punto di vista in contrapposizione a quello dell’elaborazione dell’informazione – cioè un approccio che considera la cognizione come “simbolica, razionale, incapsulata, strutturata e algoritmica”. Per i sostenitori del punto di vista dei sistemi emergenti la cognizione è invece sostanzialmente emergente, auto-organizzata e dinamica (Kelso, 1999; Esther Thelen & Smith, 1998).

Come vedremo più avanti, l’enfasi dell’approccio cognitivista e di quello emergente differiscono profondamente e sostanzialmente: distinzione che va ben oltre una semplice differenza fondata sulla manipolazione simbolica. Una possibilità per descrivere questa differenza è quella di confrontare i vari paradigmi secondo una tassonomia organizzata su dodici diversi aspetti: operazioni elementari, struttura rappresentazionale, legame semantico con l’ambiente, restrizioni temporali, epistemologia, incarnazione, percezione, azione, anticipazione, adattamento, motivazioni e autonomia⁴. La tabella I è un tentativo per descrivere questi diversi aspetti in maniera sintetica che verranno in seguito trattati singolarmente.

Paradigmi cognitivisti ed emergenti alla cognizione		
Caratteristiche	Cognitivista	Emergente
Operazioni elementari	Manipolazione sintattica di simboli	Organizzazione autonoma e concorrente di una rete
Struttura rappresentazionale	Insiemi di elementi simbolici	Stati globali del sistema
Legame semantico con l’ambiente	Associazione tra percetti e simboli	Costruzione di capacità
Restrizioni temporali	Non legato all’ambiente	Sincronicità con l’ambiente
Epistemologia	Indipendente dall’agente	Dipendente dall’agente
Incarnazione	Non necessario	La cognizione richiede l’incarnazione
Percezione	Rappresentazioni astratte e simboliche	Risposta a una perturbazione
Azione	Conseguenza causale della manipolazione di simboli	Perturbazione dell’ambiente da parte del sistema
Anticipazione	Ragionamento procedurale o simbolico che utilizza di solito un modello a priori	Esplorazione autonoma dello spazio degli stati

⁴ Esistono molte definizioni possibili di autonomia: esse spaziano dall’abilità di un sistema di contribuire alla propria persistenza come in (Bickhard, 2000), fino alle caratteristiche organizzative di auto-mantenimento delle creature viventi – sistemi dissipativi lontani dall’equilibrio – che utilizzano le proprie capacità per regolare l’interazione con il mondo e internamente con il proprio corpo, in modo da rimanere adattati all’ambiente stesso (Christensen & Hooker, 2000a).

Adattamento	Apprendimento di nuova conoscenza	Sviluppo di nuove dinamiche
Motivazioni	Risoluzione di un' <i>impasse</i>	Aumento dello spazio delle interazioni
Rilevanza dell'autonomia	Non necessaria	La cognizione richiede l'autonomia

Tabella I: confronto del paradigma cognitivista e di quello emergente.

Operazioni elementari. I sistemi cognitivisti si basano su regole (cioè processi sintattici) che manipolano elementi simbolici normalmente, ma non necessariamente, in un modo sequenziale. I sistemi emergenti sfruttano processi di auto-organizzazione, auto-produzione, auto-mantenimento e auto-sviluppo, attraverso l'interazione concorrente di una rete di componenti distribuiti che interagiscono tra di loro.

Struttura rappresentazionale. I sistemi cognitivisti utilizzano insiemi di elementi simbolici che si riferiscono a eventi del mondo esterno. Questi consistono tipicamente nel prodotto di una descrizione⁵ di un progettista umano, di solito ma non necessariamente, discreta⁶ e locale. Le rappresentazioni dei sistemi emergenti consistono in stati del sistema globale codificati nell'organizzazione dinamica della rete di componenti del sistema.

Legame semantico con l'ambiente. Le rappresentazioni simboliche dei sistemi cognitivisti si fondano sull'identificazione percelto-simbolo sia attraverso il lavoro del progettista, sia tramite una fase di apprendimento delle associazioni. Tali rappresentazioni sono accessibili all'interpretazione umana diretta. I sistemi emergenti fondano le proprie rappresentazioni sulla costruzione di capacità adattive e anticipatorie atte a preservare l'autonomia del sistema in quanto tale. Tali rappresentazioni hanno un senso solamente e finché contribuiscono alla funzionalità continuata del sistema e non sono accessibili all'interpretazione umana diretta.

Restrizioni temporali. I sistemi cognitivisti non sono necessariamente sincroni con gli eventi del mondo esterno. Quelli emergenti sono invece legati al mondo esterno e operano in modo sincronico, in tempo reale, con gli eventi che accadono nel loro ambiente circostante.

Epistemologia. I sistemi cognitivisti posseggono un'epistemologia assoluta condivisa tra agenti distinti che è garantita in virtù della loro visione positivista della realtà: ogni agente si trova inserito in un ambiente circostante, la cui struttura e semantica sono indipendenti dalla cognizione del sistema. L'epistemologia dei sistemi emergenti consiste nel risultato soggettivo di una storia di esperienze condivise in modo consensuale tra agenti filogeneticamente compatibili.

⁵ "Descrizione" nel senso che il progettista è un osservatore esterno della relazione tra un sistema cognitivo e il suo ambiente circostante; in tal modo la struttura rappresentazionale coincide col modo in cui il progettista vede tale relazione.

⁶ "Discreta" come opposto di "continua".

Incarnazione. In linea di massima, i sistemi cognitivisti non hanno bisogno di essere incarnati in virtù delle proprie radici nel funzionalismo (che asserisce che la cognizione è indipendente dalla piattaforma fisica nella quale viene implementata (Freeman & Nunez, 1999)). D'altra parte, i sistemi emergenti sono intrinsecamente incarnati e l'istanziamento fisico gioca un ruolo costitutivo diretto nel processo cognitivo (Gardner, 1993; Krichmar & Edelman, 2006; Vernon, 2006b).

Percezione. Nei sistemi cognitivisti la percezione fornisce un'interfaccia tra il mondo esterno e la rappresentazione simbolica di quel mondo. La percezione astrae rappresentazioni spazio-temporali fedeli del mondo esterno a partire dai dati sensoriali. Nei sistemi emergenti la percezione è un cambiamento nello stato del sistema in risposta alle perturbazioni dell'ambiente esterno, al fine di mantenere la stabilità.

Azione. Nei sistemi cognitivisti le azioni sono conseguenze causali dell'elaborazione simbolica di rappresentazioni interne. Nei sistemi emergenti le azioni sono perturbazioni dell'ambiente effettuate dal sistema.

Anticipazione. Nei sistemi cognitivisti l'anticipazione consiste, tipicamente, nella pianificazione, utilizzando qualche forma di ragionamento di tipo procedurale o probabilistico attraverso un modello *a priori*. L'anticipazione nel paradigma emergente richiede che il sistema visiti un numero di stati all'interno del proprio spazio degli stati percezione-azione, senza necessariamente eseguire le azioni associate.

Adattamento. Per il cognitivismo l'adattamento implica l'acquisizione di una nuova conoscenza, mentre per i sistemi emergenti comporta un'alterazione strutturale o riorganizzazione per mettere in atto un nuovo insieme di dinamiche (Schoener, 2006).

Motivazioni. Le motivazioni influenzano la percezione (attraverso l'attenzione), l'azione (attraverso la selezione delle stesse), l'adattamento (tramite i parametri che governano i cambiamenti del sistema) e la risoluzione di un'*impasse* in un sistema cognitivista o l'ampliamento dello spazio di interazione in un sistema emergente (Von Hofsten, 2003, , 2004).

Autonomia. L'autonomia non è necessariamente implicita nel paradigma cognitivista, mentre gioca un ruolo cruciale in quello emergente, dato che la cognizione è il processo nel quale un sistema autonomo diventa funzionale ed efficace.

Nel seguito, descriviamo il paradigma cognitivista e quello emergente, oltre ad alcuni approcci ibridi; entriamo nei dettagli e presentiamo una serie di esempi di sistemi realizzati seguendo i vari approcci. È bene notare che ancora oggi, nonostante il lento spostamento della comunità scientifica verso gli approcci emergenti, i sistemi cognitivisti sono quelli più avanzati mentre esiste poco riguardo all'approccio enattivo che sembra essere più un metodo di analisi che uno di sintesi di sistemi cognitivi.

2.1 Modelli cognitivisti

2.1.1 Una panoramica dei modelli cognitivisti

Le scienze cognitive hanno la loro origine nella cibernetica, nei primi sforzi di formalizzazione di ciò che fino a quel momento era stato un trattamento metafisico della cognizione (F. J. Varela, 1992). L'intenzione dei pionieri della cibernetica era quello di creare una scienza della mente, fondata sulla logica. Questi progenitori includono McCulloch e Pitts con il loro articolo seminale 'Un calcolo logico immanente nell'attività del sistema nervoso' (McCulloch & Pitts, 1943). A tale ondata iniziale dello sviluppo di una scienza della cognizione seguì nel 1956 lo sviluppo di un approccio denominato *cognitivismo*. Il cognitivismo asserisce che la cognizione richiede calcoli definiti su rappresentazioni interne da cui deriva la conoscenza, in un processo nel quale l'informazione del mondo viene astratta dalla percezione e rappresentata utilizzando una struttura di dati simbolici appropriati, sui quali si ragiona, e quindi utilizzata per pianificare e agire nel mondo. Tale approccio è stato anche denominato da molti come approccio cognitivo dell'*elaborazione dell'informazione* (o manipolazione simbolica) (Haugland, 1982; Kelso, 1999; Marr, 1977; Newell & Simon, 1975; Pinker, 1984; Esther Thelen & Smith, 1998; F. J. Varela, 1992). Senza dubbio il cognitivismo è stato l'approccio predominante della cognizione fino a oggi, e ancora risulta quello prevalente. La disciplina della scienza cognitiva viene spesso identificata, in particolare, con questo approccio (Freeman & Nunez, 1999; Kelso, 1999). Vedremo più avanti che questo non è affatto l'unico paradigma e, infatti, esistono indizi sull'allontanamento della disciplina dalle sue interpretazioni più forti (A. Clark, 2001). Per i sistemi cognitivisti, la cognizione è rappresentazionale in un modo particolare e forte: porta con sé la manipolazione di rappresentazioni simboliche dello stato e del comportamento del mondo esterno, al fine di facilitare interazioni appropriate, adattive, anticipatorie ed efficaci, e l'immagazzinamento della conoscenza ottenuta da tale esperienza per ragionare in modo ancora più efficace in futuro (Hollnagel & Woods, 1999). La percezione ha a che vedere con l'astrazione di rappresentazioni spazio-temporali fedeli del mondo esterno, a partire da dati sensoriali. Anche il ragionamento è simbolico: un sistema procedurale nel quale le rappresentazioni esplicite del mondo esterno vengono manipolate per inferire cambiamenti probabili nella configurazione del mondo (e la percezione che accompagna l'alterazione della configurazione) come conseguenza causale delle azioni. Nella maggior parte degli approcci cognitivisti, ove si è considerata la creazione di sistemi cognitivi artificiali, le rappresentazioni simboliche (o la struttura della rappresentazione nei sistemi capaci di apprendimento) sono il prodotto descrittivo di un progettista umano. Questo è significativo poichè indica che possono essere direttamente esaminate e comprese o interpretate dagli esseri umani e che la conoscenza semantica può venire incorporata direttamente nel sistema, ed estratta direttamente dal medesimo. È stato discusso come questo sia proprio il fattore chiave limitante dei sistemi cognitivisti: tali rappresentazioni dipendenti dal programmatore effettivamente viziano il sistema (o "lo rendono cieco" (Winograd & Flores, 1986)) e lo costringono a una descrizione idealizzata dipendente e conseguente ai requisiti cognitivi dell'attività umana. Tale approccio funziona finchè il sistema non deve lavorare o agire troppo lontano dalle condizioni sotto le quali tali descrizioni sono state formulate. Quanto più ci si allontana, maggiore è la 'distanza

semantica⁷ (Smeulders, Worrying, Santini, Gupta, & Jain, 2000) tra la percezione e la possibile interpretazione, una distanza che normalmente viene ridotta attraverso l'aggiunta di conoscenza ulteriore da parte del programmatore o il mantenimento di vincoli che sono funzione delle attese (del programmatore) (Pauli & Sommer, 2002), al fine di rendere il sistema adatto a un determinato ambito di problemi. Il problema fondamentale è che questa procedura di aggiunta e raffinamento del sistema sembra essere ricorsiva e non arrivare mai a un sistema generale per vari ambiti di funzionamento. Il cognitivismo immagina in modo positivista che “il mondo che percepiamo sia isomorfo con le nostre percezioni del medesimo, come in un ambiente geometrico” (Shepard & Hurwitz, 1984). L'obiettivo finale del cognitivismo, per un cognitivista, consiste nel ragionare in modo simbolico su tali rappresentazioni, in modo da generare il necessario comportamento adattivo, anticipatorio e, possibilmente, adatto a raggiungere degli obiettivi. Tale approccio alla cognizione metterà in campo, tipicamente, un arsenale di tecniche incluso l'apprendimento automatico, la modellazione probabilistica e altre tecniche in un tentativo di affrontare la natura inerentemente incerta, tempo-variante e incompleta dei dati sensoriali che guidano la struttura rappresentazionale. Questo non cambia il fatto che la struttura rappresentazionale venga ancora definita dalle descrizioni dei progettisti. La rilevanza di questo aspetto diventerà palese nelle sezioni successive.

2.1.2 Il cognitivismo e l'intelligenza artificiale

Considerato che la ricerca sul cognitivismo e quella sull'intelligenza artificiale hanno connessioni molto forti⁸, vale la pena soffermarsi sulla relazione esistente tra gli approcci cognitivisti e quelli classici dell'Intelligenza Artificiale, in modo particolare quello del “Sistema simbolico⁹” di Newell e Simon (Newell & Simon, 1975), che ha avuto una straordinaria influenza sul modo con cui riflettiamo sull'intelligenza, sia naturale che artificiale.

Nell'articolo del 1976 di Newell e Simon, vengono formulate due ipotesi.

- A. *Ipotesi del Sistema Simbolico*. Un sistema simbolico possiede i mezzi necessari e sufficienti per generare azioni generiche intelligenti.
- B. *Ipotesi della ricerca euristica*. Le soluzioni ai problemi vengono rappresentate come strutture simboliche. Un sistema simbolico esercita la propria intelligenza nella soluzione di problemi attraverso la ricerca, ossia attraverso la generazione e la progressiva modifica delle strutture simboliche, finchè non viene prodotta una struttura che rappresenta la soluzione.

⁷ Semantic gap.

⁸ Alcuni considerano l'Intelligenza Artificiale come un discendente diretto del cognitivismo: “[...] lo studio positivista e riduttivo della mente ottenne una straordinaria popolarità attraverso una dottrina relativamente recente chiamata *Cognitivismo*, un approccio che plasma la creazione di un nuovo ambito –la Scienza Cognitiva– e il suo prodotto principale: l'Intelligenza Artificiale” (enfasi nell'originale). (Freeman & Nunez, 1999)

⁹ Physical symbol system.

La prima ipotesi implica che qualunque sistema che mostri un'intelligenza di tipo generale è un sistema simbolico e qualunque sistema fisico simbolico di dimensioni sufficienti può essere configurato in qualche modo ('organizzato ulteriormente') affinché esibisca un'intelligenza di tipo generale.

La seconda ipotesi fa riferimento all'asserzione che i sistemi simbolici risolvono i problemi attraverso la ricerca euristica, cioè "strutture composte da successive generazioni di potenziali soluzioni", in un modo efficace ed efficiente. "Il ruolo dell'intelligenza, quindi, consiste nell'evitare la sempre presente minaccia dell'esplosione esponenziale della ricerca".

Un sistema fisico simbolico è equivalente a un sistema formale automatico (Haugland, 1982). Si tratta di "una macchina che evolve nel tempo una collezione di strutture simboliche". Un simbolo è un elemento fisico che può comparire come componente di un altro tipo di entità chiamata *espressione*: strutture di espressioni/simboli sono particolari configurazioni di simboli ed elementi. Oltre alle strutture simboliche, il sistema comprende processi che operano su espressioni per produrre altre espressioni: "i processi di creazione, modifica, riproduzione e distruzione". Un'espressione può designare un oggetto e quindi il sistema può sia "modificare l'oggetto in sé o comportarsi in un determinato modo, dipendendo dall'oggetto", oppure, se l'espressione designa un processo, il sistema interpreta l'espressione portando a termine il processo (vedere Figura 1).

Nelle parole di Newell e Simon: *"I sistemi simbolici sono collezioni di configurazioni e processi, questi ultimi essendo in grado di produrre, distruggere e modificare i primi. Le proprietà più importanti di queste configurazioni consistono nel fatto che possono designare oggetti, processi, oppure altre configurazioni e, quando designano processi, possono essere interpretati. Interpretare significa portare a termine il processo designato. Le due classi più rilevanti di sistemi simbolici con le quali interagiamo abitualmente sono gli esseri umani e i computer"*.

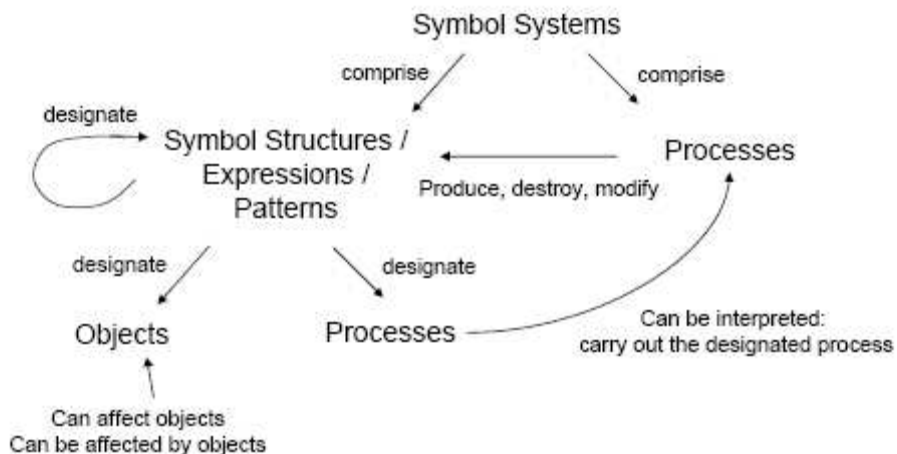


Figura 1: L'essenza di un sistema simbolico (Newell & Simon, 1975).

Quello che ci sembra importante nella descrizione di un sistema simbolico è il fatto che si tratta di un approccio più generale rispetto al solito *ritratto* dei sistemi di manipolazione

simbolica, nel quale i simboli designano solamente oggetti, per cui ci troviamo di fronte a un sistema di processi che produce, distrugge e modifica simboli, e niente di più. Il punto di vista originale di Newell e Simon è molto più sofisticato. In esso, esistono due aspetti ricorsivi: i processi possono produrre processi e le configurazioni possono designare altre configurazioni (che, certamente, possono essere processi). Questi due cicli ricorsivi sono strettamente correlati fra loro. Non solamente il sistema può costruire rappresentazioni ancora più astratte e ragionare su tali rappresentazioni, ma può anche modificare se stesso come una funzione sia della propria elaborazione, che è rappresentata dallo stato corrente, sia delle proprie rappresentazioni.

I sistemi simbolici possono venire istanziati e il loro comportamento dipende a sua volta dai dettagli del sistema simbolico, dai suoi simboli, operazioni e interpretazioni, e non dalla particolare forma della istanziazione.

L'ipotesi del sistema fisico di simboli asserisce che un sistema simbolico possiede i mezzi necessari e sufficienti per un'intelligenza generale. Da quanto appena detto sui sistemi simbolici, ne deriva che i sistemi intelligenti, sia naturali che artificiali, sono effettivamente equivalenti, poichè l'istanziamento è, almeno in linea di principio, priva di conseguenze.

La grande maggioranza dei sistemi cognitivisti è identica e identificabile con i sistemi basati sull'ipotesi di Newell e Simon.

2.1.3 Alcuni sistemi cognitivisti

L'uso della conoscenza simbolica esplicita è stato applicato a molti sistemi cognitivisti, e portiamo qui a esempio, un sistema cognitivo che supporta un sistema di visione artificiale (Nagel, 2004) sviluppato per l'interpretazione di sequenze video di traffico cittadino e la generazione di una descrizione in linguaggio naturale delle scene osservate. Questo sistema funziona trasferendo rappresentazioni dei segnali video (le immagini) alle rappresentazioni simboliche attraverso diversi strati di processi e, infine, rappresentando il comportamento dei veicoli attraverso alberi e grafi, detti *Situation Graph Tree* (SGT). L'interpretazione automatica di tale rappresentazione del comportamento viene generata attraverso la traduzione degli SGT in un programma di logica (basato sulla logica di Horn¹⁰) (Arens & Nagel, 2002, , 2005; Arens, Ottlick, & Nagel, 2002).

Le ipotesi cognitiviste sono inoltre rappresentate e istanziate in un approccio basato su modelli (Moeller, Neumann, & Wessel, 1999; Neumann & Moeller, 2005) che utilizza delle logiche descrittive, basate a loro volta sulla logica dei predicati del primo ordine, per rappresentare e ragionare su concetti di alto livello, come la configurazione spazio-temporale degli oggetti e degli eventi.

Sono state proposte delle strutture probabilistiche in alternativa (o qualche volta in aggiunta a (Neumann & Moeller, 2005)) ai sistemi di ragionamento di tipo deterministico. Per esempio, Buxton e i suoi collaboratori descrivono un sistema cognitivo con elementi di visione artificiale per interpretare le attività di operatori umani esperti. Questo sfrutta reti di decisione dinamiche (DDN) – un'estensione delle reti di

¹⁰ Fuzzy metric temporal Horn logic.

Bayes che incorpora dipendenze dinamiche e la teoria dell'utilità (Buxton, 2002) – per riconoscere e ragionare sulle attività, reti di funzioni base (*time-delay radial basis function*, TDRBFN) e modelli Markoviani a variabili latenti (*hidden markov model*, HMM) per il riconoscimento dei gesti. Sebbene il sistema utilizzi tecniche di apprendimento per creare modelli dei gesti, il processo generale di ragionamento simbolico, anche se probabilistico, necessita ancora che il progettista umano identifichi i vincoli del contesto e le loro dipendenze causali (la ricerca attuale è orientata all'apprendimento automatico delle strategie di controllo basate sia su informazione contestuale che su quella relativa al compito finale) (Buxton & Howell, 2002; Buxton, Howell, & Sage, 2002; Sage, Howell, & Buxton, 2005)¹¹. Progressi recenti, nella costruzione automatica e nell'utilizzo di modelli simbolici di comportamento, sono stati descritti a partire da dati sensoriali usando la logica induttiva (Cohn et al., 2005).

La dipendenza degli approcci cognitivisti dalle rappresentazioni del mondo decise dal progettista viene anche esemplificata bene dai sistemi basati sulla conoscenza, come per esempio quelli fondati sulle ontologie. In particolare, Maillot e i suoi collaboratori (Maillot, Thonnat, & Boucher, 2003) descrivono una struttura per un sistema visivo cognitivo basato sull'ontologia indirizzato alla mappatura tra la conoscenza di dominio e quella risultante dall'elaborazione delle immagini, utilizzando un'ontologia dei concetti visivi che include i concetti di spazio-tempo, di tessitura visiva e quelli di colore.

Un'altra architettura che portiamo come esempio (Chella, Frixione, & Gaglio, 1997) per un sistema visivo cognitivo comprende un livello sub-simbolico, che sfrutta una rappresentazione di tipo 2½D incentrata sull'osservatore e fondata sui dati sensoriali, un livello concettuale pre-linguistico fondato su rappresentazioni 3D tramite superquadriche incentrate sull'oggetto, e infine un livello linguistico che utilizza una base di conoscenza simbolica. Un processo attentivo connette il livello concettuale con quello linguistico.

Sono stati proposti sistemi adattivi per l'osservazione e l'interpretazione dell'attività umana che configurano dinamicamente i propri processi per operare in un contesto specifico (Crowley, 2005) come per esempio quello che usa la visione per controllo automatico di veicoli (Dickmanns, 2004).

Town e Sinclair presentano una struttura cognitiva che coniuga i processi di basso livello (stima del moto, inseguimento di contorni, classificazione delle regioni, riconoscimento di volti, modelli della forma, e operatori per il raggruppamento percettivo) con il processo di alto livello, utilizzando un'ontologia fondata sul linguaggio e reti Bayesiane adattive. Il sistema è autoreferenziale nel senso che mantiene una rappresentazione interna dei propri obiettivi e ipotesi attuali. Quindi le inferenze visive possono essere eseguite elaborando la struttura delle frasi in questo linguaggio. Il suddetto linguaggio adotta un approccio che è la quintessenza del cognitivismo simbolico e rappresentazionalista, anche se utilizza modelli probabilistici, dato che richiede che il progettista identifichi le “giuste ipotesi strutturali” e le distribuzioni delle probabilità a priori.

¹¹ Si veda (Buxton, 2002) per una panoramica sui modelli probabilistici generativi utilizzati nel caso di apprendimento e la comprensione della dinamica di situazioni osservate attraverso telecamera.

2.2 Approcci emergenti

Gli approcci emergenti partono da un approccio alla cognizione molto diverso. In tale approccio la cognizione consiste nel processo tramite il quale un sistema autonomo diventa funzionale ed efficace nel proprio ambiente. Ottiene tale risultato attraverso un processo di auto-organizzazione attraverso il quale il sistema si ricostruisce continuamente in tempo reale, in modo da mantenere la propria identità operativa attraverso la moderazione dell'interazione mutua sistema-ambiente circostante e la co-determinazione (R. H. Maturana & Varela, 1998). La co-determinazione implica che l'agente cognitivo venga specificato tramite il proprio ambiente circostante e, allo stesso tempo, che il processo cognitivo determini ciò che è reale o significativo per l'agente. In qualche modo, la co-determinazione significa che l'agente costruisce la propria realtà (il proprio mondo) come risultato delle proprie operazioni nel medesimo. In tale contesto, il comportamento cognitivo viene qualche volta definito come l'induzione automatica di un'ontologia: tale ontologia sarà inerentemente specifica dell'incarnazione del sistema e dipendente dalla storia delle interazioni del sistema stesso, cioè, dalle proprie esperienze. In tal senso, per gli approcci emergenti, la percezione è implicita nell'acquisizione di dati sensibili, in modo da attivare l'azione più efficace (R. H. Maturana & Varela, 1998) e dipende dalla ricchezza dell'interfaccia dell'azione (G. H. Granlund, 1999). Non si tratta di un processo nel quale la struttura di un ambiente esterno assoluto viene astratto e rappresentato in un modo più o meno isomorfo.

Sandini e colleghi hanno sostenuto che la cognizione è anche il complemento della percezione (Sandini, Metta, & Vernon, 2004). La percezione *ha a che fare* con l'immediato mentre la cognizione permette di considerare intervalli di tempo più lunghi. La cognizione riflette il meccanismo tramite il quale un agente compensa la natura immediata della percezione e può quindi "adattarsi a", ma anche anticipare, l'azione dell'ambiente circostante che avviene su intervalli di tempo molto maggiori. In altre parole, la cognizione è intrinsecamente legata all'abilità di un agente di agire in modo predittivo: di operare nel futuro e affrontare quello che potrebbe avvenire, non solamente con quanto esiste nel presente.

In contrasto con l'approccio cognitivista, molti approcci emergenti asseriscono che il modello primario per l'apprendimento cognitivo consiste nella costruzione di abilità anticipative, piuttosto che l'acquisizione di conoscenza, e che i processi che guidano l'azione e migliorano la capacità di guidare le azioni mentre le effettuano, sono considerati le capacità basilari di ogni sistema intelligente (Christensen & Hooker, 2000a). Al contrario, il cognitivismo implica un modello astratto contenuto in se stesso che è in principio disincarnato; l'istanziamento fisico dei sistemi non gioca nessun ruolo nel modello della cognizione (Vernon, 2006a, 2006b). D'altra parte, gli approcci emergenti sono intrinsecamente incarnati e l'istanziamento fisico gioca un ruolo essenziale nella cognizione.

2.2.1 Modelli connessionisti

I sistemi connessionisti si basano su processi paralleli di modelli di attivazione distribuita non simbolici, utilizzando proprietà statistiche, piuttosto che regole logiche, per elaborare l'informazione e ottenere un comportamento efficace (Medler, 1998). In tal senso, le istanze delle reti neurali del modello connessionista “sono sistemi dinamici che stimano quelle funzioni che catturano al meglio le regolarità statistiche nei dati” (Smolensky, 1996b). Una trattazione esaustiva del connessionismo non rientra tra gli obiettivi del presente articolo. Per una panoramica delle basi del connessionismo e una selezione di articoli seminali sull'argomento, si possono consultare i libri di Anderson e Rosenfeld: *Neurocomputing: Foundations of Research* (J. A. Anderson & Rosenfeld, 1988) e *Neurocomputing 2: Directions of Research* (J. A. Anderson & Rosenfeld, 1991). Medler fornisce un'introduzione succinta sullo sviluppo del connessionismo in (Medler, 1998), mentre Smolensky rilegge l'argomento da una prospettiva matematica, affrontando temi computazionali, dinamici e statistici (Smolensky, 1996a, 1996b, 1996c, 1996d). Il libro curato da Arbib *Handbook of Brain Theory and Neural Networks* fornisce brevi articoli sulla maggior parte della letteratura rilevante su questo argomento (Arbib, 2002). Le origini del connessionismo risalgono a molto prima dell'era computazionale. Sebbene Feldman e Ballard (Feldman & Ballard, 1982) siano normalmente considerati coloro che introdussero per primi il termine “modelli connessionisti” nel 1982, il termine ‘connessionismo’ venne utilizzato già nel 1932 nel campo della psicologia, da Thorndike (Thorndike, 1932, 1949), per segnalare una forma estesa di associazionismo fondato, per esempio, sui principi connessionisti chiaramente evidenti nel modello di memoria associativa¹² di William James, ma anche anticipando meccanismi come l'apprendimento Hebbiano. Infatti, l'introduzione del libro di Hebb *The Organization of Behaviour* (Hebb, 1949), dove l'autore presenta un algoritmo neurale di addestramento non supervisionato nel quale la forza sinaptica viene aumentata se due neuroni (indicati come pre- e post-sinaptico) sono attivi allo stesso tempo, contiene uno dei primi usi del termine connessionismo (J. A. Anderson & Rosenfeld, 1988, p. 43). Abbiamo già notato come il cognitivismo abbia origine anche dai primi lavori di scienza cognitiva, in particolare quelli di McCulloch e Pitts (McCulloch & Pitts, 1943). McCulloch e Pitts dimostrarono che ogni dichiarazione all'interno di una logica preposizionale può essere rappresentata da una rete di unità di elaborazione semplici e, inoltre, che tali reti hanno, in principio, la capacità computazionale di una macchina di Turing universale. A seconda di come si interpreta la suddetta equivalenza, McCulloch e Pitts contribuirono alla fondazione sia del cognitivismo sia del connessionismo. L'approccio connessionistico venne portato avanti in modo significativo alla fine degli anni 50, con l'introduzione del modello di apprendimento del *perceptrone* di Rosenblatt (Rosenblatt, 1958) e di quello del *Pandemonium* di Selfridge (Selfridge, 1959). Rosenblatt dimostrò che qualunque problema di classificazione, espresso in notazione binaria, può essere risolto attraverso una rete di perceptroni. Sebbene i modelli di apprendimento migliorarono intorno al 1960 con l'introduzione della regola di Widrow-Hoff, o regola delta, per l'apprendimento supervisionato nel modello *Adeline* (Widrow & Hoff, 1960), il problema con le reti di perceptroni risiedeva nel fatto che non esisteva nessun algoritmo di apprendimento che

¹² La raccolta di articoli seminali su *Neurocomputing* (J. A. Anderson & Rosenfeld, 1988) di Anderson e Rosenfeld si apre con il Capitolo XVI ‘Associazione’ di William James’ 1890 Psychology, Briefer Course (James, 1890).

permettesse la determinazione dei pesi delle connessioni tra le unità di ingresso e quelle associative nascoste. Di conseguenza, le reti di perceptron erano effettivamente reti a un solo livello, dato che gli algoritmi di apprendimento potevano solamente regolare la forza della connessione tra le unità nascoste e quelle in uscita, rimanendo determinati in fase di progetto i pesi che regolavano la forza della connessione tra l'ingresso e le unità nascoste. Nel 1969 Minsky e Papert (Minsky & Papert, 1969) dimostrarono che il perceptrone può essere allenato solamente per risolvere problemi linearmente separabili, e non per problemi più generali. In conseguenza di ciò, la ricerca sulle reti neurali e sui modelli connessionistici venne in qualche modo abbandonata (almeno in larga parte). Con le apparenti limitazioni del perceptrone a oscurare il lavoro sulle reti neurali, la ricerca si concentrò maggiormente sulla memoria e sul recupero dell'informazione e, in particolare, sui modelli paralleli di memoria associativa (Hinton & Anderson, 1981). Contributi storici in tale periodo includono il modello dell'Attivazione Interattiva e Competizione (IAC) di McClelland (McClelland, 1981) che introdusse l'idea di associazioni competitive di neuroni mutuamente inibitori e dimostrò l'abilità dei sistemi connessionistici di recuperare informazione sia specifica sia generica a partire da conoscenza memorizzata su casi specifici. Inoltre, durante tale periodo vennero introdotti alcuni modelli connessionistici alternativi, come la teoria della risonanza adattiva (ART) di Grossberg (Grossberg, 1976) e le mappe autorganizzanti (SOM) di Kohonen (Kohonen, 1982) spesso chiamate più semplicemente reti di Kohonen. La teoria ART, introdotta nel 1976, si è evoluta e ampliata in modo considerevole negli ultimi trenta anni, al fine di gestire l'apprendimento di categorie sia in maniera supervisionata sia la classificazione di forme e la predizione (Carpenter & Grossberg, 1995). Le reti di Kohonen producono mappe topologiche nelle quali punti vicini nello spazio d'ingresso sono mappati da un processo di apprendimento, non supervisionato e auto organizzante, in uno stato della rete interna che preserva tale topologia: in altre parole, i punti d'ingresso (punti nello spazio di riferimento) che sono vicini vengono rappresentati nella mappa da punti (nello spazio dei pesi) che sono anch'essi vicini. Una volta che l'auto organizzazione è completa, la rete di Kohonen può essere usata o come memoria auto associativa oppure come classificatore di forme. Le reti neurali simili al perceptrone subirono una rinascita a metà degli anni 80, con lo sviluppo dell'architettura dei processi paralleli distribuiti (PDP) (Rumelhart, McClelland, & Group, 1986) in generale e, in particolare, con l'introduzione dell'algoritmo della propagazione retrograda (*back propagation*) da parte di Rumelhart, Hinton e Williams (Rumelhart, Hinton, & Williams, 1986a, 1986b). L'algoritmo di apprendimento della *back propagation* – conosciuto anche come la regola delta generalizzata o GDR, dato che si tratta della generalizzazione della regola delta di Widrow-Hoff per unità Adaline – superò il limite citato da Minsky e Papert permettendo l'addestramento dei pesi delle connessioni tra le unità di ingresso e quelle nascoste, rendendo possibile quindi l'utilizzo di reti multilivello per l'apprendimento di problemi linearmente non separabili. Sebbene la regola di apprendimento *back propagation* ebbe il maggior impatto attraverso il lavoro di Rumelhart *et al.*, essa fu sviluppata in maniera indipendente anche da Werbos (Werbos, 1974).

Nella scienza cognitiva l'architettura PDP portò a un significativo passaggio dal punto di vista sequenziale (utilizzato nei modelli computazionali della mente fino a quel momento) verso un altro basato su reti di unità operative funzionanti in cooperazione o

competizione, possibilmente in parallelo. Essa contribuì inoltre all'aumentare la consapevolezza dell'importanza della struttura del sistema computazionale nel risultato della computazione. Il modello standard PDP rappresenta una mappatura statica tra i vettori di ingresso come conseguenza della configurazione di propagazione dell'informazione diretta (*feedforward*). D'altra parte, reti ricorrenti che hanno connessioni anche all'indietro per formare circuiti, cioè reti nelle quali i segnali dell'attivazione dell'uscita o delle unità nascoste vengono forniti alla rete come ingresso, sono caratterizzate dalla possibilità di un comportamento dinamico¹³. Forse il tipo più conosciuto di rete ricorrente è quello di Hopfield (Hopfield, 1982). Le reti di Hopfield sono reti ricorrenti completamente connesse che si comportano come una memoria auto-associativa¹⁴, o memoria indirizzabile per il contenuto, che può effettuare un completamento di informazioni anche parziale sulla base del precedente addestramento. Altre reti ricorrenti includono le reti di Elman (Elman, 1990) (con connessioni ricorrenti dalle unità nascoste a quelle d'ingresso) e le reti di Jordan (Jordan, 1986) (con connessioni ricorrenti dalle unità output a quelle input). Le macchine di Boltzmann (Hinton & Sejnowski, 1986) sono varianti delle reti di Hopfield che utilizzano procedure stocastiche piuttosto che procedure di aggiornamento dei pesi di tipo deterministico; questo per evitare problemi con i minimi locali durante l'apprendimento. Le reti di perceptron multilivello e le altre di tipo PDP utilizzano funzioni monotone (funzione a gradino o soglia) oppure le sigmoidi (come funzione di attivazione dei neuroni). L'uso di funzioni di attivazione non monotone, come le Gaussiane, può offrire vantaggi computazionali, per esempio una convergenza più rapida e una maggiore stabilità per problemi che non sono linearmente separabili. Le reti con funzioni base radiali (RBF) (Moody & Darken, 1989) utilizzano altresì le funzioni Gaussiane ma differiscono dai perceptron multilivello per il fatto che la funzione Gaussiana viene utilizzata solamente per il livello nascosto, con i livelli di ingresso e uscita che utilizzano invece un'attivazione lineare. I sistemi connessionisti continuano ad avere una grande influenza sulla scienza cognitiva, sia nello spirito delle reti PDP come per l'approccio di McClelland e Rogers alla semantica (McClelland & Rogers, 2003), o in riferimento ai sistemi ibridi come l'architettura cognitiva connessionista/simbolica di Smolensky e Legendre (Smolensky, 1995; Smolensky & Legendre, 2006).

Una delle motivazioni originali per lavorare sui sistemi emergenti fu la disaffezione per il carattere sequenziale, atemporale e localizzato del cognitivismo fondato sulla manipolazione dei simboli (F. J. Varela, 1992). I sistemi emergenti, d'altra parte, dipendono da architetture parallele, che agiscono in tempo reale e in maniera distribuita. In se stesso tuttavia, tale cambiamento nell'enfasi non è sufficiente per costituire un nuovo paradigma e, come abbiamo visto, esistono molte altre caratteristiche cruciali dei sistemi emergenti. Infatti, Freeman e Nunez hanno sostenuto che i sistemi più recenti – chiamati sistemi neo-cognitivisti – utilizzano calcolo parallelo, distribuito nella forma di connessioni neurali artificiali e memorie associative ma, tuttavia, aderiscono ancora alle

¹³ Questa connessione ricorrente non ha niente a che vedere con la risposta ai segnali di errore propri della *back propagation*, che è utilizzata per realizzare la regolazione dei pesi durante l'apprendimento.

¹⁴ La memoria etero-associativa – o semplicemente la memoria associativa – produce un vettore di uscita che è differente da quello di ingresso.

ipotesi originali del cognitivismo (Freeman & Nunez, 1999). Un'osservazione simile è stata effettuata da Van Gelder e Port (van Gelder & Port, 1995).

Uno dei fattori chiave dei sistemi emergenti in generale, e del connessionismo in particolare, consiste nel fatto che “la connettività del sistema diventa inseparabile *dalla propria storia delle trasformazioni* e in relazione al tipo di compito definito per il sistema” (F. J. Varela, 1992). Inoltre, i simboli non giocano nessun ruolo:¹⁵ mentre nell'approccio cognitivista i simboli sono distinti da ciò che rappresentano, nell'approccio connessionista “il significato si relaziona con lo stato globale del sistema” (F. J. Varela, 1992). Infatti il significato è qualcosa che viene attribuito da un osservatore esterno alla corrispondenza di uno stato del sistema con quello del mondo nel quale il sistema emergente è inserito. Se il significato è una descrizione attribuita da un agente esterno allora chiaramente non è qualcosa di intrinseco al sistema cognitivo, eccetto nel senso che le dinamiche del sistema riflettono l'efficacia della propria abilità di interazione con il mondo.

Esempi dell'applicazione dei sistemi di apprendimento associativi nella robotica possono essere trovati in (Jones & Vernon, 1994; Mel, 1988), dove il coordinamento mano-occhio viene appreso attraverso una connessione neurale di Kohonen associando stimoli propriocettivi ed estero-cettivi. Oltre a cercare di modellare il comportamento cognitivo, i sistemi connessionisti possono auto-organizzarsi per produrre capacità di analisi delle caratteristiche simili a quelle dei primi stadi di elaborazione del sistema visivo dei mammiferi (le cellule centro-bordo e quelle selettive all'orientazione) (Linsker, 1988). Un esempio di un sistema connessionista che utilizza l'interdipendenza tra percezione e azione in un ambito di sviluppo può essere trovato in (Metta, Sandini, & Konczak, 1999). Si tratta di un sistema motivato biologicamente che impara a raggiungere oggetti utilizzando la visione, adottando una segmentazione basata sul colore per identificare gli oggetti e una telecamera con sensore di tipo logaritmico-polare simile alla retina umana. Il sistema abbraccia l'approccio dello sviluppo: iniziando con sinergie motorie precostituite e passando attraverso una fase di apprendimento del coordinamento sensomotorio. Le reti con funzione base di tipo radiale sono state anche usate nei sistemi di visione cognitiva, per esempio per effettuare il riconoscimento dei volti (Buxton & Howell, 2002).

2.2.2 Modelli di sistemi dinamici

La teoria dei sistemi dinamici è stata utilizzata sia per complementare gli approcci classici nell'ambito dell'intelligenza artificiale (Reiter, 2001) che per modellare sistemi cognitivi naturali e artificiali (Kelso, 1999; Esther Thelen & Smith, 1998; van Gelder & Port, 1995). Gli scienziati che abbracciano l'approccio dei sistemi dinamici alla cognizione sostengono che i sistemi motorio e percettivo siano entrambi sistemi dinamici, ognuno dei quali si auto-organizza in modelli di comportamento meta-stabili.

¹⁵ Sarebbe più accurato dire che i simboli non dovrebbero giocare nessun ruolo, dato che è stato rilevato che i sistemi connessionisti spesso ritornano nel paradigma cognitivistico quando considerano i pesi neurali come una rappresentazione simbolica distribuita (van Gelder & Port, 1995).

In generale, un sistema dinamico è un sistema aperto dissipativo non-lineare e lontano dall'equilibrio: è un sistema nel senso che è formato da un gran numero di componenti che interagiscono tra loro, con un gran numero di gradi di libertà; è dissipativo nel senso che disperde energia (il suo spazio delle fasi decresce in volume nel tempo, implicando sotto-spazi preferenziali); non è all'equilibrio perchè non è in grado di mantenere la propria struttura o funzione senza una fonte esterna di energia, di materiale o d'informazione (e, quindi, aperto). L'aspetto di non-linearità è cruciale: oltre a generare un comportamento complesso, significa che la dispersione non è uniforme e che solamente un piccolo numero di gradi di libertà del sistema contribuisce al suo comportamento. Questi sono denominati *parametri d'ordine* (oppure *variabili collettive*). Ogni parametro d'ordine definisce l'evoluzione del sistema, portando gli stati meta-stabili in uno spazio di stati multi-stabili (o spazio delle fasi). Questa capacità di caratterizzare il comportamento di un sistema di elevata dimensionalità con un modello a bassa dimensionalità, è una delle caratteristiche che distingue i sistemi dinamici da quelli connessionisti (Kelso, 1999).

Alcune condizioni devono essere rispettate prima che un sistema si qualifichi come un sistema cognitivo dinamico. I componenti del sistema devono essere in relazione e interagire l'uno con l'altro; qualunque cambiamento in un componente o aspetto del sistema deve dipendere solamente dagli stati degli altri componenti: “devono essere interattivi e autocontenuti” (van Gelder & Port, 1995).

I proponenti dell'approccio dei sistemi dinamici fanno notare il fatto che questi forniscono direttamente molte delle caratteristiche inerenti ai sistemi cognitivi naturali, come la multi-stabilità, l'adattamento, la formazione di schemi e il riconoscimento di modelli, l'intenzionalità e l'apprendimento. Queste sono ottenute come funzione di leggi dinamiche e la conseguente auto-organizzazione; non richiedono alcun ricorso alle rappresentazioni simboliche, specialmente di quelle che sono il risultato dell'opera del progettista.

Tuttavia Clark (A. Clark, 2001) ha segnalato che l'antipatia che dimostrano i difensori degli approcci dei sistemi dinamici verso quelli cognitivisti si fonda su motivazioni piuttosto deboli, nella misura in cui gli scenari che utilizzano per sostenere il proprio punto di vista non richiedono livelli di ragionamento complessi: non sono “affamati di rappresentazione” e, quindi non sono molto adatti per essere utilizzati in un discorso generale anti-rappresentazionalista (o anti-cognitivista). Allo stesso tempo, Clark nota anche che tale antipatia non è neanche indirizzata alla rappresentazione *per se* (i sistemi dinamici prontamente ammettono stati interni che possono essere equivalenti a rappresentazioni), ma piuttosto al tipo di rappresentazioni oggettive che formano un surrogato simbolico isomorfo di una realtà esterna assoluta.

È stato sostenuto che i sistemi dinamici tengono conto dello sviluppo di funzioni cognitive di ordine più elevato, come l'intenzionalità e l'apprendimento, in modo diretto, almeno in linea di principio. Per esempio l'intenzionalità – comportamento premeditato – è ottenuta attraverso la sovrapposizione di una funzione potenziale¹⁶ intenzionale sulla funzione potenziale intrinseca (Kelso, 1999). Allo stesso modo, l'apprendimento è visto come la modifica di schemi di comportamento già esistenti che ha luogo in un contesto relativo alla storia del sistema nel quale la forma dell'intero attrattore (la configurazione

¹⁶ Dove il sistema dinamico ammetta una funzione potenziale.

dello spazio delle fasi) del sistema dinamico viene modificata. Pertanto l'apprendimento cambia l'intero sistema, in quanto consente lo sviluppo di nuovi attrattori.

Sebbene i modelli dinamici possano rappresentare diversi comportamenti non banali che richiedono l'integrazione di stimoli visivi e controllo motorio – includendo la percezione delle *affordances* e quella del tempo all'impatto, e alcuni effetti di bi-stabilità figura-sfondo (J. J. Gibson, 1950, 1979; Kelso, 1999; Koehler, 1940; Warren, 1984) – la possibilità di realizzare facoltà cognitive di livello elevato deve ancora essere dimostrata.

Le implicazioni dei modelli dinamici sono molte: “la cognizione è non simbolica, non rappresentazionale e l'intera attività mentale è emergente, situata, storica e incarnata” (Esther Thelen & Smith, 1998). È anche costruita socialmente, nel senso che alcuni livelli di cognizione emergono dall'interazione dinamica tra gli agenti cognitivi. Inoltre i sistemi cognitivi dinamici sono necessariamente incarnati. Tale requisito scaturisce direttamente dal fatto che la dinamica dipende da processi auto-organizzanti, nei quali il sistema differenzia se stesso come un'entità diversa attraverso la propria configurazione dinamica e la propria esplorazione interattiva dell'ambiente.

In riferimento ai sistemi emergenti in generale, e in particolare a quelli dinamici, uno dei fattori chiave consiste nel fatto che i processi cognitivi sono processi temporali che si ‘sviluppano’ in tempo reale e in modo sincronico, con gli eventi nel proprio ambiente. Tale requisito dello sviluppo sincronico nel contesto dell'ambiente circostante si collega all'approccio dei sistemi enattivi che verrà discusso nel prossimo paragrafo. È significativo per due motivi. In primo luogo, colloca un forte limite alla velocità con la quale l'apprendimento ontogenetico¹⁷ del sistema cognitivo può procedere: risulta limitato dalla velocità dell'accoppiamento (l'interazione), ma non dalla velocità secondo la quale possono avvenire i cambiamenti interni (Winograd & Flores, 1986). I sistemi cognitivi naturali possiedono un ciclo di apprendimento misurato in settimane, mesi e anni e, se è possibile per un sistema artificiale comprimerlo in tempi dell'ordine dei minuti e ore grazie ad aumenti della velocità dell'adattamento interno, esso non può essere ridotto al di sotto della scala dei tempi dell'interazione (o accoppiamento strutturale). Se il sistema deve sviluppare un'abilità cognitiva che, per esempio, gli permetta di anticipare o predire azioni ed eventi che hanno luogo durante una scala dei tempi estesa (ore), sarà necessario almeno quel periodo di tempo per l'apprendimento. In secondo luogo, considerato insieme al requisito dell'incarnazione, constatiamo che la natura di conseguenza storica e situata di questi sistemi significa che non è possibile “cortocircuitare” lo sviluppo ontogenetico. Più in particolare, non si può iniziare un sistema emergente dinamico in uno stato avanzato di comportamento o funzionalità.

Detto ciò, ci ricollegiamo all'introduzione, nel punto in cui si asseriva che una caratteristica importante dei sistemi cognitivi è la propria capacità di anticipare eventi futuri: la loro abilità di “liberarsi del presente”. Sembra esserci una contraddizione a questo punto. Da una parte abbiamo detto che i sistemi cognitivi emergenti sono sincroni agli eventi nell'ambiente e che il loro sviluppo deve procedere in tempo reale in modo sincronico con l'ambiente circostante, ma allo stesso tempo (asseriamo) che possono liberarsi da questo legame con il tempo reale. Di fatto, non esiste nessuna contraddizione: la sincronia è associata all'interazione del sistema con l'ambiente circostante, ma la capacità anticipatoria scaturisce dalla dinamica interna del sistema cognitivo. La capacità di auto-organizzazione e auto-sviluppo, coinvolgendo processi di copia e simulazione

¹⁷ L'ontogenetica ha a che vedere con lo sviluppo del sistema durante l'intera durata della propria esistenza.

degli eventi basati sull'esperienza precedente (generati nella storia del sistema dall'interazione sincronica), non è legata alla scala dei tempi dell'ambiente esterno in quanto opera internamente, auto-perturbandosi, libera dalle perturbazioni sincrone della percezione e dell'azione.

Sebbene gli approcci della teoria dei sistemi dinamici spesso differiscano dai sistemi connessionisti in diversi punti (Kelso, 1999; Esther Thelen & Smith, 1998; van Gelder & Port, 1995), forse sarebbe meglio considerarli come forme complementari di descrizione dei sistemi cognitivi, essendo i primi riferiti al comportamento macroscopico a un livello emergente, e i secondi riferiti a un comportamento microscopico a un livello meccanicistico (McClelland & Vallabha, 2006). Gli stessi sistemi connessionisti sono, in realtà, sistemi dinamici con proprietà temporali, e stati attrattori, instabilità e transizioni (Wilson, 1999). Tuttavia, i sistemi connessionisti descrivono la dinamica in uno spazio dei potenziali d'attivazione e delle connessioni a dimensionalità elevata, mentre i modelli della teoria dei sistemi dinamici descrivono la dinamica in uno spazio a bassa dimensionalità, nel quale un piccolo numero di variabili di stato cattura il comportamento del sistema come una singola unità. Schönner (Schoener, 2006) osserva che ciò è possibile perché lo stato macroscopico con dinamica a dimensionalità elevata e l'evoluzione dello stato a lungo termine, vengono catturati dalla dinamica in quel determinato spazio nel quale le instabilità avvengono: un sottospazio differenziale a bassa dimensionalità chiamato *Center-Manifold* (Schoener, 2006). Gran parte del potere della visione dei sistemi dinamici deriva da questa astrazione ad alto livello della dinamica (Smolensky, 1996c). La natura complementare dei sistemi dinamici e le descrizioni connessioniste viene enfatizzata da Schönner e Kelso (Kelso, 1999; Schoener, 2006), i quali asseriscono che i sistemi dinamici non-lineari dovrebbero essere modellati in modo simultaneo a tre diversi livelli: un livello delle condizioni al contorno che determina il compito o gli obiettivi (condizioni iniziali, condizioni non specificate); un livello di variabili collettive che caratterizzano gli stati coordinati e un livello di componenti che forma il sistema realizzato (per esempio, oscillatori non lineari accoppiati o reti neurali). Tutto ciò è significativo perché contrasta fortemente con l'approccio cognitivista, meglio riepilogato da quelli che adottano l'approccio di David Marr in una gerarchia di astrazione a tre livelli (teoria computazionale, rappresentazioni e algoritmi, e, infine, l'implementazione dell'hardware), essendo la modellazione al livello della teoria computazionale portata a termine senza riferimenti stringenti ai livelli inferiori (Marr, 1982). Questa prospettiva complementare della teoria dei sistemi dinamici e del connessionismo permette lo studio delle proprietà dinamiche emergenti dei sistemi connessionisti in termini di attrattori, meta-stabilità e transizione di stato; tutto ciò proviene dalla dinamica sottostante meccanicistica e offre la possibilità di implementare i modelli della teoria dei sistemi dinamici con architetture connessioniste.

2.2.3 Modelli di sistemi enattivi

I sistemi enattivi spingono il paradigma emergente ancora oltre. In contrapposizione al cognitivismo, che fa riferimento a un approccio della cognizione che richiede la rappresentazione di un determinato mondo oggettivo pre-determinato, i sistemi enattivi (H. Maturana, 1975; R. H. Maturana, 1980; R. H. Maturana & Varela, 1980; F. J. Varela, 1992; Winograd & Flores, 1986) asseriscono che la cognizione è un processo nel quale le

cose importanti per l'esistenza (la sopravvivenza) dell'entità cognitiva sono sviluppate in maniera efficace: esse sono co-determinate dall'entità mentre interagisce con l'ambiente circostante nel quale è immersa.

Tuttavia, niente è pre-determinato, quindi non c'è nessuna necessità di rappresentazioni simboliche. Al loro posto, c'è l'interpretazione enattiva: una scelta di importanza fondata sulla base del contesto eseguita in tempo reale.

Per il cognitivismo, il ruolo della cognizione consiste nell'astrarre la struttura oggettiva e il significato attraverso la percezione e il ragionamento. Per i sistemi enattivi, l'obiettivo della cognizione consiste nello scoprire regolarità e ordine in maniera non specifica, che possano in seguito essere utilizzati come significativi perchè facilitano l'operazione, l'esistenza e l'evoluzione del sistema cognitivo. L'approccio enattivo sfida l'ipotesi convenzionale che considera il mondo, così come lo sperimenta il sistema, indipendente dal sistema cognitivo ('il conoscitore'). Al contrario, colui che conosce e il conosciuto "si trovano in relazione l'uno con l'altro, con mutua specificazione: crescono insieme" (F. J. Varela, 1992).

L'unico requisito necessario per un sistema enattivo è l'azione efficace, cioè l'azione che permette l'integrità continuata nel tempo del sistema in questione. Si tratta essenzialmente di una posizione molto neutrale: ipotizza solamente che esista un'ordine nell'ambiente nel quale il sistema cognitivo è tenuto a operare e che, quindi, il sistema possa sfruttare l'ordine dell'ambiente per apprendere e svilupparsi. Da questo punto di vista, la *cognizione* è esattamente il processo attraverso il quale tale ordine o qualche aspetto del medesimo è scoperto (o costruito) dal sistema. Questo comporta in modo immediato l'esistenza di diverse forme della realtà (o d'importanza) che dipendono direttamente dalla natura della formazione dinamica del sistema cognitivo.

Non si tratta di una posizione solipsista di un soggettivismo infondato, ma nemmeno è la posizione comunemente adottata di un unico – rappresentabile – realismo. Siamo di fronte, fondamentalmente, a una posizione fenomenologica.

La ricerca nei sistemi enattivi risale ai primi anni 70, al lavoro dei biologi Maturana e Varela; venne poi ripresa da altri, inclusi alcuni autori del filone dell'approccio classico all'intelligenza artificiale (Winograd & Flores, 1986).

L'obiettivo della ricerca dei sistemi enattivi consiste nel trattamento completo della natura e dell'emergenza di sistemi autonomi, cognitivi e sociali. Si fonda sul concetto di autopoiesi – letteralmente auto-produzione –, secondo il quale un sistema emerge come un'entità coerente, separata dal proprio ambiente, in conseguenza di processi di auto-organizzazione. Tuttavia, l'enazione riguarda diversi livelli di autopoiesi, quindi si possono distinguere tre tipi di sistema.

I sistemi appartenenti al primo tipo di autopoiesi corrispondono a entità cellulari che ottengono un'identità fisica attraverso l'accoppiamento strutturale con il proprio ambiente circostante. Nel momento in cui il sistema è accoppiato con il proprio contesto, interagisce con esso nel senso che le perturbazioni scatenano cambiamenti strutturali 'che gli permettono di continuare a operare'. I sistemi del secondo ordine sono sistemi meta-cellulari che si legano in accoppiamenti strutturali con il proprio contesto, ma questa volta attraverso un sistema nervoso che rende possibile l'associazione di molti stati interni con le diverse interazioni nel quale l'organismo è coinvolto. Oltre ai processi di

auto-produzione, tali sistemi possiedono anche processi di auto-sviluppo. Maturana e Varela utilizzano il termine 'chiusura operativa' in riferimento ai sistemi del secondo ordine, al posto di autopoiesi, in modo da rispecchiare un livello superiore di flessibilità (R. H. Maturana & Varela, 1998). I sistemi del terzo ordine sono costituiti da diversi agenti cognitivi del secondo ordine legati da un ulteriore accoppiamento. È significativo il fatto che i sistemi del secondo e del terzo tipo posseggano l'abilità di perturbare i propri processi organizzativi e le strutture che li accompagnano. Gli accoppiamenti del terzo tipo permettono uno spostamento ontogenetico ricorrente (comune), nel quale i sistemi sono reciprocamente accoppiati. L'adattamento strutturale risultante – condiviso mutuamente dagli elementi del sistema – consente nuovi ambiti fenomenologici: il linguaggio e una epistemologia condivisa che riflette (ma non astrae) il mezzo comune nel quale avviene l'accoppiamento. Tali sistemi sono capaci di tre tipi di comportamento: (i) i comportamenti istintivi che derivano dai principi organizzativi che definiscono il sistema autopoietico in quanto tale (e che emergono dall'evoluzione filogenetica del sistema), (ii) i comportamenti ontogenetici che derivano dallo sviluppo del sistema durante la durata della vita, e (iii) i comportamenti comunicativi che sono il risultato dell'accoppiamento strutturale di terzo tipo tra i membri della società.

Il cuore dell'approccio enattivo consiste nel fatto che la cognizione è considerata un processo nel quale un sistema identifica delle regolarità come conseguenza della co-determinazione delle stesse attività cognitive, in modo che l'integrità del sistema venga preservata. Secondo tale approccio, il sistema nervoso (e un agente cognitivo) non astrae o 'raccolge informazioni' dall'ambiente circostante e, quindi, la metafora di considerare il cervello un dispositivo per il processamento delle informazioni è 'non solo ambigua, ma chiaramente sbagliata' (R. H. Maturana & Varela, 1998). Al contrario, la conoscenza è l'effettivo uso delle contingenze sensoriali e motorie fondate sull'accoppiamento strutturale nel quale il sistema nervoso esiste. La conoscenza è specifica della storia dell'interazione del sistema. La condivisione della conoscenza all'interno di una società di agenti cognitivi, non è dovuta a nessuna universalità intrinseca e astratta, ma alla storia consensuale di esperienze condivise tra gli agenti cognitivi con una filogenesi simile e uno sviluppo ontogenetico compatibile.

Come per i sistemi dinamici, quelli enattivi operano in modo sincronico in tempo reale: i processi cognitivi devono procedere in modo sincronico rispetto agli eventi nell'ambiente circostante, come conseguenza dell'accoppiamento strutturale e della co-determinazione tra il sistema e l'ambiente circostante. L'osservazione che abbiamo fatto sui sistemi dinamici riguardo al processo complementare di anticipazione, si applica anche in questo caso. Inoltre, i sistemi enattivi sono necessariamente sistemi incarnati come conseguenza diretta del requisito di accoppiamento strutturale. Non esiste una distanza semantica nei sistemi emergenti (connessionistici, dinamici o enattivi): il sistema costruisce la propria comprensione mentre si sviluppa e la comprensione cognitiva emerge dall'apprendimento mediante l'esplorazione.

Le posizioni emergenti in generale, e quella enattiva in particolare, sono supportate da risultati recenti che hanno dimostrato come la percezione del proprio corpo in un organismo biologico e la dimensionalità e geometria dello spazio nel quale è inserito,

possono essere dedotte (apprese o scoperte) dall'organismo attraverso un'analisi delle dipendenze esistenti tra i comandi motori e i conseguenti dati sensoriali, senza alcuna conoscenza o riferimento a un modello esterno del mondo oppure alla struttura fisica dell'organismo (Philipona, O'Regan, & Nadal, 2003; Philipona, O'Regan, Nadal, & Coenen, 2004). La struttura percepita della realtà potrebbe essere una conseguenza dello sforzo di una parte del cervello per tenere conto della dipendenza tra le proprie efferenze e afferenze nei termini di un numero limitato di parametri. Non esiste pertanto in realtà nessun bisogno di contare sull'idea classica di un modello a priori del mondo esterno che sia mappato dall'apparato sensoriale per 'qualche archetipo oggettivo'. Le concezioni di spazio, geometria e il mondo dal quale il corpo distingue se stesso, nascono dall'interazione sensomotoria del sistema; siamo esattamente di fronte alla posizione sostenuta dalla psicologia dello sviluppo (Esther Thelen & Smith, 1998). È proprio l'analisi delle conseguenze sensoriali dei comandi motori che dà luogo a tali concetti. In modo significativo, i comandi motori non sono riconducibili a una funzione dei dati sensoriali. La questione principale consiste nel fatto che le informazioni sensoriali e quelle motorie sono trattate in modo simultaneo, e non da una prospettiva di stimolo oppure da un punto di vista di controllo motorio.

L'approccio enattivo si trova nel lavoro di altri studiosi. Per esempio, Bickhard (Bickhard, 2000) introduce l'idea del sistema auto-persistente (*self-maintenant*) e dei sistemi ricorsivi auto-persistenti. Egli asserisce che 'le basi della cognizione sono costituite da autonomia adattiva, lontana dall'equilibrio – autonomia che è ricorsiva e auto-persistente –, non dall'elaborazione di simboli nè da quella di segnali di ingresso di tipo connessionista. Le basi della cognizione non sono simili a quelle del calcolatore elettronico nell'esecuzione di un programma, nemmeno ai vettori di attivazione di una passività connessionista'. Bickhard definisce l'autonomia come la proprietà di un sistema di contribuire alla propria persistenza. Dato che esistono diversi gradi di contributo, di conseguenza ci saranno diversi livelli di autonomia. Bickhard introduce una distinzione tra due tipi di sistema auto-organizzante.

- 1) *I sistemi auto-persistenti* contribuiscono attivamente alla propria persistenza, ma non contribuiscono al mantenimento delle condizioni necessarie alla stessa. Bickhard usa come esempio una candela accesa. La fiamma scioglie la cera, che a sua volta brucia per alimentare la fiamma.
- 2) *I sistemi ricorsivi auto-persistenti* contribuiscono attivamente alle condizioni per la persistenza. Tali sistemi possono disporre di diversi processi di auto-mantenimento, dipendendo dalle condizioni ambientali: "essi modificano i propri processi di auto-mantenimento in modo da tenere stabile l'auto-mantenimento mentre l'ambiente circostante cambia".

Egli distingue due tipi di stabilità: (a) stabilità in buca di potenziale, che equivale alla stabilità dei sistemi in equilibrio termodinamico – nessuna interazione con il proprio ambiente circostante è richiesta per mantenere l'equilibrio – e (b) stabilità lontana dall'equilibrio, equivalente all'equilibrio non termodinamico. La persistenza di questo stato di equilibrio richiede che il processo o il sistema non raggiunga l'equilibrio termodinamico. Tali sistemi sono completamente dipendenti per la loro esistenza da

contributi continui di fattori esterni: essi richiedono l'interazione con l'ambiente e sono necessariamente processi aperti (che, ciò nonostante, mostrano un'auto-organizzazione chiusa).

I sistemi auto-persistenti e quelli ricorsivi auto-persistenti sono entrambi esempi di sistemi con stabilità del tipo "lontana dall'equilibrio". Per quanto riguarda le rappresentazioni nei sistemi emergenti, egli nota che i sistemi ricorsivi, in effetti, conducono all'emergere della rappresentazione. La funzione nasce in sistemi auto-persistenti e la rappresentazione emerge come un tipo particolare di funzione ('indicazioni di potenziali interazioni') nei sistemi auto-persistenti che sono anche ricorsivi.

2.3 Modelli ibridi

Sforzi consistenti sono stati compiuti nello sviluppo di approcci che combinano aspetti dei sistemi emergenti e di quelli cognitivisti (G. H. Granlund, 1999; G.H. Granlund, 1999; Granlund, 2002). L'approccio ibrido ha una giustificazione nelle critiche sull'uso della conoscenza basata sulla programmazione esplicita nella creazione di sistemi di intelligenza artificiale (Dreyfus, 1982) e nello sviluppo di sistemi attivi 'animati' (Ballard, 1991), nei quali l'attenzione si concentra sui cicli di percezione-azione, piuttosto che sull'astrazione percettiva delle rappresentazioni.

I sistemi ibridi utilizzano ancora rappresentazioni e invarianze rappresentazionali, ma si è proposto di fare in modo che sia il sistema stesso a costruire le rappresentazioni, attraverso l'interazione e l'esplorazione del mondo, piuttosto che attraverso una determinazione *a priori* delle specifiche o attraverso la programmazione. In questo modo gli oggetti dovrebbero essere rappresentati come 'combinazioni invarianti di percezioni e risposte, nelle quali le invarianti (che non sono ristrette a proprietà geometriche) devono essere apprese attraverso l'interazione piuttosto che specificati o programmati *a priori*' (G. H. Granlund, 1999).

Quindi l'abilità di un sistema di interpretare gli oggetti e il mondo esterno dipende dalla capacità del medesimo di interagire in modo adattabile con questi, considerando che l'interazione è un meccanismo organizzativo che può guidare coerentemente la creazione di associazioni tra percezione e azione. Dal considerare un approccio che include l'azione e la percezione scaturiscono due importanti conseguenze. In primo luogo, non si può avere nessun accesso diretto significativo alle rappresentazioni semantiche interne; in secondo luogo, i sistemi cognitivi devono essere incarnati (almeno durante la fase dell'apprendimento) (G.H. Granlund, 1999). Secondo Granlund, per esempio, l'azione precede la percezione e "i sistemi cognitivi hanno bisogno di acquisire informazione sul mondo esterno attraverso l'apprendimento o il processo di creazione delle associazioni" [...] "In ultima analisi, l'elemento chiave consiste nell'ottenere plasticità comportamentale, cioè l'abilità di un sistema incarnato di imparare a portare a termine un compito per il quale non era stato esplicitamente progettato". Quindi, i sistemi ibridi sono per molti versi "compatibili" con quelli emergenti, anche se sfruttano ancora

rappresentazioni create dal programmatore (Okuma, Taleghani, de Freitas, Little, & Lowe, 2004).

Granlund descrive un sistema cognitivo visivo fondato su tali principi (Granlund & Moe, 2004). L'architettura di questo sistema combina un componente che include un mappaggio dalla percezione all'azione implementato tramite reti neurali (nel quale la percezione è mediata da azioni attraverso l'apprendimento e l'esplorazione) con un componente simbolico (fondato sui concetti – descrizioni invariante senza contesto spaziale – che possono essere utilizzati nell'elaborazione dove sia richiesta una capacità predittiva maggiore, come nel caso della pianificazione o della comunicazione).

Un sistema biologicamente motivato, modellato sulle funzioni del cervello e sulla connettività corticale e che utilizza il flusso ottico come stimolo visivo principale, ha mostrato come si possa sviluppare in maniera autonoma la segmentazione di oggetti, il riconoscimento e le capacità di localizzazione con conoscenze a priori limitate. Questo è stato ottenuto utilizzando la capacità di interagire in maniera attiva con l'ambiente attraverso l'esplorazione e con azioni di manipolazione relativamente semplici (Metta & Fitzpatrick, 2003). Questa estensione ibrida di un sistema connessionista precedente (Metta, Sandini, & Konczak, 1999) ha inoltre mostrato l'abilità di imparare le *affordances* di oggetti semplici e di utilizzarle per imitare le azioni di un altro agente (umano).

Un approccio ibrido alternativo, fondato sull'apprendimento e su metodi di riduzione attraverso sottospazi, viene utilizzato in (Jogan, Artac, Skocaj, & Leonardis, 2003) per costruire un sistema robotico incarnato che può imparare a orientarsi in un ambiente utilizzando una telecamera panoramica e un modello dell'ambiente circostante costruito in maniera incrementale in opportuni autospazi delle immagini.

2.4 Punti di forza relativi

Ognuno degli approcci presentati possiede punti di forza, così come i propri sostenitori e critici, e si trovano a diversi stadi di maturità scientifica. Gli argomenti a favore dei sistemi dinamici e di quelli enattivi sono importanti ma, al momento, le capacità dei sistemi cognitivisti sono in realtà quelle più avanzate. Tuttavia, i sistemi cognitivisti presentano gli ormai noti problemi di fragilità.

Vari autori hanno fornito critiche dettagliate sui diversi approcci. Tra essi, ricordiamo per esempio Clark (A. Clark, 2001), Christensen e Hooker (Christensen & Hooker, 2000a) e Crutchfield (Crutchfield, 1998). Christiansen e Hooker sostengono che i sistemi cognitivisti risentono di tre problemi: quello detto del *symbol grounding*, il problema del *frame* (la necessità di differenziare il significante in un insieme di dati molto ampio e poi di generalizzare per includere nuovi dati)¹⁸; e infine, il problema dell'esplosione combinatoriale.

¹⁸ Nel paradigma cognitivistico, il problema del *frame* è stato espresso in termini leggermente diversi, ma essenzialmente equivalenti: come si può creare un programma capace di inferire gli effetti di un'azione senza ragionare in modo esplicito sui suoi possibili molti non-effetti (Shanahan & Baars, 2005).

I suddetti problemi rappresentano uno dei motivi per cui i modelli cognitivisti hanno difficoltà nel creare sistemi che mostrino interazioni sensoriali-motorie in ambienti reali, con presenza di rumore e con caratteristiche dinamiche complesse. Presentano inoltre difficoltà nel modellare le abilità cognitive di livello più alto, come la generalizzazione, la creatività e l'apprendimento (Christensen & Hooker, 2000b). Secondo Christensen e Hooker, e come abbiamo sottolineato in molte occasioni, i sistemi cognitivisti sono poveri nel loro funzionamento, al di fuori di un ambito ristretto e di un problema ben definito.

I sistemi enattivi e dinamici dovrebbero in teoria essere molto meno fragili, poiché la cognizione in questo caso emerge attraverso l'interazione reciproca e il co-sviluppo con l'ambiente, ma la nostra abilità nel costruire sistemi artificiali cognitivi sulla base di tali principi attualmente è molto limitata. Fino a oggi la teoria dei sistemi dinamici ha rappresentato più un ambito di riferimento generale, piuttosto che un vero e proprio modello di cognizione (Christensen & Hooker, 2000b); è stata inoltre utilizzata come uno strumento di analisi, piuttosto che come strumento per la progettazione e la sintesi dei sistemi cognitivi (Metta, Vernon, & Sandini, 2005). Rimane un argomento di ricerca sviluppare ulteriormente questi metodi.

Alcuni credono che gli approcci ibridi offrano il meglio dei due mondi: l'adattabilità dei sistemi emergenti (poiché essi costruiscono le rappresentazioni attraverso l'apprendimento e l'esperienza), ma anche le caratteristiche avanzate dei sistemi cognitivisti (poiché le invarianze e le strutture rappresentazionali non necessitano di essere apprese ma sono progettate all'interno di essi). Non è chiaro quanto sia possibile combinare quelle che sono in definitiva filosofie altamente antagoniste (A. Clark, 2001).

L'implementazione di un sistema cognitivo sarà inevitabilmente un sistema complesso e mostrerà sicuramente qualche forma di organizzazione, anche se non si tratterà probabilmente dell'organizzazione suggerita dagli approcci cognitivisti. La teoria dei sistemi dinamici non offre attualmente molto aiuto all'identificazione di tale organizzazione, dato che il modello consiste di una dinamica stato-spazio, che viene in realtà astratta dall'organizzazione fisica del sistema sottostante (Christensen & Hooker, 2000b). L'organizzazione richiesta non deve necessariamente seguire la decomposizione funzionale dall'alto verso il basso dell'intelligenza artificiale, ma ciò nonostante potrebbe utilizzare una qualche forma appropriata di organizzazione funzionale.

I sistemi dinamici, al momento, offrono più una struttura generale di un modello plausibile piuttosto che un modello di cognizione; tale affermazione è generalmente condivisa e diversi autori hanno sottolineato lo stesso concetto, cioè che l'approccio dei sistemi dinamici è stato utilizzato di più come uno strumento d'analisi che come uno strumento per la progettazione e la sintesi dei sistemi cognitivi (Christensen & Hooker, 2000b; Metta, Vernon, & Sandini, 2005).

Clark suggerisce che un passo in avanti è costituito dallo sviluppo di una forma di "computazionalismo dinamico", nel quale gli elementi dinamici formano parte di un sistema di elaborazione dell'informazione (A. Clark, 2001). Tale idea viene ripresa da

Crutchfield (Crutchfield, 1998) che, mentre è d'accordo sul fatto che la dinamica sia indubbiamente coinvolta nella cognizione, sostiene che la dinamica di per sé "non è un sostituto dell'elaborazione dell'informazione e della computazione nei processi cognitivi". Egli non dice neanche che i due approcci non siano nemmeno incompatibili. Dichiarò che è possibile svilupparne una sintesi per fornirne un approccio che permetta alle strutture dinamiche dello spazio degli stati di sostenere la computazione. Propone il termine 'meccanica computazionale' per riferirsi a tale sintesi di dinamica e computazione. Tuttavia, tale sviluppo richiede che la stessa dinamica sia estesa da quella deterministica, a bassa dimensionalità e asintotica nel tempo, verso una che sia stocastica, distribuita e ad alta dimensionalità, che reagisca nei transitori piuttosto che in maniera asintotica. Inoltre, il vincolo che identifica la computazione con la computazione di tipo digitale o discreto deve essere rilassato per permettere altre interpretazioni del significato della computazione.

3 Le architetture cognitive

Sebbene usato liberamente dai proponenti degli approcci cognitivisti, emergenti e ibridi, il termine architettura cognitiva si originò in realtà dal lavoro di Newell e colleghi (Newell, 1982, 1990; Rosenbloom, Laird, & Newell, 1993). Il termine possiede un significato ben preciso nel paradigma cognitivista, nel quale le architetture cognitive rappresentano tentativi di creare teorie unificate della cognizione (Newell, 1990), cioè teorie che possano coprire i vari aspetti caratteristici della cognizione, come l'attenzione, la memoria, la risoluzione di problemi, il processo di decisione, l'apprendimento, partendo da diversi aspetti come la psicologia, la neuroscienza e le scienze della computazione. L'architettura Soar di Newell (Rosenbloom, Laird, & Newell, 1993), l'architettura ACT-R di Anderson (J. R. Anderson, 1996) e la *Società della Mente* di Minsky (Minsky, 1986) sono tutte candidate a essere teorie unificate della cognizione.

Per gli approcci emergenti, quando si consideri anche lo sviluppo da uno stato iniziale primitivo a uno pienamente cognitivo durante la vita del sistema, l'architettura del sistema è equivalente alla configurazione filogenetica: lo stato iniziale dal quale successivamente si sviluppa l'agente cognitivo essendo gli stati successivi al di fuori del controllo del progettista.

Nel paradigma cognitivista, il punto focale dell'architettura cognitiva risiede negli aspetti della cognizione che rimangono costanti nel tempo e che sono relativamente indipendenti da un compito particolare (Gray, Young, & Kirschenbaum, 1997). Considerando che le architetture cognitive rappresentano la parte costante della cognizione, non possono veramente fare niente di per sé e hanno bisogno di ricevere conoscenza per portare a termine compiti precisi.

Tale combinazione tra una determinata architettura cognitiva e un particolare insieme di conoscenze è generalmente denominato *modello cognitivo*. Da notare che nella maggior parte dei sistemi cognitivi la conoscenza incorporata nel modello è determinata dal progettista (umano), sebbene si registri un crescente uso di tecniche di apprendimento automatico per aumentare e adattare tale conoscenza. Le specifiche di un'architettura

cognitiva consistono nelle ipotesi rappresentazionali, nelle caratteristiche della struttura della memoria e nei processi che operano su questa memoria. L'architettura cognitiva definisce il modo in cui un agente cognitivo gestisce le risorse a sua disposizione. Per gli approcci cognitivisti, tali risorse sono nient'altro che il sistema computazionale nel quale il sistema simbolico è realizzato. L'architettura specifica il formalismo relativo alle rappresentazioni della conoscenza e alla memoria utilizzata per archivarle, i processi che agiscono su tale conoscenza e, infine, i meccanismi di apprendimento di questa. Normalmente fornisce altresì un modo di programmazione del sistema, in modo che sistemi dotati di "intelligenza" possano essere istanziati in un qualche dominio applicativo (Langley, 2005).

Per gli approcci emergenti il bisogno di identificare un'architettura nasce dall'intrinseca complessità della cognizione e dalla necessità di fornire qualche forma di struttura nella quale inserire i meccanismi di percezione, azione, adattamento, anticipazione e motivazione che attivano lo sviluppo ontogenetico. È proprio il livello di complessità l'elemento di distinzione di un'architettura cognitiva emergente che considera elementi di sviluppo ontogenetico da un semplice sistema connessionista di controllo di un robot che tipicamente apprende solo associazioni per compiti particolari (Jones & Vernon, 1994). In un certo senso, l'architettura cognitiva di un sistema emergente corrisponde alle capacità innate che sono date dalla filogenesi e che non devono essere imparate, ma che possono essere senza dubbio sviluppate ulteriormente. Tali risorse facilitano l'ontogenesi del sistema. Rappresentano il punto di partenza per il sistema cognitivo e forniscono la base e il meccanismo per il successivo sviluppo autonomo, uno sviluppo che può condizionare direttamente l'architettura stessa. Come abbiamo già asserito, l'autonomia del sistema in questo tipo di sviluppo è importante, perché pone forti restrizioni nel modo in cui la conoscenza del sistema viene acquisita e in quello in cui la propria semantica viene fondata (tipicamente, da un apprendimento di possibilità di azione che preservino l'autonomia, che siano anticipatorie e possibilmente adattive) e attraverso la quale si ottiene un'epistemologia condivisa da una popolazione di agenti (il risultato soggettivo di una storia di esperienze consensuali condivise tra agenti filogeneticamente compatibili).

È importante enfatizzare che la presenza di capacità innate nei sistemi emergenti *non* implica assolutamente che l'architettura sia funzionalmente modulare: ovvero che il sistema cognitivo sia formato da diversi moduli, ognuno dei quali porta a termine un compito determinato. Se una modularità è presente, potrebbe essere dovuta al fatto che il sistema sviluppi tale modularità attraverso l'esperienza, come parte dell'ontogenesi o epigenesi, piuttosto che questa sia predetta dalla filogenia del sistema (teoria della "ri-descrizione rappresentazionale" Karmiloff-Smith, 1992, 1994). Fatto ancora più importante, la modularità non implica necessariamente che le capacità innate siano delle competenze cognitive solidificate, come suggerito dalla psicologia nativista (Fodor, 1983; Pinker, 1997)¹⁹. Allo stesso tempo, non implica nemmeno necessariamente che il sistema cognitivo sia una *tabula rasa*, priva di qualunque struttura cognitiva innata, come

¹⁹ Più recentemente, Fodor (Fodor, 2000) asserisce che la modularità si applica solamente alla cognizione locale (per esempio nel riconoscere una foto del Monte Whitney), ma non alla cognizione globale (decidere di scalare il John Muir Trail).

postulato dalla visione costruttivista dello sviluppo cognitivo di Piaget (Piaget, 1955)²⁰; alla fine deve esistere un meccanismo, una struttura o un'organizzazione che permetta al sistema cognitivo di essere autonomo, di agire in modo efficace ancorché limitato, e di sviluppare tale autonomia. Il paradigma emergente si colloca in opposizione con i due pilastri del cognitivismo – il dualismo che postula la separazione logica di mente e corpo, così come il funzionalismo che asserisce che i meccanismi cognitivi sono indipendenti dalla piattaforma fisica (Freeman & Nunez, 1999) – è probabile che l'architettura rispecchierà o riconoscerà in qualche modo la morfologia del corpo fisico nel quale è inserita, e del quale è parte intrinseca.

Avendo stabilito le condizioni al contorno per le architetture cognitive e per quelle emergenti, (e implicitamente, per quelle ibride), per gli obiettivi di questo articolo il termine architettura cognitiva sarà utilizzato nel senso generale e non in quello specifico. Con questo intendiamo la configurazione minima necessaria a un sistema per mostrare capacità e comportamenti cognitivi: le specifiche dei componenti in un sistema cognitivo, la loro funzione e la loro organizzazione come insieme. Detto questo, poniamo maggiore enfasi sulle necessità dei sistemi emergenti che adottano almeno in parte l'approccio dello sviluppo piuttosto che su quelli preconfigurati.

Cognitivista	Emergente	Ibrida
Soar EPIC ACT-R ICARUS ADAPT	AAR Global Workspace I-C SDAL SASE DARWIN	HUMANOID Cerebus Cog: Teoria della Mente Kismet

Tabella II: le architetture cognitive esaminate.

Nel prossimo paragrafo descriviamo un confronto tra varie architetture cognitive realizzate seguendo l'approccio cognitivista, quello emergente e infine quello ibrido. La tabella II elenca le architetture cognitive esaminate per ognuna delle tre categorie.

4 Confronto tra le architetture

La tabella III mostra un riassunto di tutte le architetture esaminate *vis-à-vis* un sottoinsieme delle dodici caratteristiche dei sistemi cognitivi delle quali abbiamo parlato nella precedentemente nella tabella I. Abbiamo omesso le prime cinque – operazioni elementari, struttura rappresentazionale, legame semantico con l'ambiente, restrizioni temporali, epistemologia – poiché queste possono essere inferite direttamente dal paradigma adottato nella progettazione del sistema cognitivo considerato: cognitivista, emergente o ibrido, indicati con le lettere C, E o H nella tabella III. Una 'x' indica che l'aspetto o il problema corrispondente è rappresentato o affrontato in maniera consistente nell'implementazione dell'architettura, '+' indica che è considerato ma non rappresenta

²⁰ Piaget fondò la scuola costruttivista dello sviluppo cognitivo, secondo la quale la conoscenza non è fissata a priori (filogeneticamente), ma viene scoperta e costruita dal bambino, attraverso la manipolazione attiva dell'ambiente.

una caratteristica chiave; uno spazio vuoto indica che non è considerato e che sostanzialmente l'architettura non risolve il problema. Una 'x' è stata collocata in corrispondenza della colonna 'Adattamento' solamente se il sistema è in grado di svilupparsi (nel senso di riuscire a creare nuove strutture rappresentative o modelli), piuttosto che apprendere semplicemente (nel senso del stima dei parametri del modello) (Weng, 2004).

Architettura	Paradigma	Incarnazione	Percezione	Azione	Anticipazione	Adattamento	Motivazioni	Autonomia
Soar	C				+	+		
Epic	C		+	+	+			
ACT-R	C		+	+	+	+		
ICARUS	C		+	+	+	+		
ADAPT	C	x	x	x	+	+		
AAR	E	x	x	x			+	x
Spazio di lavoro globale	E	+	+	+	x		x	x
I-C SDAL	E	+	+	+	+	+	x	x
SASE	E	x	x	x	+	x	x	x
Darwin	E	x	x	+		x	x	x
HUMANOID	H	x	x	x	x	+	+	
Cerebrus	H	x	x	x	+	+		
Cog: Teoria della Mente	H	x	x	x	+			
Kismet	H	x	x	x			x	

Tabella III: le architetture cognitive vis-à-vis sette delle dodici caratteristiche che abbiamo usato per caratterizzare i sistemi cognitivi.

5 Il paradigma dello sviluppo

5.1 Sviluppo

Lo sviluppo implica la progressiva acquisizione, durante la vita di un sistema cognitivo, di capacità anticipatorie e predittive attraverso l'apprendimento basato sull'esperienza. Come abbiamo avuto modo di constatare, lo sviluppo richiede una base dalla quale crescere – una configurazione filogenetica – oltre a motivazioni che possano guidarlo.

Nel paradigma emergente la configurazione filogenetica deve agevolare l'autonomia del sistema e, in modo particolare, l'accoppiamento del sistema con il proprio ambiente circostante, attraverso la percezione e l'azione; e promuovere l'auto-organizzazione del

sistema come un'entità distinta dall'ambiente. Questo accoppiamento complementare tra percezione e azione e le forme di auto-organizzazione si definisce co-determinazione. La co-determinazione ha origine dalla natura autonoma di un sistema cognitivo e rispecchia il fatto che un sistema autonomo si autodefinisce attraverso un processo di auto-organizzazione e sottomette tutti gli altri processi a quello della salvaguardia di tale autonomia (F. Varela, 1979). Qualunque sistema auto-organizzato è immerso in un ambiente, dal quale si distingue, e che è interpretato dal sistema in qualunque modo esso supporti questo processo di salvaguardia e preservazione dell'autonomia. In tale senso, il sistema e il suo ambiente sono co-specificati: l'agente cognitivo viene determinato dal suo ambiente, dalle sue necessità di sostenere la propria autonomia di fronte alle perturbazioni ambientali, ma allo stesso tempo il processo cognitivo determina ciò che è reale o significativo per l'agente. La co-determinazione significa che l'agente costruisce la propria realtà (il proprio mondo) come risultato delle proprie operazioni in quel mondo.

Maturana e Varela introducono una notazione grafica per esprimere la natura autonoma auto-organizzata di un sistema co-determinato che disturba o viene disturbato dal proprio ambiente (H. Maturana, 1975): a questo proposito si veda la figura 2. Il cerchio con la freccia indica l'autonomia e l'auto-organizzazione del sistema, la linea ondulata l'ambiente, e le frecce bi-direzionali la perturbazione reciproca.

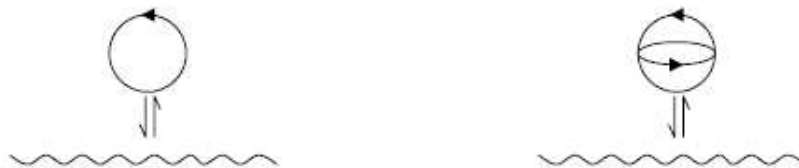


Figura 2: gli ideogrammi di Maturana e Varela come notazione per indicare sistemi autopoietici con chiusura operativa. Tali sistemi mostrano rispettivamente la co-determinazione e l'auto-sviluppo.

Il diagramma sulla sinistra rappresenta un sistema autopoietico: il cerchio con la freccia indica l'autonomia, l'auto-organizzazione, e l'auto-produzione del sistema; la linea ondulata l'ambiente, e le frecce bi-direzionali la perturbazione reciproca – l'accoppiamento strutturale – tra i due. Il diagramma sulla destra rappresenta invece un sistema autonomo con chiusura operativa, con un sistema nervoso centrale. Questo sistema è in grado di svilupparsi attraverso l'auto-perturbazione – auto-modificazione – del proprio sistema nervoso, in modo da accomodare uno spazio molto maggiore di azioni efficaci.

La co-determinazione richiede quindi che il sistema sia capace di essere autonomo come entità. In altre parole, che possieda un processo di auto-organizzazione capace di azione e percezione coerenti, e gli elementi essenziali per sopravvivere e svilupparsi. Questo è esattamente quello che intendiamo con la configurazione filogenetica di un sistema: le capacità innate di un sistema disponibili all'inizio della sua esistenza. Questo, pertanto, costituisce la base per il successivo sviluppo. Un sistema co-determinato autonomo

possiede una ristretta gamma di capacità comportamentali e quindi un limitato grado di autonomia.

Lo sviluppo autonomo è identico al processo cognitivo che stabilisce e aumenta lo spazio degli accoppiamenti possibili nei quali un sistema può inserirsi o sopravvivere mentre mantiene (o incrementa) la sua autonomia. Ciò nonostante, è lo sviluppo del sistema nel tempo in un contesto ecologico e sociale mentre avviene la crescita delle possibilità di accoppiamento, che deve essere consistente con il mantenimento dell'auto-organizzazione. Lo sviluppo autonomo richiede una plasticità addizionale e processi di auto-organizzazione. Lo spazio delle possibilità percettive non è istanziato su un ambiente assoluto e oggettivo ma nello spazio delle possibili azioni che il sistema può intraprendere mentre conserva la consistenza dell'accoppiamento con l'ambiente. Queste perturbazioni non controllano il sistema dato che non sono parti di esso (e per definizione non giocano nessun ruolo nell'auto-organizzazione) ma altresì giocano un ruolo importante nel suo sviluppo. Il sistema cognitivo costruisce la sua epistemologia personale proprio per mezzo dello sviluppo ontogenetico. Il sistema sviluppa una sua storia privata e una conoscenza specifica del suo mondo che ha senso esattamente perchè cattura la consistenza e regolarità che emerge dalla regolazione dell'accoppiamento con l'ambiente. In altre parole, le azioni del sistema determinano la percezione con il vincolo che si preservi l'accoppiamento con l'ambiente. Alla fine, tutto si riduce alla preservazione dell'autonomia, ma questa volta in uno spazio di possibili accoppiamenti che cresce di dimensione col passare del tempo.

Questo processo di sviluppo si ottiene attraverso l'auto-modificazione grazie all'esistenza del sistema nervoso centrale: non solo l'ambiente perturba il sistema (e vice versa) ma è anche il sistema stesso che genera delle perturbazioni e in ultima analisi l'adattamento del sistema nervoso centrale. Come conseguenza, il sistema cognitivo si sviluppa in modo da accomodare un numero enormemente maggiore di azioni efficaci. Tutto ciò è rappresentato graficamente dal secondo ideogramma di Maturana e Varela (figura 2) che aggiunge un secondo cerchio con una freccia al primo diagramma dell'autopoiesi per descrivere appunto il processo di auto-modificazione.

Lo sviluppo autonomo e la co-determinazione presi insieme corrispondono alla visione di Thelen che afferma appunto che la percezione, l'azione e la cognizione formano un unico processo di auto-organizzazione "nel contesto specifico delle perturbazioni ambientali del sistema" (E. Thelen, 1995). Quindi, possiamo dire, che dal punto di vista di Thelen la cognizione è inseparabile dall'azione fisica (E. Thelen, 1995): senza una specifica incarnazione che determini l'esplorazione, un sistema cognitivo non avrebbe la base dello sviluppo. I sistemi emergenti, per definizione, devono essere incarnati e inseriti nel loro ambiente, e situati in un contesto di sviluppo (Esther Thelen & Smith, 1998).

È importante enfatizzare il fatto che lo sviluppo avviene in un modo molto speciale. L'azione, la percezione, e la cognizione sono accoppiati durante lo sviluppo: non solo l'azione organizza la percezione e la cognizione ma la percezione e cognizione sono a loro volta necessarie per organizzare l'azione. I sistemi di controllo non appaiono già formati e non siamo neppure nella situazione dove questi siano appresi completamente

attraverso l'esperienza. Essi in realtà sono il risultato sia del funzionamento del sistema nervoso centrale e dell'interazione dinamica del sistema con l'ambiente. Percezione, cognizione e motivazioni sono il risultato dello sviluppo che avviene all'interfaccia dei processi cerebrali e dell'azione. Come conseguenza, la cognizione può anche essere vista come il risultato di un processo di sviluppo attraverso il quale il sistema diventa progressivamente più abile e acquisisce la capacità di comprendere gli eventi che lo circondano, di capire diversi contesti e azioni, al principio considerando solo situazioni immediate e via via aumentando le proprie capacità di predizione. Questa dipendenza dello sviluppo dall'esplorazione è una delle ragioni che portano alcuni a sostenere la tesi che i sistemi incarnati necessitano di uno spazio delle azioni ricco di possibilità quali la manipolazione e la locomozione (Sandini, Metta, & Vernon, 2004).

Notiamo che il concetto di co-determinazione è certamente basato sul lavoro di Maturana e Varela e sull'idea dell'accoppiamento strutturale dei sistemi autopoietici di livello uno²¹ (R. H. Maturana & Varela, 1998). Questo è simile anche al concetto di Kelso della causalità circolare tra azione e percezione, ognuna delle quali è una funzione dell'altra nella gestione dell'interazione col mondo eseguita dal sistema cognitivo (Kelso, 1999). È possibile notare che lo stesso concetto è anche simile a quello dei sistemi auto-persistenti di Bickhard (Bickhard, 2000). In questo caso il parallelo è anche più profondo: il concetto di auto-persistenza ricorsivo ha le sue radici nel livello due e tre dei sistemi autopoietici di Maturana e Varela (R. H. Maturana & Varela, 1998).

Per concludere, lo sviluppo dell'azione e della percezione, lo sviluppo del sistema nervoso, e lo sviluppo (come crescita) del corpo, sono processi che hanno una mutua influenza nel processo globale di crescita della complessità, delle capacità di predizione necessarie per la soluzione del problema di imparare a generare un comportamento efficace (von Hofsten, 2004).

5.2 Apprendimento e motivazioni

Lo sviluppo dipende in maniera cruciale dalle motivazioni che definiscono gli scopi dell'agire. Le due motivazioni fondamentali sono quella sociale e quella esplorativa. La motivazione sociale include il conforto, la sicurezza e la soddisfazione. Ci sono almeno due motivazioni legate all'esplorazione: una relativa alla scoperta della novità e regolarità del mondo e un'altra che riguarda la scoperta del potenziale delle proprie azioni.

La motivazione che porta a espandere il proprio repertorio di azioni è molto forte fino al punto a volte di deviare temporaneamente il sistema dallo scopo prefissato (per esempio, lo sviluppo della capacità di camminare oppure l'apprendimento della coordinazione occhio-testa in situazioni per le quali un movimento esclusivamente oculare sarebbe più efficace. Nello stesso modo, scoprire le proprietà degli oggetti o degli eventi nel contesto di una nuova azione è una motivazione molto forte.

²¹ L'autopoiesi è un tipo speciale di auto-organizzazione: un sistema autopoietico è un sistema omeostatico (che si auto regola) ma di un tipo particolare nel quale la regolazione non avviene per un parametro del sistema ma rispetto all'organizzazione del sistema stesso (H. Maturana, 1975).

La psicologia dello sviluppo supporta questo punto di vista secondo il quale l'esplorazione è un elemento cruciale dello sviluppo ontogenetico. Von Hofsten sostiene che il fatto di raggiungere uno scopo specifico non sembra motivare un neonato, bensì la scoperta di nuovi modi di interazione: la scoperta di un nuovo modo di fare qualcosa attraverso l'esplorazione (Von Hofsten, 2003). Chiaramente, l'atto di esplorare non può far altro che sospendere, almeno momentaneamente, l'azione corrente. Come conseguenza, lo sviluppo ontogenetico, differisce dal mero apprendimento per il fatto che (a) deve inibire delle capacità esistenti e (b) deve poter calcolare (forse effettuare) dei cambiamenti nella morfologia o struttura del sistema (Sandini, Metta, & Konczak, 1997). L'inibizione non è necessariamente una perdita di controllo ma semplicemente il blocco del collegamento tra certi stimoli sensoriali e le corrispondenti risposte motorie.

Oltre allo sviluppo di abilità specifiche attraverso l'esplorazione (avvicinamento, prensione e manipolazione di quello che c'è attorno), la cognizione si sviluppa in almeno altre due modalità importanti. Queste sono l'imitazione (Billard, 2002; Rao, Shon, & Meltzoff, 2004) e l'interazione sociale, che include l'insegnamento esplicito (Demiris & Hayes, 2002).

Al contrario di altri metodi (per esempio l'apprendimento con rinforzo) l'imitazione – l'apprendimento di nuovi comportamenti attraverso l'osservazione delle azioni di altri – consente un'acquisizione rapida (Rao, Shon, & Meltzoff, 2004). Diversi autori (Meltzoff, 2002; Meltzoff & Moore, 1997) suggeriscono che i bambini imparano imitando gli altri attraverso quattro fasi:

1. lallazione²², che si manifesta come una serie di movimenti di tentativo apparentemente casuali;
2. l'imitazione dei movimenti del corpo;
3. l'imitazione delle azioni verso gli oggetti;
4. l'imitazione che si basa sulla possibilità di capire che cosa fanno gli altri.

I neonati usano la lallazione (detta a volte, lallazione motoria per distinguerla da quella tradizionalmente riferita al parlato) per imparare lo spazio delle azioni nel quale nuove posture possono essere costruite per interpolazione anche se è sorprendente che già alla nascita i neonati possano imitare alcuni movimenti (Rao, Shon, & Meltzoff, 2004). Lo sviluppo dell'imitazione segue la stessa sequenza di altre capacità di comunicazione, come per esempio, la condivisione dell'attenzione, l'alternanza durante la comunicazione e il linguaggio (Nadel, Guerini, Peze, & Rivet, 1999; Speidel, 1989; Trevarthen, Kokkinaki, & Fiamenghi, 1999). L'imitazione è chiaramente un elemento cruciale verso lo sviluppo di capacità cognitive più elevate.

È importante capire che cosa voglia dire esattamente il termine "interazione". L'interazione è un'attività condivisa nella quale due o più agenti influenzano le proprie azioni a vicenda col risultato di generare una qualche forma di programma di comportamento condiviso (Ogden, Dautenhahn, & Stribling, 2002). Questa definizione è coerente con l'approccio emergente della cognizione già ampiamente discusso,

²² In senso non strettamente linguistico ma più in generale motorio.

specialmente, in questo contesto, per quanto riguarda la natura emergente dell'interazione ispirata ai concetti di autopoiesi e accoppiamento strutturale (R. H. Maturana & Varela, 1980).

L'aspetto relativo ai modelli di comportamento complementare costruiti in maniera reciproca viene anche enfatizzato nella nozione proposta da Clark dell'azione congiunta (H. H. Clark, 1994). Secondo tale definizione, il significato esplicito non è necessario per comunicare qualcosa durante un'interazione, ma è invece importante che gli agenti siano reciprocamente coinvolti in una sequenza di azioni. Il significato emerge attraverso l'esperienza condivisa e consensuale, mediata dall'interazione. Sviluppo e motivazioni a parte, i meccanismi per effettuare una modifica della struttura in modo autonomo – l'apprendimento – sono ancora necessari.

Si possono distinguere tre tipi di apprendimento: apprendimento supervisionato, nel quale i segnali d'errore hanno caratteristiche direzionali; apprendimento tramite rinforzo, nel quale i segnali d'errore sono semplicemente un segnale che indica una qualità ed è uno scalare; e, infine, apprendimento non supervisionato, senza alcuna presenza di segnali d'errore. Doya sostiene che il cervelletto sia specializzato nell'apprendimento supervisionato, i gangli della base per l'apprendimento tramite rinforzo e la corteccia cerebrale per l'apprendimento non supervisionato (Doya, 1999). Suggestisce che, nello sviluppo delle architetture cognitive, i moduli dell'apprendimento supervisionato del cervelletto possano essere utilizzati come un modello interno dell'ambiente e come un modello-bypass di altri mappaggi di ingresso-uscita che sono stati acquisiti in un'altra parte del cervello. I moduli di apprendimento tramite rinforzo nei gangli della base vengono usati per valutare un determinato stato e, di conseguenza, per selezionare un'azione. I moduli dell'apprendimento non supervisionato nella corteccia cerebrale rappresentano lo stato dell'ambiente esterno, oltre al contesto interno, fornendo anche una struttura rappresentazionale comune per il cervelletto e i gangli della base, che non possiedono connessioni anatomiche dirette tra di loro.

A prescindere dai dettagli esatti del modello di Doya, ciò che risulta significativo è che diverse regioni (del cervello) facilitino diversi tipi di apprendimento, e che tali regioni e i relativi processi di apprendimento siano interdipendenti. Per esempio, McClelland *et al.* hanno suggerito che l'ippocampo e la neo-corteccia formino un sistema complementare per l'apprendimento (McClelland, NcNaughton, & O'Reilly, 1995). L'ippocampo favorisce l'apprendimento rapido autoassociativo, ma anche etero-associativo, che viene utilizzato per riprodurre e consolidare memorie apprese nella neo-corteccia in modo graduale. In tal modo, la memoria dell'ippocampo può essere considerata non solamente come un magazzino di memorie, ma come “un segnale d'errore del sistema di elaborazione neocorticale”. Da notare inoltre che la riproduzione delle memorie può avvenire in linea, quindi rendendo possibile il controllo esplicito del comportamento, ma anche fuori linea, per esempio, la simulazione mentale, i ricordi e il sonno.

In modo simile, Rougier ha proposto e convalidato un'architettura per una memoria auto-associativa fondata sull'organizzazione dell'ippocampo, coinvolgendo la corteccia entorinale, il giro dentato, CA3 e CA1 (Rougier, 2001). Una caratteristica saliente di tale architettura consiste nel fatto che evita il problema della cosiddetta interferenza catastrofica, tipica delle memorie associative, attraverso l'uso di ridondanza, ortogonalizzazione, e rappresentazioni con codifica sparsa. Anche Rougier osserva che

l'ippocampo gioca un ruolo 'nell'insegnare' qualcosa alla neo-corteccia, cioè nella formazione delle rappresentazioni neocorticali.

Diversi tipi di sviluppo richiedono diversi meccanismi di apprendimento. I comportamenti innati sono affinati tramite l'apprendimento continuo, senza basarsi su una conoscenza esplicita e probabilmente con meccanismi di apprendimento tramite rinforzo, in un processo in qualche modo simile alla stima parametrica. D'altra parte, nuove competenze si sviluppano attraverso un diverso modo di apprendimento, guidato non solamente dalle funzioni di rinforzo tradizionali di ricompensa/punizione (risposta positiva e negativa), ma attraverso un gioco e una esplorazione spontanea non supervisionata (Sloman & Chappell, 2005a, 2005b).

Riassumendo, le competenze cognitive emergono progressivamente attraverso lo sviluppo ontogenetico, mentre si impara a dare un senso del proprio mondo tramite l'esplorazione, la manipolazione, l'imitazione e l'interazione sociale, includendo la comunicazione (Sandini, Metta, & Vernon, 2004). I fautori dell'approccio enattivo aggiungerebbero il requisito che tale sviluppo abbia luogo nel contesto di una causalità circolare di azione e percezione, essendo ognuna una funzione dell'altra, mentre il sistema governa la propria interazione con il mondo: essenzialmente lo sviluppo autonomo di azione e percezione, e co-determinazione del sistema attraverso l'auto-organizzazione in un contesto ecologico e sociale. Per concludere, Winograd e Flores (Winograd & Flores, 1986) catturano l'essenza dell'apprendimento e dello sviluppo emergente dicendo:

“L'apprendimento non è un processo di accumulo di rappresentazioni dell'ambiente; è invece un processo continuo di trasformazione del comportamento attraverso il cambiamento continuo nella capacità del sistema nervoso di sintetizzarlo. I ricordi non dipendono in una conservazione infinita di un invariante strutturale che rappresenta l'entità da ricordare (un'idea, immagine o simbolo), ma nell'abilità funzionale del sistema di creare, quando certe condizioni ricorrenti sono date, un comportamento che l'osservatore classificherebbe come la riproduzione di uno stesso comportamento precedente”.

5.3 La co-dipendenza percezione-azione

È stato mostrato che la percezione e l'azione nei sistemi biologici sono co-dipendenti. Per esempio, l'attenzione spaziale dipende dal programma oculo-motorio: quanto l'occhio è in una posizione vicina a un limite di movimento, e perciò non può eseguire una saccade in quella direzione, l'attenzione visiva nella stessa direzione è attenuata (Craigheo, Nascimben, & Fadiga, 2004). Questa teoria premotoria dell'attenzione si applica anche all'attenzione selettiva nella quale alcuni oggetti piuttosto che altri sono più visibili. Per esempio, l'abilità di rilevare un oggetto migliora quando le sue caratteristiche visive coincidono con la configurazione di prensione che il soggetto sta preparando (Craigheo, Fadiga, Rizzolatti, & Umilta', 1999). In altre parole, l'azione del soggetto condiziona la sua percezione. In maniera simile, l'esistenza di un insieme particolare di neuroni – detti neuroni specchio – è spesso citata come evidenza di una relazione tra percezione e azione (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). I neuroni specchio sono attivi sia quando un'azione è eseguita, sia quando questa o

una simile è eseguita da un altro agente. Questi neuroni sono specifici per lo scopo finale dell'azione ma non della meccanica della sua esecuzione (von Hofsten, 2004). Inoltre, lo sviluppo percettivo è determinato dalle capacità motorie durante lo sviluppo e da quali oggetti, azioni e eventi sono osservati nel contesto delle varie azioni (E. J. Gibson & Pick, 2000; von Hofsten, 2004).

Un esempio pratico di un sistema che sfrutta questa dipendenza in uno scenario che include lo sviluppo è descritto in (Metta, Sandini, & Konczak, 1999), nel quale si presenta un sistema biologicamente plausibile che impara a raggiungere alcuni oggetti colorati utilizzando immagini di tipo logaritmico-polare. Il sistema adotta l'approccio dello sviluppo iniziando con alcuni riflessi innati e precostituiti e imparando successivamente la coordinazione sensomotoria. Questo sistema opera ipotizzando che il punto di fissazione rappresenti l'oggetto da raggiungere con la mano, quindi i comandi motori sono costruiti mappando i dati della propriocezione del sistema occhio-testa con i parametri di controllo del braccio. Il controllo stesso è implementato come delle sinergie multi-giunto che modulano una combinazione lineare di campi di forza. Il campo di forza totale descrive quali forze o coppie applicare a un attuatore o a un certo insieme di essi per ottenere un particolare punto di equilibrio nello spazio di lavoro. In pratica, il comando motorio che dirige lo sguardo sull'oggetto è anche utilizzato per determinare il comando motorio per raggiungerlo, implementando quello che nell'articolo viene chiamato coordinazione motoria-motoria. La conversione tra gli angoli ai giunti che descrivono la direzione dello sguardo e quelli che descrivono la configurazione del braccio è appresa osservando la mano del robot durante una fase di ontogenesi del sistema.

Un sistema simile al precedente ma con diverse estensioni, modellato seguendo le funzioni del cervello e in particolare alcuni percorsi corticali utilizza il flusso ottico come stimolo visivo primario e mostra come si possano sviluppare autonomamente capacità quali la segmentazione visiva degli oggetti, il riconoscimento e la localizzazione senza conoscenza a priori della forma ma solo attraverso azioni esplorative e di manipolazione (Metta & Fitzpatrick, 2003). Il sistema in questione dimostra inoltre come si possano imparare le *affordance* degli oggetti (anche se semplici) e quindi utilizzarne la conoscenza per imitare l'azione di un essere umano. L'ipotesi è che l'azione sia richiesta in tutte quelle situazioni dove il sistema deve procedere attraverso lo sviluppo di classi o categorie di oggetti in maniera autonoma. L'inerente ambiguità della percezione visiva può essere risolta agendo nell'ambiente e attraverso l'esperienza. Lo sviluppo inizia attraverso l'esplorazione, la prensione e infine con il riconoscimento degli oggetti. L'addestramento che riguarda il controllo motorio e i movimenti del braccio è eseguito come in (Metta, Sandini, & Konczak, 1999) ma in questo caso invece di utilizzare dei semplici identificatori colorati, la segmentazione è fatta utilizzando l'informazione relativa al flusso ottico e analizzandone la coerenza temporale con il movimento del braccio stesso (durante l'apprendimento misurando la sincronità delle discontinuità nel movimento del braccio con le discontinuità del flusso ottico). La segmentazione degli oggetti è anch'essa effettuata tramite il flusso ottico toccando e spingendo gentilmente gli oggetti e misurando le aree dell'immagine il cui flusso ottico è correlato con il movimento della mano ma che non può essere attribuito alla mano stessa. Gli oggetti che

sono segmentati in questo modo, dal tocco, possono poi essere classificati utilizzando la tecnica degli istogrammi nello spazio del colore. Una semplice *affordance* – il rotolamento dopo il contatto – è quindi appresa calcolando la probabilità di movimento relativa all'asse principale dell'oggetto, quest'ultimo ottenuto processando le aree dell'immagine segmentate con la procedura precedente. L'effetto di differenti modi di spingere gli oggetti (gesti) è imparato per ogni possibile caso calcolando la probabilità (un istogramma) della direzione di movimento mediata su tutti i possibili oggetti. Negli esperimenti sono riportati quattro possibili gesti chiamati di avvicinamento, allontanamento, rovescio e laterale. Quando il sistema opera in modalità non esplorativa, il riconoscimento degli oggetti è basato sulla ricerca del colore, la localizzazione tramite la retroproiezione dell'istogramma, e l'orientazione dell'oggetto è calcolata stimando l'asse principale tramite il confronto con un insieme di prototipi visivi dello stesso oggetto (anch'essi acquisiti durante l'apprendimento). Il robot infine seleziona l'azione cercando la direzione preferita di rotolamento (partendo dall'*affordance*), aggiungendo l'orientazione corrente dell'oggetto e, quindi, scegliendola in modo da massimizzare la probabilità associata con la direzione corrente. La riproduzione del comportamento altrui (distinta dall'imitazione che richiede l'apprendimento di comportamenti nuovi mentre in questo caso si ripetono azioni note (Billard, 2002)) si effettua presentando un oggetto conosciuto al robot e mostrandogli una particolare azione. Questa azione che dovrà essere imitata è marcata da un'attività notevole nei pressi dell'oggetto (che è comunque osservato dal robot) che ha come conseguenza l'inibizione dell'azione di prensione da parte del robot. Durante questa dimostrazione gli effetti dell'azione sono misurati dal robot utilizzando anche in questo caso il flusso ottico (l'oggetto è segmentato utilizzando esattamente la stessa procedura già descritta). Quando l'oggetto è presentato al robot per la seconda volta, l'azione che riproduce al meglio quella appena osservata è selezionata. Si pensa che questo sia esattamente quello che ci si dovrebbe aspettare da una rappresentazione percezione-azione simile a quella dei neuroni specchio. I neuroni specchio possono essere pensati come una mappa di “associazioni tra l'osservazione di un'azione di manipolazione eseguita da qualcun altro e la rappresentazione neurale della propria azione”.

6 Conclusioni

Concludiamo questo articolo cercando di collegare tra loro i vari temi presentati e riassumendo alcune delle caratteristiche fondamentali che dovrebbe mostrare un sistema cognitivo artificiale per essere definito tale. In particolare, cerchiamo di mettere in evidenza quali sono i principi necessari affinché un sistema si possa dire di “implementare” un approccio basato sullo sviluppo. È forse necessario però sottolineare la dicotomia tra cognitivismo e sistemi emergenti. Come abbiamo visto ci sono alcune differenze fondamentali nell'adottare uno di questi due paradigmi – l'incarnazione è solo un principio teorico nei sistemi simbolici mentre è fondamentale che i sistemi emergenti posseggano un corpo fisico (Vernon, 2006a) e, soprattutto, il modo in cui i sistemi cognitivisti possano evitare lo sviluppo ontogenetico spesso attraverso l'inserimento di conoscenza esterna, specifica per il dominio in esame, o con la realizzazione di strutture *ad-hoc* a marcare la differenza. Differenze che comunque sembrano oggi essere minori di un tempo. Questo perché (i) una parte dei proponenti del cognitivismo riconoscono ora il

ruolo fondamentale dell'azione e della percezione per la realizzazione di sistemi cognitivi; (ii) perchè ci si è allontanati dalle forme simboliche della rappresentazione come l'unica valida alternativa (A. Clark, 2001); e (iii) per l'indebolirsi dell'idea di inserire della conoscenza a priori e, invece, per l'aumentare dell'utilizzo di tecniche di apprendimento automatico nella regolazione di alcuni parametri, nell'acquisizione di nuova conoscenza, per la rappresentazione di oggetti e la formazione stessa di nuove rappresentazioni. Nonostante ciò, i sistemi cognitivisti non hanno ancora coperto la distanza che li separa dallo sviluppo ontogenetico propriamente detto con tutte le sue conseguenze di autonomia, necessità di un corpo, plasticità dell'architettura, e la costruzione di conoscenza mediata dall'esplorazione, dalle motivazioni sociali e da sistemi di valori innati o appresi.

Krichmar e colleghi identificano sei principi di progetto per sistemi che sono capaci di sviluppo (Krichmar & Edelman, 2005). Anche se presentano questi principi nel contesto dei sistemi basati su modelli plausibili del cervello, essi sono in larga misura rilevanti anche per i sistemi emergenti in generale.

Primo, Krichmar suggerisce che l'architettura deve considerare la dinamica degli elementi neurali in differenti regioni del cervello, la struttura di queste regioni, e in particolare la connettività e l'interazione tra queste regioni.

Secondo, essi ci fanno notare come un sistema cognitivo dovrebbe essere in grado di effettuare una categorizzazione percettiva, ovvero organizzare i segnali sensoriali, in qualunque modalità essi siano in opportune categorie senza che ci sia una conoscenza a priori o istruzioni esterne. A tutti gli effetti questo vuol dire che il sistema è autonomo, e come dice Weng (Weng, 2004, p. 206), un sistema dotato di sviluppo dovrebbe essere un generatore di modelli piuttosto che adattare i parametri di modelli esistenti (Olsson, Nehaniv, & Polani, 2006).

Terzo, un sistema che si sviluppa deve essere istanziato fisicamente, deve essere incarnato, in modo da essere legato sia alla sua morfologia sia all'esplorazione dell'ambiente circostante.

Quarto, il sistema deve poter muoversi nell'ambiente e quindi deve possedere un insieme minimale di comportamenti innati che gli consentano di sopravvivere nella sua nicchia ecologica iniziale. Partendo da questo insieme iniziale, chiaramente, il sistema può apprendere e adattarsi in modo che col tempo il suo comportamento migliori²³.

Quinto, i sistemi che si sviluppano devono avere dei meccanismi per effettuare l'adattamento. Questo implica necessariamente la presenza di un sistema di valori (motivazioni che guidano e governano lo sviluppo). Questi ultimi dovrebbero essere segnali modulatori non specifici²⁴ che cambiano la dinamica in modo tale da soddisfare i bisogni globali del sistema cognitivo: in effetti, in modo che l'autonomia si preservi e venga possibilmente aumentata. Il sistema di valori può, per esempio, essere modellato sulla falsariga dei corrispondenti sistemi nervosi: i sistemi di segnali dopaminergico, colinergico e noradreninergico, sulla base degli stimoli sensoriali, sulla predizione dell'errore, l'incertezza e la novità. Krichmar e colleghi notano anche che questi sistemi dovrebbero sempre essere confrontati con i sistemi biologici corrispondenti.

²³ Krichmar dice "ottimizzare" invece di "migliorare".

²⁴ Nel senso che non specificano l'azione da intraprendere.

A questo punto possiamo chiederci quali considerazioni si possano trarre da questa enunciazione di principi quando si volesse passare all'implementazione in sistemi reali.

In primo luogo, un sistema cognitivo con possibilità di sviluppo dovrà essere costituito da una rete di sotto-sistemi (o circuiti corticali) distribuiti e multi-funzionali, in competizione ma anche cooperanti allo stesso tempo; ognuno dovrà essere dotato della propria rappresentazione limitata, ma in grado, insieme agli altri sotto-sistemi, di raggiungere l'obiettivo di un comportamento efficace. Tale rete costituisce la configurazione filogenetica del sistema: le sue abilità innate.

In secondo luogo, un'architettura cognitiva che includa lo sviluppo deve essere capace di portare a termine una qualche forma di auto-modificazione, non semplicemente adattando i parametri del sistema (come nel caso dell'uso di tecniche di apprendimento automatico per effettuare la stima di un certo insieme di parametri), ma facilitando la modifica della struttura stessa e dell'organizzazione del sistema, in modo che questo sia capace di cambiare la propria dinamica interna, ampliando il proprio repertorio di azioni, e quindi adattandosi alle nuove circostanze. Tale sviluppo dovrebbe essere guidato da motivazioni sia esplorative che sociali. La prima motivazione è legata alla necessità di scoprire le regolarità nel mondo ma anche il potenziale delle azioni proprie del sistema, mentre la seconda motivazione ha a che vedere con le interazioni tra agenti, le attività condivise e le strategie di comportamento comune possibilmente realizzate grazie a una epistemologia condivisa. Per fare ciò, è probabilmente necessario far ricorso a una varietà di paradigmi di apprendimento, includendo apprendimento non supervisionato, quello supervisionato e con rinforzo.

Poichè i sistemi cognitivi non sono solamente adattivi ma anche *anticipativi*, diventa cruciale che abbiano (in virtù della loro filogenesi) o sviluppino (in virtù della loro ontogenesi) alcuni meccanismi di simulazione di scenari ipotetici – in modo esplicito, come l'architettura ACT-R di Anderson (Anderson, 1996), oppure in modo implicito, come l'architettura dello spazio di lavoro globale di Shanahan (Shanahan, 2006) – e utilizzino questi per modulare il comportamento reale del sistema.

I sistemi cognitivi, per poter effettuare lo sviluppo, devono possedere un corpo, almeno nel senso dell'accoppiamento strutturale con l'ambiente e probabilmente in qualche forma organismoide più forte (Ziemke, 2001, 2003). Se poi la comprensione epistemologica dei sistemi artificiali deve essere consistente con quella di altri agenti cognitivi, come gli esseri umani, questo è ancora più vero (Vernon, 2006b). Sembra evidente come la complessità e il livello di complessità del comportamento cognitivo dipendano dalla ricchezza e diversità dell'interazione con l'ambiente e quindi dalla ricchezza potenziale delle azioni del sistema.

Infine, sia per il paradigma cognitivista sia per quello emergente, lo sviluppo (l'ontogenesi) dipende dalla configurazione filogenetica del sistema, così come dalla sua storia di interazioni e attività. Quale configurazione filogenetica sia esattamente necessaria e sufficiente per i sistemi cognitivi artificiali rimane una questione aperta.

Rimane da dire che qualunque studio approfondito della cognizione dovrebbe considerare il ruolo (o la rilevanza) della coscienza, così come i fondamenti filosofici, psicologici e neuroscientifici della cognizione.

Ringraziamenti

Questo lavoro è supportato dal progetto europeo RobotCub IST-004370 dell'Obiettivo Strategico 2.3.2.4. Cognitive Systems. Gli autori vogliono ringraziare Vincenzo Tagliasco per i commenti sulle versioni precedenti dell'articolo nonché i revisori per i numerosi suggerimenti. Gli autori ringraziano Sabrina Celli per la traduzione di alcune parti del testo da una precedente versione inglese.

Riferimenti bibliografici

- Anderson, J. A., & Rosenfeld, E. (Eds.). (1988). *Neurocomputing: Foundations of Research*. Cambridge, MA: MIT Press.
- Anderson, J. A., & Rosenfeld, E. (Eds.). (1991). *Neurocomputing 2: Directions for Research*. Cambridge, MA: MIT Press.
- Anderson, J. R. (1996). Act: A simple theory of complex cognition. *American Psychologist*, *51*, 335-365.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review*, *111*(4), 1036-1060.
- Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, *149*(1), 99-130.
- Arbib, M. A. (Ed.). (2002). *The Handbook of Brain Theory and Neural Networks* (Second Edition). Cambridge, MA: MIT Press.
- Arens, M., & Nagel, H.-H. (2002). *Representation of behavioral knowledge for planning and plan recognition in a cognitive vision system*. Paper presented at the 25th German Conference on Artificial Intelligence, Aachen, Germany.
- Arens, M., & Nagel, H.-H. (2005). Quantitative movement prediction based on qualitative knowledge about behaviour. *KI-Zeitschrift Kunstliche Intelligenz, Special Issue on Cognitive Computer Vision*, 5-11.
- Arens, M., Ottlick, A., & Nagel, H.-H. (2002). *Natural language texts for a cognitive vision system*. Paper presented at the 15th European Conference on Artificial Intelligence (ECAI 2002), Amsterdam, Holland.
- Ballard, D. H. (1991). Animate vision. *Artificial Intelligence*, *48*, 57-86.
- Berthoz, A. (2000). *The Brain's Sense of Movement*. Cambridge, MA: Harvard University Press.
- Bickhard, M. H. (2000). Autonomy, function, and representation. *Artificial Intelligence, Special Issue on Communication and Cognition*, *17*(3-4), 111-131.
- Billard, A. (2002). Imitation. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 566-569). Cambridge, MA: MIT Press.
- Brachman, R. J. (2002). Systems that know what they're doing. *IEEE Intelligent Systems*, *17*(6), 67-71.
- Buxton, H. (2002). *Generative Models for Learning and Understanding Dynamic Scene Activity*. Paper presented at the ECCV Workshop on Generative Model Based Vision, Copenhagen, Denmark.
- Buxton, H., & Howell, J. (2002). *Active Vision Techniques for Visually Mediated Interaction*. Paper presented at the International Conference on Pattern Recognition, Quebec City, Canada.
- Buxton, H., Howell, J., & Sage, K. (2002). *The Role of Task Control and Context in Learning to Recognize Gesture*. Paper presented at the Workshop on Cognitive Vision, Zurich, Switzerland.
- Carpenter, G. A., & Grossberg, S. (1995). Adaptive resonance theory (ART). In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 79-82). Cambridge, MA: MIT Press.

- Chella, A., Frixione, M., & Gaglio, S. (1997). A cognitive architecture for artificial vision. *Artificial Intelligence*, 89(1-2), 73-111.
- Christensen, W. D., & Hooker, C. A. (2000a). An interactivist-constructivist approach to intelligence: self-directed anticipative learning. *Philosophical Psychology*, 12(1), 5-45.
- Christensen, W. D., & Hooker, C. A. (2000b). *Representation and the meaning of life*. Paper presented at the Representation in Mind: New Approaches to Mental Representation. 27-29th June 2000 at The University of Sydney, Australia.
- Clark, A. (2001). *Mindware: an introduction to the philosophy of cognitive science*. Oxford: Oxford University Press, UK.
- Clark, H. H. (1994). Managing problems in speaking. *Speech Communication*, 15, 243-250.
- Cohn, A. G., Hogg, D. C., Bennett, B., Devin, V., Galata, A., Megee, D. R., et al. (2005). Cognitive vision: Integrating symbolic qualitative representations with computer vision. In H. I. Christensen & H.-H. Nagel (Eds.), *Cognitive Vision Systems: Sampling the Spectrum of Approaches* (pp. 211-234). Heidelberg: Springer-Verlag.
- Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. A. (1999). Movement for perception: a motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance*.
- Craighero, L., Nascimben, M., & Fadiga, L. (2004). Eye Position Affects Orienting of Visuospatial Attention. *Current Biology*, 14, 331-333.
- Crowley, J. L. (2005). Things that see: Context-aware multi-modal interaction. *KI-Zeitschrift Kunstliche Intelligenz, Special Issue on Cognitive Computer Vision*.
- Crutchfield, J. P. (1998). Dynamical embodiment of computation in cognitive processes. *Behavioural and Brain Sciences*, 21(5), 635-637.
- Demiris, J., & Hayes, G. (2002). Imitation as a Dual-Route Process Featuring Predictive and Learning Components: A Biologically-Plausible Computational Model. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts* (pp. 327-361). Cambridge, MA, USA: MIT Press.
- Dickmanns, E. D. (2004). Dynamic vision-based intelligence. *AI Magazine*, 25(2), 10-29.
- Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks*, 12, 961-974.
- Dreyfus, H. L. (1982). From micro-worlds to knowledge representation. In J. Haugland (Ed.), *Mind Design: Philosophy, Psychology, Artificial Intelligence* (Excerpted from the Introduction to the second edition of the author's *What Computers Can't Do*, Harper and Row, 1979. ed., 161-204). Cambridge, MA: MIT Press.
- Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Feldman, R. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205-254.
- Fodor, J. A. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.
- Fodor, J. A. (2000). *The Mind Doesn't Work that Way*. Cambridge, MA: MIT Press.
- Freeman, W. J., & Nunez, R. (1999). Restoring to cognition the forgotten primacy of action, intention and emotion. *Journal of Consciousness Studies*, 6(11-12), ix-xix.

- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593-609.
- Gardner, H. (1993). *Multiple Intelligences: The Theory in Practice*. New York: Basic Books.
- Gibson, E. J., & Pick, A. (2000). *An Ecological Approach to Perceptual Learning and Development*. Oxford: Oxford University Press.
- Gibson, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton Mifflin.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Granlund, G. H. (1999). The complexity of vision. *Signal Processing*, *74*, 101-126.
- Granlund, G. H. (1999). *Does vision inevitably have to be active?* Paper presented at the SCIA99, Scandinavian Conference on Image Analysis. June 7-11, 1999. Kangerlussuaq, Greenland.
- Granlund, G. H. (2002). *Cognitive vision – background and research issues*. Unpublished manuscript, Linköping.
- Granlund, G. H., & Moe, A. (2004). Unrestricted recognition of 3D objects for robotics using multilevel triplet invariants. *AI Magazine*, *25*(2), 51-67.
- Gray, W. D., Young, R. M., & Kirschenbaum, S. S. (1997). Introduction to this special issue on cognitive architectures and human-computer interaction. *Human-Computer Interaction*, *12*, 301-309.
- Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*, 121-134.
- Haugland, J. (1982). Semantic engines: An introduction to mind design. In J. Haugland (Ed.), *Mind design: Philosophy, Psychology, Artificial Intelligence*, 1-34. Cambridge, MA: MIT Press.
- Hebb, D. O. (1949). *The Organization of Behaviour*. New York: John Wiley & Sons.
- Hinton, G. E., & Anderson, J. A. (Eds.). (1981). *Parallel models of associative memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hinton, G. E., & Sejnowski, T. J. (1986). Learning and relearning in boltzmann machines. In D. E. Rumelhart, J. L. McClelland & T. P. R. Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 282-317. Cambridge, MA: MIT Press.
- Hollnagel, E., & Woods, D. D. (1999). Cognitive systems engineering: New wind in new bottles. *International Journal of Human-Computer Studies*, *51*, 339-356.
- Hopfield, J. J. (1982). Neural neural network and physical systems with emergent collective computational abilities. *Proceedings of National Academy of Sciences*, *79*(8), 2554-2588.
- James, W. (1890). *The Principles of Psychology* (Vol. 1).
- Jogan, M., Artac, M., Skocaj, D., & Leonardis, A. (2003). A framework for robust and incremental self-localization of a mobile robot. In J. Crowley, J. Piater, M. Vincze & L. Paletta (Eds.), *Proceedings of the Third International Conference on Computer Vision Systems, ICVS 2003*, 2626, 460-469. Berlin Heidelberg: Springer-Verlag.
- Jones, M., & Vernon, D. (1994). Using neural networks to learn hand-eye co-ordination. *Neural Computing and Applications*, *2*(1), 2-12.

- Jordan, M. I. (1986). *Attractor dynamics and parallelism in a connectionist sequential machine*. Paper presented at the Eighth Conference of the Cognitive Science Society. Amherst, MA, USA, 531-546. Hillsdale: Erlbaum.
- Karmiloff-Smith, A. (1992). *Beyond Modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Karmiloff-Smith, A. (1994). Precis of beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences*, 17(4), 693-745.
- Kelso, J. A. (1999). *Dynamic Patterns: the self organization of brain and behavior*. Cambridge, MA: MIT Press.
- Koehler, W. (1940). *Dynamics in Psychology*. New York: Liveright.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.
- Krichmar, J. L., & Edelman, G. M. (2005). Brain-based devices for the study of nervous systems and the development of intelligent machines. *Artificial Life*, 11, 63-77.
- Krichmar, J. L., & Edelman, G. M. (2006). *Principles underlying the construction of brain-based devices*. Paper presented at the AISB 06 Adaptation in Artificial and Biological Systems, Bristol.
- Langley, P. (2005). *An adaptive architecture for physical agents*. Paper presented at the IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Copiegne, France.
- Linsker, R. (1988). Self-organization in a perceptual network. *Computer*, 21(3), 105-117.
- Maillot, N., Thonnat, M., & Boucher, A. (2003). Towards ontology based cognitive vision. In J. Crowley, J. Piater, M. Vincze & L. Paletta (Eds.), *Proceedings of the Third International Conference on Computer Vision Systems (ICVS 2003)*, 2626, 44-53. Berlin: Springer-Verlag.
- Marr, D. (1977). Artificial Intelligence - A personal view. *Artificial Intelligence*, 9, 37-48.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- Maturana, H. (1975). The organization of the living: a theory of the living organization. *International Journal of Man-Machine Studies*, 7(3), 313-332.
- Maturana, R. H. (1980). Biology of Cognition. In *Autopoiesis and Cognition: The Realization of the Living*, 5-58, Dordrecht: D. Reidel Publishing Co.
- Maturana, R. H., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: D.Reidel Publishing Co.
- Maturana, R. H., & Varela, F. J. (1998). *The tree of knowledge, the biological roots of human understanding* (R. Paolucci, Trans., Revised Edition ed.). Boston & London: Shambhala Publications, Inc.
- McClelland, J. L. (1981). *Retrieving general and specific information from stored knowledge of specifics*. Paper presented at the Third Annual Meeting of the Cognitive Science Society, Berkeley, CA, USA.
- McClelland, J. L., NcNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419-457.
- McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature*, 4, 310-322.

- McClelland, J. L., & Vallabha, G. (2006). Connectionist models of development: Mechanistic dynamical models with emergent dynamical properties. In J. P. Spencer, M. S. C. Thomas & J. L. McClelland (Eds.), *Toward a New Grand Theory of Development? Connectionism and Dynamic Systems Theory Re-Considered*, New York: Oxford University Press.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- Medler, D. A. (1998). A brief history of connectionism. *Neural Computing Surveys*, 1, 61-101.
- Mel, B. W. (1988). *MURPHY: A robot that learns by doing*. Paper presented at the Neural Information Processing Systems, New York, 1988, 544-553.
- Meltzoff, A. N. (2002). The elements of a developmental theory of imitation. In A. N. Meltzoff & W. Prinz (Eds.), *The Imitative Mind: Development, Evolution, and Brain Bases*, 19-41. Cambridge: Cambridge University Press.
- Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting*, 6, 179-192.
- Metta, G., & Fitzpatrick, P. (2003). Early Integration of Vision and Manipulation. *Adaptive Behavior*, 11(2), 109-128.
- Metta, G., Sandini, G., & Konczak, J. (1999). A Developmental Approach to Visually-Guided Reaching in Artificial Systems. *Neural Networks*, 12(10), 1413-1427.
- Metta, G., Vernon, D., & Sandini, G. (2005). *The robotcub approach to the development of cognition: Implications of emergent systems for a common research agenda in epigenetic robotics*. Paper presented at the Fifth International Workshop on Epigenetic Robotics (EpiRob2005), Nara, Japan.
- Minsky, M. (1986). *The Society of Mind*. New York: Simon & Schuster.
- Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.
- Moeller, R., Neumann, B., & Wessel, M. (1999). *Towards computer vision with description logics: Some recent progress*. Paper presented at the Integration of Speech and Image Understanding, Corfu, Greece.
- Moody, J., & Darken, C. J. (1989). Fast learning in networks of locally tuned processing units. *Neural Computation*, 1, 281-294.
- Nadel, J., Guerini, C., Peze, A., & Rivet, C. (1999). The evolving nature of imitation as a format for communication. In J. Nadel & G. Butterworth (Eds.), *Imitation in Infancy*, 209-234. Cambridge: Cambridge University Press.
- Nagel, H.-H. (2004). Steps toward a cognitive vision system. *AI Magazine*, 25(2), 31-50.
- Neumann, B., & Moeller, R. (2005). On scene interpretation with description logics. In H. I. Christensen & H.-H. Nagel (Eds.), *Cognitive Vision Systems: Sampling the Spectrum of Approaches* (In press, 235-260). Heidelberg: Springer-Verlag.
- Newell, A. (1982). The knowledge level. *Artificial Intelligence*, 18(1), 87-127.
- Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.
- Newell, A., & Simon, H. A. (1975). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery*, 19, 113-126.

- Ogden, B., Dautenhahn, K., & Stribling, P. (2002). Interactional structure applied to the identification and generation of visual interactive behaviour: Robots that (usually) follow the rules. In I. Wachsmuth & T. Sowa (Eds.), *Gesture and Sign Languages in Human-Computer Interaction*, 2298, 254-268: Springer-Verlag.
- Okuma, K., Taleghani, A., de Freitas, N., Little, J., & Lowe, D. (2004). A boosted particle filter: Multitarget detection and tracking. In T. Pajdla & J. Matas (Eds.), *Proceedings of the 8th European Conference on Computer Vision, ECCV 2004, 3021*, 28-39: Springer-Verlag.
- Olsson, L., Nehaniv, C. L., & Polani, D. (2006). From unknown sensors and actuators to actions grounded in sensorimotor perceptions. *Connection Science*, 18(2).
- Pauli, J., & Sommer, G. (2002). Perceptual organization with image formation compatibilities. *Pattern Recognition Letters*, 23(7), 803-817.
- Philipona, D., O'Regan, J. K., & Nadal, J.-P. (2003). Is there something out there? Inferring space from sensorimotor dependencies. *Neural Computation*, 15(9).
- Philipona, D., O'Regan, J. K., Nadal, J.-P., & Coenen, O. (2004). Perception of the structure of the physical world using unknown multimodal sensors and effectors. In S. Thrun, L. Saul & B. Schoelkopf (Eds.), *Advances in Neural Information Processing Systems, 16*. Cambridge, MA: MIT Press.
- Piaget, J. (1955). *The Construction of Reality in the Child*. London: Routledge and Kegan Paul.
- Pinker, S. (1984). Visual congnition: An introduction. *Cognition*, 18, 1-63.
- Pinker, S. (1997). *How the Mind Works*. New York: W. W. Norton and Company.
- Pylyshyn, Z. W. (1984). *Computation and Cognition* (Bradford Books, 2nd ed.). Cambridge, MA: MIT Press.
- Rao, R., Shon, A., & Meltzoff, A. (2004). A bayesian model of imitation in infants and robots. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation and Social Learning in Robots, Humans, and Animals: Behaviour, Social and Communicative Dimensions*. Cambridge: Cambridge University Press.
- Reiter, R. (2001). *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*. Cambridge, MA: MIT Press.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2), 131-141.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386-408.
- Rosenbloom, P., Laird, J., & Newell, A. (Eds.). (1993). *The Soar Papers: Research on Integrated Intelligence*. Cambridge, MA: MIT Press.
- Rougier, N. P. (2001). *Hippocampal auto-associative memory*. Paper presented at the International Joint Conference on Neural Networks. Washington DC, USA.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986a). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland & T. P. R. Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (pp. 318-362). Cambridge, MA: MIT Press.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986b). Learning representations by back-propagating errors. *Nature*, 323, 533-536.

- Rumelhart, D. E., McClelland, J. L., & Group, T. P. R. (Eds.). (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.
- Sage, K., Howell, J., & Buxton, H. (2005). Recognition of action, activity and behaviour in the actIPret project. *KI-Zeitschrift Kunstliche Intelligenz, Special Issue on Cognitive Computer Vision*, 30-34.
- Sandini, G., Metta, G., & Konczak, J. (1997). *Human Sensori-motor Development and Artificial Systems*. Paper presented at the AIR&IHAS '97, Tsukuba, Japan.
- Sandini, G., Metta, G., & Vernon, D. (2004). *RobotCub: An Open Framework for Research in Embodied Cognition*. Paper presented at the IEEE-RAS/RJS International Conference on Humanoid Robotics, Santa Monica, CA.
- Schoener, G. (2006). Development as change of dynamic systems: stability, instability, and emergence. In J. P. Spencer, M. S. C. Thomas & J. L. McClelland (Eds.), *Toward a New Grand Theory of Development? Connectionism and Dynamic Systems Theory Re-Considered*. New York: Oxford University Press.
- Selfridge, O. G. (1959). *Pandemonium: A paradigm for learning*. Paper presented at the Symposium on Mechanization of Thought Processes, London.
- Shanahan, M. P. (2006). A cognitive architecture that combines internal simulation with a global workspace. *Consciousness and Cognition*, to appear.
- Shanahan, M. P., & Baars, B. (2005). Applying global workspace theory to the frame problem. *Cognition*, 98(2), 157-176.
- Shepard, R. N., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, 18, 161-193.
- Sloman, A., & Chappell, J. (2005a). *The altricial-precocial spectrum for robots*. Paper presented at the IJCAI '05 – 19th International Joint Conference on Artificial Intelligence, Edinburgh, UK.
- Sloman, A., & Chappell, J. (2005b). *Altricial self-organising information processing systems*. Paper presented at the International Workshop on the Grand Challenge in Non-classical Computation, York, UK.
- Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of early years. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(12), 1349-1380.
- Smolensky, P. (1995). Structure and explanation in an integrated connectionist/symbolic cognitive architecture. In C. Macdonald & G. Macdonald (Eds.), *Connectionism: Debates on psychological explanation*, 2, 221-290). Basel: Blackwell.
- Smolensky, P. (1996a). Computational perspectives on neural networks. In P. Smolensky, M. C. Mozer & D. E. Rumelhart (Eds.), *Mathematical perspectives on neural networks*, 1-15: Lawrence Erlbaum Associates, Florence, Kentucky, USA.
- Smolensky, P. (1996b). Computational, dynamical, and statistical perspectives on the processing and learning problems in neural network theory. In P. Smolensky, M. C. Mozer & D. E. Rumelhart (Eds.), *Mathematical perspectives on neural networks*, 1-15: Lawrence Erlbaum Associates, Florence, Kentucky, USA.
- Smolensky, P. (1996c). Dynamical perspectives on neural networks. In P. Smolensky, M. C. Mozer & D. E. Rumelhart (Eds.), *Mathematical perspectives on neural networks*, 245-270: Lawrence Erlbaum Associates, Florence, Kentucky, USA.

- Smolensky, P. (1996d). Statistical perspectives on neural networks. In P. Smolensky, M. C. Mozer & D. E. Rumelhart (Eds.), *Mathematical perspectives on neural networks*, 453-496: Lawrence Erlbaum Associates, Florence, Kentucky, USA.
- Smolensky, P., & Legendre, G. (2006). *The Harmonic Mind: From Neural Computation To Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press.
- Speidel, G. S. (1989). Imitation: a bootstrap for learning to speak. In G. E. Speidel & K. E. Nelson (Eds.), *The many faces of imitation in language learning*, 151-180: Springer-Verlag.
- Thelen, E. (1995). Time-scale dynamics and the development of embodied cognition. In R. F. Port & T. van Gelder (Eds.), *Mind as Motion – Explorations in the Dynamics of Cognition*, 69-100. Cambridge, MA: MIT Press.
- Thelen, E., & Smith, L. B. (1998). *A Dynamic System Approach to the Development of Cognition and Action* (3rd ed.). Cambridge, MA: MIT Press.
- Thorndike, E. L. (1932). *The Fundamentals of Learning*. New York: Teachers College, Columbia University.
- Thorndike, E. L. (1949). *Selected Writings from a Connectionist Psychology*. New York: Greenwood Press.
- Trevarthen, C., Kokkinaki, T., & Fiamenghi, G. A. J. (1999). What infants' imitations communicate: with mothers, with fathers and with peers. In J. Nadel & G. Butterworth (Eds.), *Imitation in Infancy*, 61-124. Cambridge: Cambridge University Press.
- van Gelder, T., & Port, R. F. (1995). It's about time: An overview of the dynamical approach to cognition. In R. F. Port & T. van Gelder (Eds.), *Mind as motion - Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.
- Varela, F. (1979). *Principles of Biological Autonomy*. New York: Elsevier North Holland.
- Varela, F. J. (1992). Whence perceptual meaning? A cartography of current ideas. In *Understanding Origins - Contemporary Views on the Origin of Life, Mind and Society*, 253-263. Boston: Kluwer Academic Publishers.
- Vernon, D. (2006a). Cognitive vision: The case for embodied perception. *Image and Vision Computing*, In press, 1-14.
- Vernon, D. (2006b). The space of cognitive vision. In H. I. Christensen & H.-H. Nagel (Eds.), *Cognitive Vision Systems: Sampling the Spectrum of Approaches*, In press, 7-26. Heidelberg: Springer-Verlag.
- Von Hofsten, C. (2003). On the development of perception and action. In J. Valsiner & K. J. Connolly (Eds.), *Handbook of Developmental Psychology*, 114-140. London: Sage.
- von Hofsten, C. (2004). An action perspective on motor development. *Trends in cognitive sciences*, 8(6), 266-272.
- Warren, W. H. (1984). Perceiving affordances: Visual guidance of stairclimbing. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 683-703.
- Weng, J. (2004). Developmental robotics: Theory and experiments. *International Journal of Humanoid Robotics*, 1(2), 199-236.
- Werbos, P. (1974). *Beyond regression: new tools for prediction and analysis in the behavioural sciences*. Boston, MA, Harvard University.

- Widrow, B., & Hoff, M. E. (1960). *Adaptive switching circuits*. Paper presented at the 1960 IRE WESCON Convention Record, New York, Woods.
- Wilson, H. R. (1999). *Spikes, Decisions, and Actions: Dynamical Foundations of Neurosciences*. New York: Oxford University Press.
- Winograd, T., & Flores, F. (1986). *Understanding Computers and Cognition - A New Foundation for Design*. Reading, MA: Addison-Wesley Publishing Company.
- Wray, R., Chong, R., Phillips, J., Rogers, S., & Walsh, B. (1994). A Survey of Cognitive and Agent Architectures (<http://ai.eecs.umich.edu/cogarch0/>): University of Michigan.
- Ziemke, T. (2001). *Are robots embodied?* Paper presented at the First International Workshop on Epigenetic Robotics — Modeling Cognitive Development in Robotic Systems, Lund, Sweden.
- Ziemke, T. (2003). What's that thing called embodiment? In A. a. Kirsh (Ed.), *25th Annual Conference of the Cognitive Science Society*, 1134-1139. Mahwah, NJ: Lawrence Erlbaum.

Biografie e foto

David Vernon è professore all'Etisalat University College (Emirati Arabi Uniti) dove insegna ingegneria dei sistemi elettronici. Collabora inoltre con l'Università di Genova nell'ambito di vari progetti europei. David Vernon è il coordinatore di euCognition, una rete europea dedicata allo studio dei sistemi cognitivi artificiali e membro del progetto RobotCub. Quest'ultimo ha tra i suoi obiettivi quello di creare un sistema robotico umanoide *open source*. Nel passato, ha lavorato alla Westinghouse Electric, al Trinity College a Dublino, alla Commissione Europea, alla National University in Irlanda (Maynooth), e per la Science Foundation sempre in Irlanda. I suoi interessi includono aspetti di visione artificiale e dell'uso del paradigma enattivo per la realizzazione di modelli dei sistemi cognitivi.



Giorgio Metta è ricercatore all'Università di Genova dove insegna Robotica Antropomorfa e Sistemi Operativi per il corso di laurea in Bioingegneria. Collabora inoltre come ricercatore con l'Istituto Italiano di Tecnologia anch'esso con sede a Genova. Giorgio Metta ha un dottorato in ingegneria elettronica conseguito presso l'Università di Genova. Dal 2001 al 2002 ha lavorato nel corso di un postdottorato presso il laboratorio di Intelligenza Artificiale del MIT. Dal 1993 è al LIRA-Lab dell'Università di Genova dove si è occupato dello sviluppo di varie piattaforme robotiche con l'obiettivo di implementare modelli biomorfi della coordinazione sensomotoria. Fa parte del gruppo di lavoro del progetto RobotCub nel quale si occupa della realizzazione di una piattaforma robotica umanoide. Per maggiori informazioni: <http://pasa.liralab.it>



Giulio Sandini è direttore di ricerca all'Istituto Italiano di Tecnologia e professore ordinario in bioingegneria all'Università di Genova. La sua ricerca è nel campo della neuroscienza computazionale e nella robotica con l'obiettivo di comprendere i meccanismi della coordinazione sensomotoria e dello sviluppo cognitivo. Giulio Sandini è ingegnere elettronico (bioingegneria) ed è stato ricercatore presso la Scuola Normale

Superiore di Pisa fino al 1984 lavorando nel laboratorio di neurofisiologia del CNR. È stato inoltre ricercatore presso il dipartimento di neurologia della Harvard Medical School e “Visiting Professor” all’Artificial Intelligence lab al MIT. Dal 2006 è direttore di ricerca all’Istituto Italiano di Tecnologia dove guida il dipartimento di “robotica, scienze cognitive e del cervello”.

