

IL FONDAMENTO DEI SIMBOLI  
ATTRAVERSO L'ESPERIENZA CON IL MONDO:  
UN APPROCCIO CONNESSIONISTA

1. IL PROBLEMA DEL FONDAMENTO DEI SIMBOLI NELLA METODOLOGIA  
SIMULATIVA

La simulazione ha costituito un'importante evoluzione metodologica nello studio dei processi cognitivi. Tali processi, infatti, sono il prodotto di sistemi complessi che è talvolta difficile o impossibile manipolare direttamente attraverso la metodologia sperimentale. La metodologia simulativa offre in alternativa la costruzione di sistemi artificiali che presentano analogie con quelli naturali da spiegare e che ovviamente possono essere controllati e manipolati a piacimento.

Se i vantaggi della metodologia simulativa sono evidenti (cfr. Greco 1988; Parisi 2001; Cangelosi e Parisi 2002 per altri esempi), tuttavia essa presenta una maggiore criticità rispetto a quella sperimentale. Come nella metodologia sperimentale, nella simulazione la manipolazione avviene allo scopo di mettere alla prova una teoria e quindi per entrambe le metodologie è fondamentale identificare e specificare le variabili in gioco. Ma se già nelle indagini sperimentali l'interpretazione della procedura e dei risultati deve affrontare il nodo critico della correttezza delle variabili individuate e delle loro effettive relazioni, nella metodologia simulativa a ciò si aggiunge un ulteriore momento critico, in quanto i fenomeni manipolati e osservati sono riferiti a un modello e non alla realtà. Appare perciò fondamentale la considerazione della natura delle analogie su cui il modello si fonda.

In linea generale, in un modello cognitivo le analogie possono riguardare l'input, il processo, l'output. Di solito si assume che i dati forniti in ingresso al modello siano analoghi agli stimoli e che l'output sia analogo alle risposte; da ciò si inferisce che l'elaborazione che il modello compie sia *analogica* al processo cognitivo supposto dalla teoria, che ne ricaverebbe pertanto un sostegno. Questo era, almeno, lo schema interpretativo dei primi tentativi di costruzione di quei «sistemi artificiali» che costituiscono il fulcro della metodologia simulativa e che oggi vanno sotto il nome di modelli «simbolici». Tali modelli usano tradizionali programmi per computer e in quanto tali hanno natura

simbolica e algoritmica: si tratta di definire come i dati in ingresso devono essere rappresentati simbolicamente e quindi elaborati secondo una sequenza di passi il cui percorso è fondamentalmente determinato dall'esito di una serie di test condizionali.

Uno dei problemi delle simulazioni di questo tipo è che il significato dei simboli attraverso cui le variabili sono rappresentate (a qualunque livello, input-processo-output) dipende dall'interpretazione di chi costruisce o usa il modello. Consideriamo, ad esempio, uno dei primi e più noti sistemi di comprensione del linguaggio, SHRDLU (Winograd 1972) che pur nella sua semplicità appare paradigmatico. Quel sistema era in grado di analizzare la struttura grammaticale di un input linguistico, relativo a un semplice mondo fatto di cubi e piramidi, per determinarne il «significato»; poteva rispondere a domande sulla posizione di oggetti o elaborare piani per eseguire azioni (in effetti l'idea di Winograd ha in qualche modo anticipato gli attuali *agenti*). Se ci si chiede in che modo gli oggetti su cui il sistema era in grado di conversare e su cui poteva agire fossero rappresentati, è evidente che si trattava di espressioni puramente simboliche (del tipo «È 001 CUBO» oppure «HA COLORE 001 VERDE», supponendo che 001 fosse il codice corrispondente a un elemento della lista di oggetti definiti come presenti in un certo ambiente). Il significato della stringa «CUBO» non poteva essere legato in alcun modo a un'esperienza diretta del sistema con le proprietà percettive dei cubi e la conoscenza di certe caratteristiche di questo tipo di oggetti rispetto ad altri come le piramidi (ad esempio il fatto che ai cubi si possono sovrapporre altri oggetti mentre alle piramidi no) derivava da una definizione costituita semplicemente dall'associazione di alcuni simboli con altri simboli. La definizione dei simboli del sistema rinvia o ad altri simboli definienti o, in ultima analisi, al significato assegnato dal costruttore ai simboli primitivi.

L'esempio che abbiamo considerato non era stato proposto in realtà quale modello simulativo della comprensione del linguaggio, ma come un sistema con finalità di intelligenza artificiale, ed era stato superato più per le limitazioni delle sue prestazioni che per il problema che abbiamo esposto. Ma per i modelli con finalità simulative il problema è effettivo perché le spiegazioni basate sui modelli simbolici funzionano solo se i simboli sono interpretati come inteso da chi li costruisce. Tipicamente, la spiegazione sta nel fatto che fra l'input e l'output ci sia luogo per certe specifiche *rappresentazioni*, che tuttavia per il sistema sono codici simbolici (ad es. variabili rappresentate come stringhe o come numeri) e null'altro. Il fatto che una stringa rappresenti una particolare variabile invece che un'altra è solo una questione di plausibilità di interpretazione nel contesto da parte di chi usa il sistema.

Questo problema si presenta in tutti i modelli – e sono la maggioranza – in cui *per il sistema* la definizione di ciascun simbolo si trova in altri simboli da definire oppure la cui definizione ultima è nell'interpre-

tazione del costruttore o utente<sup>1</sup>. Il caso sarebbe diverso se i simboli primitivi del sistema fossero collegati a esperienze dirette, percettive o motorie, del sistema stesso – che ne costituirebbero il fondamento (*grounding*). Il problema che abbiamo presentato viene dunque detto del *symbol grounding*<sup>2</sup>. Esso è stato evidenziato per la prima volta con chiarezza da Stevan Harnad (1990), che – riprendendo il famoso argomento della stanza cinese con cui Searle (1980) aveva posto il cognitivismo di fronte alla difficoltà di modelli della mente puramente simbolici – ha paragonato i sistemi che ne sono afflitti alla situazione di una persona a cui fosse chiesto di tradurre dal cinese avendo come definizione per ogni parola altre parole cinesi. Sarebbe un continuo rinvio da un simbolo all'altro, in una «giostra» senza fine.

La condizione che i simboli linguistici siano direttamente, e intrinsecamente al modello, legati a un'esperienza percettiva e cognitiva di costruzione di referenti semantici (ad esempio categorie) è particolarmente importante nei modelli computazionali di produzione e comprensione del linguaggio. La *performance* dei modelli in cui il significato dei simboli linguistici è semplicemente costituito da altri simboli astratti non è generalizzabile alle abilità linguistiche di soggetti umani. Invece i sistemi cognitivi con fondamento dei simboli sono in grado di costruire, attraverso l'esperienza, delle rappresentazioni categoriali e di usarle come referente semantico per i simboli lessicali usati per comunicare. Questo è esattamente quanto avviene nei bambini, che devono essere in grado di costruire categorie e poi di imparare ad associare i suoni uditi dagli adulti alle proprie categorie.

In questo articolo esamineremo come il problema del *symbol grounding* influenzi la capacità esplicativa dei modelli simulativi, sia di quelli tradizionalmente definiti «simbolici» che di quelli «non simbolici» che fanno uso di reti neurali. In particolare discuteremo in quali modalità e a quali condizioni l'implementazione di sistemi connessionisti possa fornire una soluzione a tale problema. Infine presenteremo a titolo esemplificativo una serie di simulazioni realizzate dal nostro gruppo di ricerca, volte a realizzare modelli i cui simboli siano fondati, in cui si è tentato di conseguire questo obiettivo attraverso l'adozione di un approccio connessionista.

<sup>1</sup> Naturalmente questo può essere considerato un vero «problema» solo per i modelli simulativi e robotici che sono usati per studiare i sistemi cognitivi naturali (degli animali e degli uomini). Nei modelli computazionali di intelligenza artificiale, invece, non vi è alcuna necessità di assicurarsi che lo stesso sistema sia in grado di costruire i simboli e i suoi significati, perché il loro scopo è solo strumentale, cioè di fornire un'applicazione informatica che aiuti gli utenti nello svolgimento di compiti pratici.

<sup>2</sup> In questo articolo useremo in maniera intercambiabile le espressioni *symbol grounding* e «fondamento dei simboli».

## 2. DIVERSI APPROCCI AL *SYMBOL GROUNDING*

Abbiamo detto che possono essere considerati immuni dal problema del *symbol grounding* quei sistemi i cui simboli sono collegati a esperienze percettive o motorie del sistema stesso. Cosa sono queste «esperienze»?

Una prima risposta a questa domanda è considerare presupposto per fondare i simboli qualunque percezione o azione nel mondo reale. In questo senso, qualunque sistema che sia posto in un ambiente e interagisca con esso (un *agente*), costruendosene autonomamente delle rappresentazioni, dovrebbe essere per sua natura fondato. Questo approccio, che possiamo definire «robotico», è quello adottato ad esempio da Steels e collaboratori (Steels 2002; Vogt 2002). Secondo questa prospettiva, il problema del *symbol grounding* dipenderebbe solo dal fatto di considerare i simboli come equivalenti alle rappresentazioni, mentre invece dovrebbero essere considerati come il processo che costruisce le rappresentazioni attraverso l'accoppiamento dei nomi con i loro referenti. In questo senso, una rappresentazione è fondata se i dati sensoriali posti in input sono trasformati in rappresentazioni attraverso un processo autonomo dell'agente e dunque le rappresentazioni non sono poste nel sistema direttamente dal costruttore come avveniva nei modelli della prima intelligenza artificiale. Questo processo è innescato attraverso un'interazione fra uomo e robot, cioè attraverso *giochi linguistici* in cui ogni simbolo linguistico è usato in un contesto (ad es. indicare, mostrare un oggetto) e associato al referente per ostensione.

Secondo questo approccio, il problema potrebbe essere risolto semplicemente evitando di considerare il concetto di simbolo, come si fa nell'uso più comune, come uno stato di rappresentazione ma sottolineandone invece l'aspetto di processo di rappresentazione (cfr. Greco 1995 per un'analisi di questi due sensi del concetto di rappresentazione; cfr. Brooks 1999, 109). Nella prospettiva robotica l'accento è dunque spostato dalle rappresentazioni al processo che le produce, in questo caso algoritmi che consentono di costruire, percorrere, potare rami di alberi decisionali e compiere scelte<sup>3</sup>.

Tuttavia questo approccio non risolve del tutto il problema, perché appare illusorio pensare ai processi che operano in un sistema senza fare qualche riferimento agli stati del sistema stesso, cioè – nel caso dei sistemi cognitivi – a qualche forma di rappresentazione interna a cui gli

<sup>3</sup> Gli alberi decisionali sono una metodologia di intelligenza artificiale per classificare diversi stimoli o processi, secondo la quale si arriva alla rappresentazione di un concetto attraverso una serie di scelte successive tra caratteristiche alternative (ad esempio, per un «cerchio rosso», tra varie possibili forme si sceglie il cerchio, quindi tra vari possibili colori si sceglie il rosso e così via); dall'idea delle ramificazioni tra alternative derivano le metafore «arboree» come quella della potatura.

stimoli sensoriali danno origine. Se il riferimento al concetto di rappresentazione non è eliminabile, è allora evidente che, perché il problema del *symbol grounding* sia davvero superato e si possa uscire dalla giostra, qualunque cosa sia ciò che fornisce fondamento ai simboli, non può trattarsi di altri simboli ma deve trattarsi di qualche rappresentazione non simbolica.

Queste considerazioni fanno ritenere preferibile considerare le «esperienze» del sistema come rappresentazioni degli input di tipo analogico che il sistema ha ricevuto in passato, ad esempio l'attivazione di certi recettori sensoriali o l'informazione che arriva in *feedback* da movimenti muscolari. Gli input analogici sono per loro natura espressione di variazioni di energia fisica (ad es. luminosa) in un continuum; è possibile ipotizzare una corrispondente rappresentazione analogica all'interno di un sistema (ad es. la carica di un condensatore proporzionale all'intensità di un segnale elettrico ricevuto in input) e, per estensione, considerare analogica qualunque rappresentazione che abbia una relazione di somiglianza diretta con ciò che rappresenta (come un orologio che esprime il passare del tempo con il movimento di una lancetta). L'uso di simboli linguistici per denotare queste esperienze percettivo-motorie è invece in qualche modo discreto e arbitrario, come la *parola* «gatto» rispetto al disegno di un gatto<sup>4</sup>.

La questione del passaggio dall'analogico al digitale, dal non simbolico al simbolico, è in effetti un altro modo di esprimere uno degli argomenti più affascinanti delle scienze cognitive attuali e cioè la nascita del linguaggio.

### 3. ASPETTI SIMBOLICI E NON SIMBOLICI NELLE RETI NEURALI

Le reti neurali sono sembrate fin dal loro apparire le candidate ideali per la simulazione degli aspetti non simbolici della cognizione. Infatti, almeno nel caso delle reti distribuite, le singole unità non sono simbolo di alcunché in quanto la rappresentazione si trova nel *pattern* dei pesi delle connessioni tra le unità e viene attivata nel suo complesso da una particolare configurazione di attivazione delle singole unità.

Si potrebbe dunque pensare che i modelli connessionisti siano di per sé immuni dal problema del *symbol grounding*, ma le cose non stanno

<sup>4</sup> Una distinzione simile a quella qui presentata è quella posta da Palmer (1978) tra sistemi rappresentativi *intrinseci*, che non sono soggetti a vincoli arbitrari ma presentano caratteristiche inerenti ai sistemi rappresentati, e sistemi *estrinseci*, che invece introducono aspetti arbitrari, non inerenti a ciò che rappresentano. Anche molta parte del dibattito sulla rappresentazione delle immagini mentali (Pylyshyn 1981) e lo stesso concetto di *modello mentale* di Johnson-Laird (1983, 263), naturalmente, sono stati basati su tale distinzione.

così. In primo luogo, in alcuni sistemi connessionisti esistono a volte rappresentazioni «localistiche» che vengono usate e interpretate dagli autori del modello esattamente come nei modelli simbolici. In questi casi ovviamente il problema torna a presentarsi, perché si attribuisce a certe unità del sistema un significato concettuale estraneo al sistema stesso. D'altra parte, qualora le rappresentazioni fossero totalmente distribuite e non interpretate, il problema del fondamento dei simboli non esisterebbe per il semplice fatto che non ci sarebbero più simboli. In questo caso rimarrebbe però da mostrare l'utilità euristica di un ipotetico sistema totalmente privo di rappresentazioni simboliche. Un motivo che renderebbe poco interessante dal punto di vista simulativo un tale modello è il fatto che le analogie su cui abbiamo visto fondarsi il concetto stesso di simulazione sono precisabili solo in termini simbolici; altrimenti sarebbe difficile sostituire il linguaggio dell'attivazione di unità o di pesi di connessioni con riferimenti a variabili che abbiano un significato dal punto di vista psicologico (Greco 1997). Di fatto, nella pratica, non esistono reti neurali le cui unità fisiche o moduli funzionali non abbiano a qualche livello (input-strati intermedi-output) un'interpretazione simbolica e quindi che non pongano il problema del *symbol grounding*.

L'interpretabilità di qualche elemento del sistema in termini simbolici e arbitrari è anche la condizione essenziale per poter simulare un agente dotato di linguaggio. Proprio sul terreno del linguaggio ci sono state le controversie più accese tra i fautori dell'approccio cognitivo simbolico tradizionale e i connessionisti (cfr. Fodor e Pylyshyn 1988; Smolensky 1988). Queste classiche controversie si sono scontrate soprattutto sul problema della *composizionalità*. Con questo termine si fa riferimento al fatto che i significati vengono espressi nel linguaggio attraverso la combinazione e la possibile ricombinazione di costituenti elementari discreti (fonemi, morfemi, parole, sintagmi) secondo certe regole che indicano l'accettabilità delle forme risultanti (grammatica). Le reti neurali, secondo la critica classica di Fodor e Pylyshyn, non possedendo simboli discreti non potrebbero consentire di riprodurre né di spiegare questa caratteristica del linguaggio.

Molta parte di questa discussione naturalmente dipende da che cosa si intende per *simbolo*. Per quanto la definizione di che cosa sia un simbolo sia una delle più controverse, l'aspetto su cui più o meno tutti sono disposti a trovarsi d'accordo è il fatto che un simbolo «sta al posto di» qualcos'altro. Secondo una classica distinzione di Peirce (1978), anche le *icone* e gli *indizi* stanno al posto di qualcos'altro, ma le icone hanno una somiglianza fisica con ciò che rappresentano e gli indizi sono contigui nello spazio o nel tempo con ciò che rappresentano; i simboli invece hanno la caratteristica di essere arbitrari ed essere associati a ciò che rappresentano solo in virtù di una convenzione. In questo senso, i simboli più tipici sono quelli del linguaggio.

Come si è detto, nelle reti neurali distribuite le singole unità non stanno al posto di unità concettualmente identificabili e separate (singoli oggetti, singole caratteristiche, ecc.), che tutt'al più possono essere rappresentate da *patterns* di unità e dai pesi delle relative connessioni. In qualche caso una rete può organizzarsi per effetto dell'apprendimento in modo tale che certe unità o *cluster* di unità costituiscano, di fatto, rappresentazioni di tratti distintivi a livello microscopico (Smolensky 1988, ha usato il termine *microfeatures* e ha parlato di rappresentazioni *subsimboliche*), ma queste rappresentazioni non hanno comunque caratteristiche di arbitrarietà e non possono essere usate in maniera componenziale come nel linguaggio.

Un ulteriore problema presentato dalle reti neurali distribuite, corollario di quanto si è appena visto, è che le conoscenze acquisite da queste reti per affrontare un certo problema – se non hanno natura simbolica – non possono essere riutilizzate per altri problemi, anche leggermente diversi da quello precedente; di fatto è necessario un addestramento specifico per ogni compito (Clark e Karmiloff-Smith 1993; Greco e Cangelosi 1999).

Riassumendo quanto si è detto, emerge che il valore euristico di sistemi connessionisti totalmente distribuiti è dubbio se essi restano privi di interpretabilità simbolica; d'altra parte, un modello privo di rappresentazioni simboliche sarebbe poco interessante e non avrebbe la capacità di simulare il funzionamento del linguaggio. Se però una qualunque parte di un modello è simbolica o interpretabile come simbolica, allora il problema del *symbol grounding* torna a presentarsi. Non è evidentemente il fatto di non possedere simboli che rende un modello «fondato», ma il fatto che le rappresentazioni simboliche usate nel modello siano riferite a rappresentazioni non simboliche presenti nello stesso modello e non solo nella mente di chi lo interpreta.

#### 4. L'APPROCCIO CONNESSIONISTA AL *SYMBOL GROUNDING*

Come si è detto, la prima chiara formulazione del problema del *symbol grounding* è di Harnad (1990), che già ipotizzava una possibile soluzione basata sull'uso di reti neurali. I primi modelli di Harnad e collaboratori (Harnad, Hanson e Lubin 1992; Harnad 1993) erano costituiti da reti neurali a tre strati (input, unità nascoste, output) a cui venivano presentate linee poste in un continuum di diverse lunghezze. In una prima fase le reti venivano sottoposte a un pre-addestramento di tipo autoassociativo (in cui com'è noto l'output è identico all'input), in modo che imparassero essenzialmente a discriminare le linee una dall'altra. Nella seconda fase, alle linee veniva associato un diverso input (una sorta di nome) in modo da insegnare loro (tramite l'algoritmo della

*backpropagation*, che fornisce un *feedback* correttivo alle risposte) a classificare le linee in due categorie (corte e lunghe).

Si trattava di una classificazione banale per un modello di quel tipo; ma la cosa che colpì Harnad fu il fatto che l'analisi delle rappresentazioni interne delle lunghezze (si calcolava la distanza euclidea fra i vettori di attivazione delle unità nascoste per ciascuna linea) rivelava un'interessante differenza tra la prima fase della semplice discriminazione e la seconda fase di categorizzazione. Succedeva che nella fase di discriminazione le linee venivano rappresentate in un continuum (le distanze tra i vettori erano uguali), ma dopo la categorizzazione le distanze tra tutte le linee «corte» erano di misura inferiore rispetto alle distanze tra linee «corte» e linee «lunghe»; inoltre, le distanze tra le linee «lunghe» e «corte» erano aumentate per via dell'apprendimento categoriale. In altri termini, le distanze *all'interno delle* categorie erano state compresse e le distanze *tra le* categorie erano state espanse.

Questo effetto, che è noto come «percezione categoriale» (*categorical perception*), è esattamente quello che si verifica nella percezione e categorizzazione umana. È classico il caso della percezione dei colori (Berlin e Kay 1969) e dei suoni linguistici (Liberman *et al.* 1957); l'effetto è stato trovato di recente da Nakisa e Plunkett (1998) anche in una simulazione dell'apprendimento fonologico e, come vedremo, è presente anche nei nostri modelli. Secondo una definizione molto suggestiva di Harnad, in questi casi lo spazio della somiglianza fra gli esempi percepiti viene in qualche modo «deformato» in modo tale da non riflettere più l'uniformità di stimoli che per loro natura sono posti in un continuum.

Il fenomeno della percezione categoriale esprime la differenza tra la semplice *discriminazione*, cioè la capacità di distinguere uno stimolo dagli altri in un certo contesto, e l'*identificazione*, cioè la capacità di attribuire un'identità stabile a un certo stimolo rispetto a tutti gli altri. La prima nasce da un giudizio relativo ed è di tipo sensoriale, la seconda nasce da un giudizio assoluto e ha natura categoriale. La percezione categoriale riguarda solo gli stimoli categorizzati, cioè i casi in cui non si hanno risposte a stimoli isolati, ma a *classi* di stimoli accomunati da qualche somiglianza. La deformazione dello spazio che si verifica in questi casi costituisce un primo effetto della rappresentazione, più compatta rispetto alle prime «proiezioni» sensoriali degli stimoli; la sua natura è tuttavia ancora analogica e non simbolica.

Secondo Harnad, le proiezioni analogiche prossimali degli stimoli distali (che chiama rappresentazioni *iconiche*) sono uno dei due tipi di rappresentazioni non simboliche che fondano «dal basso» le rappresentazioni simboliche; le altre sono le rappresentazioni *categoriali*, che contengono le caratteristiche invarianti estratte dalle proiezioni sensoriali ma non possono essere ancora considerate simboliche perché, pur essendo associate a nomi, costituiscono una semplice tassonomia di eti-



chette indipendenti, che per esempio non è possibile combinare tra loro per costruire ed esprimere nuovi significati. Soltanto quando i nomi sono combinati in proposizioni, che possono avere un valore di verità, essi hanno pieno valore simbolico.

Harnad ha definito le simulazioni con reti neurali che esibiscono la capacità del *symbol grounding* come sistemi ibridi (Harnad 1993; 2002), ma questa espressione non va intesa nel senso più tradizionale di sistemi formati da moduli simbolici e non simbolici. Infatti né le prime simulazioni di cui si è appena riferito, né le successive (comprese quelle che il nostro gruppo di ricerca ha compiuto e di cui riferiremo in seguito) comprendono alcuno specifico modulo di elaborazione simbolica. La peculiarità di questi sistemi è che, essendo i simboli fondati su esperienze sensomotorie compiute in precedenza dallo stesso modello, essi possono essere «messi insieme» in combinazioni di più alto livello in un modo del tutto nuovo rispetto alla classica manipolazione di simboli, il cui unico vincolo era la sintassi, cioè le regole per combinare i simboli solo sulla base di aspetti formali e arbitrari. Nei sistemi fondati, invece, esiste un secondo vincolo, quello degli aspetti non arbitrari degli invarianti sensoriali che collegano i simboli alle proiezioni sensoriali, di natura analogica, degli oggetti a cui si riferiscono. Questo tipo di elaborazione, come mostreremo nel seguito di questo articolo, può essere compiuto anche da un sistema puramente connessionista.

## 5. LA REALIZZAZIONE DI MODELLI «FONDATI» CON L'USO DI RETI NEURALI

In questo paragrafo presenteremo degli esempi di simulazioni connessioniste realizzate con l'obiettivo di superare il problema del *symbol grounding*. Lo scopo di tali modelli<sup>5</sup> è, in generale, di creare sistemi cognitivi autonomi i cui simboli siano intrinsecamente connessi con l'abilità di acquisire categorie dall'esperienza che hanno dell'ambiente. I simboli così fondati vengono successivamente impiegati in proposizioni che descrivono categorie nuove: attraverso questo processo, che definiamo di «trasferimento del fondamento» (*grounding transfer*), i sistemi acquisiscono l'abilità di esprimere significati più complessi, che vanno al di là della percezione immediata. Per queste finalità, facciamo uso di reti neurali in cui esistono sia rappresentazioni di tipo analogico sia rappresentazioni simboliche e arbitrarie.

Le reti ricevono il loro input da una retina su cui vengono proiettate delle immagini che vengono categorizzate sfruttando rappresentazioni analogiche. Inoltre possono ricevere un input simbolico rappresentato

<sup>5</sup> Alcuni dei dati qui riportati sono stati presentati al «I Convegno Nazionale della Società Italiana di Scienze Cognitive» (Rovereto, 20-21 settembre 2002) e alla «13<sup>th</sup> International Conference on Artificial Neural Networks» (Istanbul, 26-29 giugno 2003).

da *patterns* arbitrari. L'input visivo viene messo in relazione con input e output simbolici attraverso un processo di apprendimento diretto in cui sono presentati contemporaneamente alle reti stimoli visivi e simbolici. I simboli vengono fondati su queste rappresentazioni analogiche, cioè sull'esperienza del sistema del suo ambiente circostante. Otteniamo quindi dei sistemi ibridi utilizzando comunque un'architettura interamente connessionista.

I nostri modelli vengono addestrati in modo incrementale: nella prima fase le reti imparano a categorizzare gli stimoli che vengono proiettati sulla retina. Nella seconda fase acquisiscono simboli che denotano le categorie sensomotorie apprese attraverso un processo in cui vengono presentati sia stimoli retinici che simbolici. Durante la terza fase le reti ricevono esclusivamente input simbolici, proposizioni che descrivono categorie di stimoli mai incontrati prima. Il *grounding* dei simboli che costituiscono queste proposizioni viene trasferito ai simboli che denotano le categorie nuove ed è questo il processo di *grounding transfer*. Per testare se i simboli di cui sono composte le descrizioni trasmettono effettivamente il *grounding* ai simboli nuovi, gli stimoli retinici descritti da queste proposizioni sono presentati al modello per la prima volta nell'ultima fase.

### 5.1. Simulazione I

*Stimoli.* In questo modello, dieci reti neurali sono state addestrate a categorizzare e denominare immagini di forme geometriche (cerchi, ellissi, quadrati e rettangoli).

*Architettura.* Sono state usate 10 reti a tre strati, di tipo *feed-forward*, con i pesi iniziali casuali. Lo strato di input conteneva due gruppi di unità: 49 unità che simulavano la retina e 6 unità linguistiche (una per ciascun nome di categoria: «cerchio», «ellissi», «quadrato», «rettangolo», «simmetrico», «asimmetrico»). Lo strato nascosto era formato da 5 unità connesse a entrambi i gruppi di unità di input. Lo strato di output aveva la stessa organizzazione di quello di input, con 49 unità corrispondenti a quelle retiniche più 6 unità per i nomi.

*Procedura.* La procedura di *training* era costituita da tre fasi di apprendimento supervisionato mediante la tecnica di *backpropagation*. Nel primo compito (*prototype sorting*), le reti imparavano a categorizzare le figure: avendo in input le forme, producevano in output il prototipo della categoria. Nella seconda fase (*entry-level naming*), le reti imparavano ad associare ciascuna forma con il corrispondente nome. La parte più interessante della simulazione era il terzo compito (*higher-level*), in cui alle reti venivano presentati i nomi già fondati (ad esempio «quadrato» o «rettangolo») in congiunzione con nuovi nomi arbitrari («simmetrico», «asimmetrico»). Questo equivale in un certo senso ad insegnare

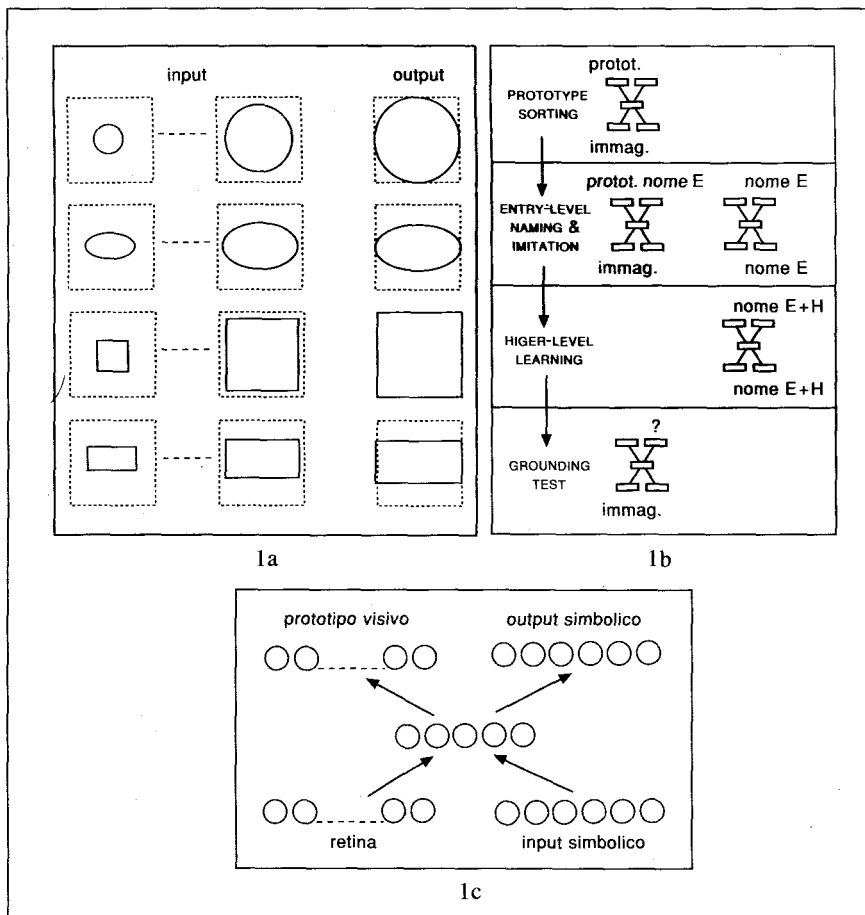


FIG. 1. Simulazione I. 1a: immagini usate per l'addestramento della rete; 1b: procedura di apprendimento e di test; 1c: architettura della rete neurale.

che «cerchio e quadrato [sono figure] simmetriche» e che «ellissi e rettangolo [sono figure] asimmetriche». È stata così insegnata alle reti una categoria di ordine superiore, associando le nuove stringhe a quelle relative ai simboli fondati in precedenza.

## 5.2. Simulazione II

La prima simulazione ha mostrato che è possibile trasferire il fondamento posseduto da simboli elementari che si riferiscono a *individui* (ad es. quadrato) a categorie di livello superiore (ad es. simmetrico). Tale

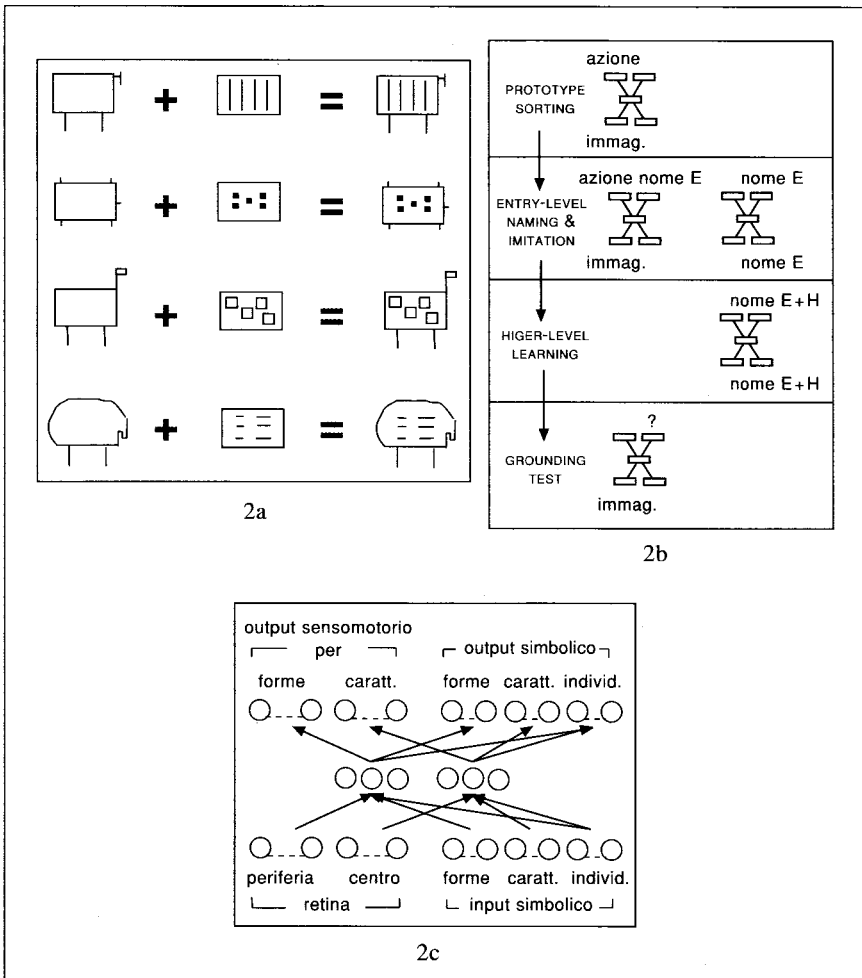


FIG. 2. Simulazione II. 2a: immagini usate per l'addestramento della rete; 2b: procedura di apprendimento e di test; 2c: architettura della rete neurale.

modello è stato replicato ed esteso ad una congiunzione *intensionale*, cioè basata sulla combinazione di *caratteristiche*. Abbiamo addestrato dieci reti a categorizzare e denominare immagini di forme animali (ad es. cavallo, tartaruga) e di loro caratteristiche (ad es. strisce, puntini). È stata quindi insegnata alle reti una categoria di ordine superiore basata sulla combinazione di caratteristiche per definire un nuovo animale (ad es. zebra).

*Architettura.* Sono state usate 10 reti simili alle precedenti. Lo strato di input era anche in questo caso diviso in un gruppo di unità retiniche

e un gruppo di unità linguistiche. A differenza del precedente modello, le unità retiniche sono state suddivise in una parte centrale e una periferica. Lo strato nascosto, costituito da 6 unità, era connesso a quello di input in modo che le prime tre unità nascoste ricevessero input dalla periferia della retina e le altre tre unità dalla parte centrale della retina stessa. In questo modo si consentiva alle unità nascoste di specializzarsi nel riconoscere le forme o le trame. Per l'apprendimento, in questa simulazione non veniva richiesto in output il prototipo delle categorie, ma una risposta considerata di tipo sensomotorio. Lo strato di output era suddiviso in due gruppi di unità come quello di input: il primo conteneva 8 unità che rappresentavano l'output sensomotorio e il secondo conteneva 12 unità linguistiche strutturate in maniera simile a quelle di input. Lo strato di output era connesso con quello nascosto in modo particolare per ottenere una struttura modulare, in cui tre unità si specializzavano nel rappresentare forme animali e le altre tre nel rappresentare le loro caratteristiche.

*Procedura.* La procedura è stata del tutto simile a quella adottata nella prima simulazione. Dopo le fasi di categorizzazione e di denominazione delle categorie di base, in cui era sempre presente l'input visivo, nella terza fase (apprendimento di alto livello) sono state presentate alle reti soltanto stringhe di simboli, costituite dalla combinazione di nomi di forme animali e di loro caratteristiche (ad es. cavallo, strisce, zebra), che definivano 4 categorie di alto livello.

Nella fase di test sono stati presentati 180 nuovi stimoli visivi, che combinavano insieme forme e *texture* per formare immagini di nuovi animali (ad es. cavallo + strisce = zebra). Anche se questi stimoli non erano mai stati presentati alle reti, queste sono state in grado di denominarli correttamente, mostrando che anche in questo caso il *grounding* era stato trasferito dai simboli di base a quelli di ordine superiore.

### 5.3. Simulazione III

I risultati della precedente simulazione, basata su un'architettura in cui le parti periferica e centrale dello stimolo retinico erano separate, potevano dare adito al dubbio che la corretta *performance* fosse indotta da un artificio. Infatti, gli stimoli erano costruiti in modo tale che le forme occupassero sempre la zona periferica della retina e le *texture* in esse contenute si trovassero sempre in quella centrale. Nella terza simulazione sono stati usati stimoli completamente diversi e la retina non è stata più divisa. Un secondo problema, condiviso da entrambe le prime simulazioni, era il fatto che non fosse del tutto chiaro quale output le reti dovessero fornire durante la prima fase dell'apprendimento, quella riguardante le categorie di base. Nella prima simulazione l'output era una sorta di duplicazione della retina e veniva richiesto di fornire il prototi-

po della categoria che veniva appresa. Nella seconda simulazione, l'output era di tipo localistico e veniva interpretato come una sorta di risposta sensomotoria. La plausibilità di queste interpretazioni non era tuttavia del tutto soddisfacente. La terza simulazione ha ovviato al problema lasciando solo l'output simbolico.

*Stimoli.* Sono stati usati quali stimoli 3 forme geometriche (quadrato, croce, punti) e 3 colori (rosso, verde, blu). In totale sono stati presentati 81 stimoli (derivanti dalla combinazione delle 3 forme con i 3 colori e ponendo ogni forma in 9 posizioni diverse), di cui 54 presentati nella fase di *training* e i rimanenti 27 nella fase di test. Ogni pixel dell'immagine occupava tre unità nella retina, una per ogni colore primario.

*Architettura.* Sono state usate 30 reti *feed-forward* con tre strati (input, output, intermedio). Per quanto riguarda l'input, la retina conteneva 75 unità e le unità simboliche erano 9. Lo strato nascosto era costituito da 6 unità completamente interconnesse con le unità di input. Lo strato di output comprendeva 3 unità per i colori, 3 per le forme e 3 per gli oggetti composti presentati durante il test. Abbiamo ottenuto una specializzazione delle unità nascoste attraverso la loro connessione selettiva con le unità di output: tre unità erano connesse alle unità di output per il colore e tre alle unità per la forma. Tutte le unità *hidden* erano invece connesse alle unità per gli oggetti composti.

*Procedura.* La procedura di apprendimento è stata quella delle precedenti simulazioni, con la differenza che durante la prima fase le reti rispondevano agli stimoli visivi corrispondenti a forme di un certo colore (ad es. quadrato rosso) direttamente con un output simbolico, lo stesso che poi è stato associato nella seconda fase ai nomi delle forme e dei colori. Nella terza fase le reti hanno imparato a denominare figure non ancora viste attraverso descrizioni simboliche costituite dai nomi fondati nella fase precedente associati a nuove etichette arbitrarie (ad es. una proposizione potrebbe avere la forma «quadrato + rosso = DAX»).

Oltre a simboli che denotano oggetti già recepiti abbiamo presentato anche descrizioni di oggetti mai presentati. Per esempio, se proiettiamo immagini sulla retina di quadrati rossi e blu, triangoli verdi e rossi, ma mai di quadrati verdi o di triangoli di colore blu, forniremo comunque descrizioni simboliche della forma «quadrato + verde = SOD». Se le reti riescono a rispondere con i simboli che denotano queste categorie di oggetti nuovi quando le proiettiamo sulla retina per la prima volta, allora possiamo concludere che anche questi simboli sono fondati sull'esperienza che queste reti hanno dell'ambiente, e che il *grounding* viene effettivamente trasferito dai simboli di cui sono composte le proposizioni ai simboli nuovi.

In effetti quando, nella fase di test, sono state presentate per la prima volta sulla retina le figure definite solo simbolicamente nella terza fase

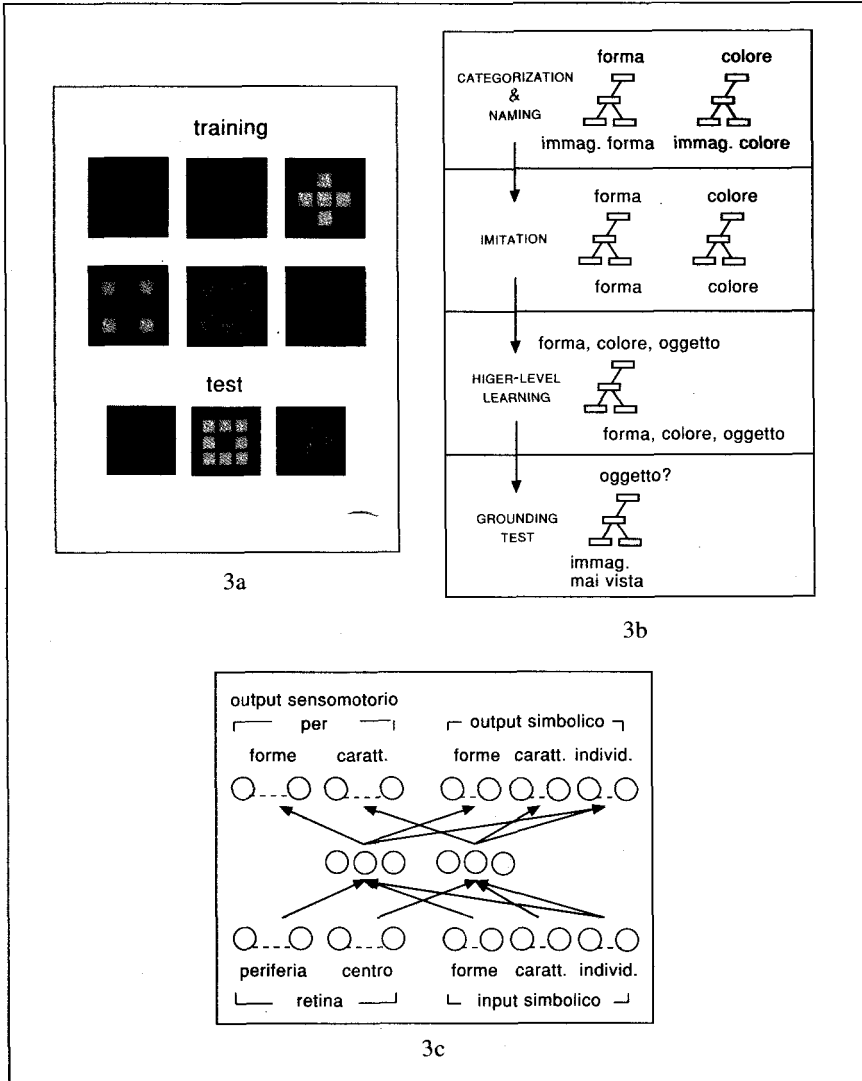


FIG. 3. Simulazione III. 3a: immagini usate per l'addestramento della rete; 3b: procedura di apprendimento e di test; 3c: architettura della rete neurale.

di *training*, le reti hanno fornito in output nell'85,3% dei casi l'etichetta ad esse arbitrariamente associata, mostrando ancora una volta di aver appreso il *grounding transfer*.

#### 5.4. *Discussione*

Nel loro complesso, le simulazioni qui presentate hanno la caratteristica di far uso di simboli il cui significato è fondato in esperienze sensomotorie fatte dal sistema stesso, attraverso un apprendimento diretto per prove ed errori, supervisionato dalle conseguenze di una corretta o errata categorizzazione. Si è mostrato come possano essere appresi simboli di ordine superiore, attraverso un meccanismo, definito «furto simbolico» (Cangelosi, Greco e Harnad 2000), di trasferimento del significato dei simboli di base a categorie più complesse senza più bisogno di un faticoso apprendimento diretto. Ciò è quanto avviene normalmente attraverso l'uso simbolico del linguaggio, al di là degli aspetti di coordinazione di schemi sensomotori che anch'esso possiede.

Questi modelli presentano inoltre spontaneamente il fenomeno della percezione categoriale, per cui si verificano effetti di compressione ed espansione all'interno dello spazio di attivazione delle unità nascoste. Analizzando le attivazioni si può notare che le rappresentazioni di stimoli appartenenti alla stessa categoria diventano più simili durante l'apprendimento, mentre quelle di categorie diverse diventano più dissimili. Questo fenomeno, che come abbiamo visto era presente anche nelle prime simulazioni di Harnad e coll. (Harnad, Hanson e Lubin 1992; Harnad 1993), si verifica a causa del fatto che le reti, per identificare gli oggetti presentati, estraggono le caratteristiche invarianti dagli stimoli, producendo in questo modo delle rappresentazioni astratte che vengono definite rappresentazioni categoriali (Harnad 1987). Queste rappresentazioni preservano alcune parti della «forma» delle proiezioni sensoriali che però vengono trasformate per permettere la categorizzazione. In particolare la similarità all'interno della categoria viene compressa mentre quella tra categorie viene espansa in modo da istituire un confine tra categorie che garantisce una corretta classificazione. Come abbiamo visto, la categorizzazione umana esibisce ampiamente l'effetto della rappresentazione categoriale e ciò rivela la plausibilità psicologica di questi modelli.

I nostri risultati confermano che questi fenomeni possono essere simulati all'interno di un'architettura interamente connectionista. Nei nostri modelli le reti apprendono in modo incrementale: affrontano in ognuna delle tre fasi compiti sempre più difficili sfruttando le rappresentazioni acquisite nella fase precedente. Per realizzare ciò usiamo reti modulari in cui le rappresentazioni per diverse chiavi di classificazione sono localizzate in zone diverse permettendo così la ricombinazione di costituenti elementari discreti in altri di più alto livello.

Le nostre simulazioni mostrano anche l'importanza della riutilizzabilità delle rappresentazioni per poter affrontare il problema della composizionalità discusso nei paragrafi precedenti. Questi modelli, infatti, danno la possibilità di formare proposizioni formate da diversi



simboli per esprimere significati più complessi; siccome questi simboli sono fondati, il problema della composizionalità implica un trasferimento di fondamento dai simboli definienti ai simboli che vengono definiti. Per fare ciò abbiamo bisogno di rappresentazioni riutilizzabili in compiti diversi e questo è uno dei problemi che rimangono aperti con l'uso di reti neurali (Clark e Karmiloff-Smith 1993).

## 6. CONCLUSIONE

Il presente lavoro ha discusso il problema del fondamento dei simboli in modelli cognitivi di rilevanza psicologica delineando una possibile soluzione mediante l'uso di reti neurali. Abbiamo messo in evidenza l'importanza del fatto che i simboli utilizzati da sistemi cognitivi costruiti o usati con finalità simulate, cioè di spiegazione di sistemi naturali a cui corrispondono per analogia, siano collegati a esperienze dirette dei sistemi stessi. Infatti le variabili che forniscono valore esplicativo ai modelli sono per lo più simboliche, come quelle che fanno riferimento a *rappresentazioni* interne, ma perdono tale valore se i simboli che le costituiscono fanno riferimento a stati mentali di chi usa il modello e non sono intrinseci al modello stesso. Abbiamo anche sostenuto che le esperienze dirette che danno fondamento ai simboli dovrebbero essere considerate rappresentazioni di natura analogica derivanti direttamente dall'interazione sensomotoria del sistema con l'ambiente. Abbiamo quindi preso in considerazione l'approccio connessionista per la costruzione di modelli «fondati», in quanto l'uso di reti neurali consente una simulazione più realistica del passaggio dagli aspetti analogici a quelli simbolici. Abbiamo sottolineato, in particolare, il fatto che in qualunque sistema – anche totalmente distribuito – esistono sempre parti o aspetti interpretati simbolicamente in modo da esprimere le analogie con il sistema simulato e da dare allo stesso modello valore esplicativo. Da questo punto di vista la scelta di un'architettura completamente distribuita o ibrida (simbolica-non simbolica) sembrerebbe non avere molta importanza se non dal punto di vista della convenienza della strategia di implementazione; tuttavia esistono vantaggi dal punto di vista teorico che rendono preferibile adottare un approccio puramente connessionista, come abbiamo fatto nei modelli qui presentati.

Nelle nostre simulazioni una stessa rete neurale viene utilizzata per gestire diverse abilità cognitive come quelle percettive, di categorizzazione e linguistico-simboliche. Questo approccio presenta il vantaggio di permettere l'integrazione diretta tra la capacità di interagire con il mondo esterno attraverso la percezione e categorizzazione e quella di acquisire e usare simboli. La plausibilità psicologica di questo approccio è confermata dall'ampia letteratura che considera l'interazione tra

aspetti espliciti e impliciti nell'apprendimento e in altri processi psicologici (cfr. ad es. Berry 1997; Goschke 1997; Shanks 1997). Inoltre, questa prospettiva consente di analizzare meglio le relazioni tra gli aspetti tradizionalmente definiti di basso e alto livello o tra le ottiche *bottom-up* e *top-down* (ad esempio dal riconoscimento di configurazioni al ragionamento e viceversa).

Nei modelli che abbiamo presentato, la stessa rete è in grado di utilizzare due diversi metodi per il fondamento dei simboli. Il primo, detto di «fondamento diretto», è basato su un'esperienza per prove ed errori e di apprendimento tramite *feedback* diretto durante l'interazione con gli stimoli del mondo esterno. Tale metodo viene utilizzato per imparare gli stimoli di base su cui verrà successivamente basato tutto il sistema linguistico. Il secondo metodo, detto di «trasferimento del fondamento», è basato sull'acquisizione di nuove conoscenze sul mondo (ad es. categorie di livello superiore) attraverso la combinazione di simboli di base precedentemente fondati nell'esperienza diretta con il mondo.

Oltre a indicare un nuovo approccio al problema generale del fondamento dei simboli, questo lavoro ha anche presentato delle nuove simulazioni con risultati importanti per l'uso di reti neurali in compiti di apprendimento e manipolazione di simboli. In particolare, le simulazioni con reti ad architettura modulare mostrano che il sistema cognitivo ha bisogno di differenziare tra la parte del modello dedicata al fondamento diretto degli stimoli e quella specializzata per la successiva combinazione di simboli di nuove categorie e il trasferimento del fondamento dai simboli di base a quelli utilizzati per descrivere nuove conoscenze. La separazione di tali funzioni in diversi gruppi di unità nascoste e connessioni è stata necessaria per permettere al modello di acquisire numerose categorie di basso e alto livello, superando i limiti dell'architettura totalmente connessa utilizzata nelle simulazioni di Cangelosi, Greco e Harnad (2000).

I modelli connessionisti del tipo che abbiamo discusso in questo articolo possono anche essere combinati con i modelli robotici (ad es. Steels 2002; Vogt 2002) in modo da permettere il fondamento dei simboli attraverso un'interazione sensomotoria con il mondo esterno. Usando entrambi gli approcci, connessionista e robotico, si possono sviluppare complessi sistemi cognitivi intelligenti, psicologicamente plausibili e immuni dal problema del fondamento dei simboli. Per esempio Marocco, Cangelosi e Nolfi (2003) utilizzano una rete neurale per controllare il comportamento sensomotorio e cognitivo-linguistico di robot evolutivi, mentre Cangelosi (1998) utilizza il metodo delle reti neurali ecologiche con agenti simulati. Questo tipo di modelli permette di beneficiare sia degli aspetti positivi dell'utilizzo delle reti neurali nel fondamento dei simboli, come dimostrato in questo lavoro, che dei vantaggi nell'adoperare dei sistemi cognitivi basati su robot, e più in generale su sistemi

detti di *embodied-situated cognition* (Sharkey e Ziemke 2000; Pfeifer e Scheier 1999; per una discussione vedi Anderson 2003).

Le aree di ricerca e di sviluppo di modelli cognitivi linguistici che utilizzano simboli fondati sull'esperienza diretta sono molteplici. Per esempio, sono necessari ulteriori studi e simulazioni per incrementare la complessità dell'ambiente di interazione e il numero e tipo di stimoli in esso presenti. A questo deve anche corrispondere una certa complessità del sistema cognitivo stesso, perché sia in grado di apprendere e costruire sistemi gerarchici di categorie e di simboli ad esse associati. Questo si è finora rivelato uno dei limiti degli attuali sistemi robotici e di agenti simulati, nei quali alla complessità della gestione delle abilità comportamentali del robot bisogna aggiungere quella del suo sistema cognitivo e linguistico.

Per quanto riguarda, appunto, il sistema simbolico-linguistico del modello, è importante sviluppare dei metodi di manipolazione di simboli prettamente basati sul linguaggio umano. Questo può essere ottenuto attraverso reti neurali in grado di utilizzare strutture sintattiche per la combinazione di simboli di base e la generazione di nuove conoscenze sul mondo attraverso il meccanismo del trasferimento del fondamento. Le simulazioni qui presentate combinano i nomi di due categorie esistenti con quello della nuova categoria di ordine superiore (es. cavallo, strisce, zebra) senza utilizzare un vero sistema sintattico. In modelli futuri sarà necessario che siano utilizzate delle reti con vera capacità combinatoria, per esempio attraverso i meccanismi di emergenza di strutture sintattiche, come nei modelli di evoluzione di nomi e verbi (Cangelosi 2001; Cangelosi e Parisi 2001).

Occorre sottolineare il fatto che le capacità linguistiche dei modelli simbolicamente «fondati» non rivestono un interesse di pura esibizione di capacità di interazione robotica ma sono importanti anche dal punto di vista del loro valore esplicativo, perché il modo più ovvio per attribuire un'etichetta non arbitraria alle variabili interne di un modello è che sia lo stesso modello a suggerirla attraverso un'operazione squisitamente linguistica. Da questo punto di vista, la prassi attuale di esaminare gli stati interni (pesi delle connessioni) di una rete, ad esempio attraverso metodologie di *clustering*, dovrebbe essere considerata un procedimento provvisorio in attesa di una generazione di modelli che possano imparare a comunicare in termini simbolici, anche schematici ed elementari, le rappresentazioni che si sono costruite autonomamente.

Infine, un'interessante direzione di ricerca per i modelli sul fondamento degli stimoli è quella di guardare alle origini evolutive del simbolo stesso. Per esempio, di recente sono state proposte delle evidenti similarità e parallelismi tra i meccanismi di acquisizione dei simboli e i fenomeni di evoluzione del linguaggio. Nel lavoro di Cangelosi e Harnad (2000), il meccanismo ontogenetico del «fondamento diretto» degli stimoli corrisponde a quello del *sensorimotor toil* per l'evoluzione

delle prime parole, mentre il meccanismo del «trasferimento dei simboli» corrisponde a quello evolutivo del «furto simbolico». Lo sviluppo di modelli cognitivi complessi in grado di simulare la ricchezza del linguaggio umano permetterà anche la comprensione dei meccanismi di origine ed evoluzione delle nostre abilità simboliche e linguistiche.

#### RIFERIMENTI BIBLIOGRAFICI

- ANDERSON M.L. (2003), *Embodied Cognition: A Field Guide*, in «Artificial Intelligence», 149, pp. 91-130.
- BERLIN B. e KAY P. (1969), *Basic Color Terms: Their Universality and Evolution*, Berkeley, University of California Press.
- BERRY D.C. (1997), *Introduction*, in D.C. Berry (a cura di), *How Implicit is Implicit Learning?*, Oxford, Oxford University Press, pp. 1-12.
- BROOKS R. (1999), *Cambrian Intelligence: The Early History of the New AI*, Cambridge, Mass, The MIT Press.
- CANGELOSI A. (1998), *Simulazione ed evoluzione del linguaggio in un contesto ecologico*, in «Sistemi Intelligenti», X, 3, pp. 409-442.
- CANGELOSI A. (2001), *Evolution of Communication and Language Using Signals, Symbols and Words*, in «IEEE Transactions on Evolutionary Computation», 5, 2, pp. 93-101.
- CANGELOSI A., GRECO A. e HARNAD S. (2000), *From Robotic Toil to Symbolic Theft: Grounding Transfer from Entry-Level to Higher-Level Categories*, in «Connection Science», 12, 2, pp. 143-162.
- CANGELOSI A. e HARNAD S. (2000), *The Adaptive Advantage of Symbolic Theft over Sensorimotor Toil: Grounding Language in Perceptual Categories*, in «Evolution of Communication», 4, 1, pp. 117-142.
- CANGELOSI A. e PARISI D. (2001), *How Nouns and Verbs Differentially Affect the Behavior of Artificial Organisms*, in J.D. Moore e K. Stenning (a cura di), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, London, Lawrence Erlbaum Associates, pp. 170-175.
- CANGELOSI A. e PARISI D. (2002), *Computer Simulation: A New Scientific Approach to the Study of Language Evolution*, in A. Cangelosi e D. Parisi (a cura di), *Simulating the Evolution of Language*, London, Springer-Verlag, pp. 3-28.
- CLARK A. e KARMILOFF-SMITH A. (1993), *The Cognizer's Innards: A Psychological and Philosophical Perspective on the Development of Thought*, in «Mind & Language», 8, pp. 487-568.
- FODOR J. e PYLYSHYN Z. (1988), *Connectionism and Cognitive Architecture: A Critical Analysis*, in «Cognition», 28, pp. 3-71.
- GOSCHKE T. (1997), *Implicit Learning and Unconscious Knowledge: Mental Representation, Computational Mechanisms, and Brain Structures*, in K. Lamberts e D. Shanks (a cura di), *Knowledge, Concepts, and Categories*, Hove, UK, Psychology Press, pp. 247-333.
- GRECO A. (1988), *Introduzione alla simulazione come metodologia di ricerca in psicologia*, Milano, Franco Angeli.

- GRECO A. (1995), *The Concept of Representation in Psychology*, in A. Greco (a cura di), *A Special Issue on «Representation»*, in «Cognitive Systems», 4-2, pp. 247-256.
- GRECO A. (1997), *Osservazioni su «Scienza cognitiva oggi» di Domenico Parisi*, in «Giornale Italiano di Psicologia», XXIV, 3, pp. 641-647.
- GRECO A. e CANGELOSI A. (1999), *Language and the Acquisition of Implicit and Explicit Knowledge: A Pilot Study Using Neural Networks*, in «Cognitive Systems», 5-2, pp. 148-165.
- HARNAD S. (1990), *The Symbol Grounding Problem*, in «Physica D», 42, pp. 335-346.
- HARNAD S. (1993), *Grounding Symbols in the Analog World with Neural Nets. A Hybrid Model*, in «Think», 2, 1, pp. 12-78 (*Special Issue on «Connectivism versus Symbolism»*, a cura di D.M.W. Powers e P.A. Flach).
- HARNAD S. (2002), *Symbol Grounding and the Origin of Language*, in M. Scheutz (a cura di), *Computationalism: New Directions*, Cambridge, Mass., The MIT Press, pp. 143-158.
- HARNAD S., HANSON S.J. e LUBIN J. (1991), *Categorical Perception and the Evolution of Supervised Learning in Neural Nets*, in D.W. Powers e L. Reeker (a cura di), *Working Papers of the AAAI Spring Symposium on Machine Learning of Natural Language and Ontology*, pp. 65-74.
- JOHNSON-LAIRD P.N. (1983), *Mental Models*, Cambridge, Cambridge University Press, trad. it. *Modelli mentali*, Bologna, Il Mulino, 1988.
- LIBERMAN A.M., HARRIS K.S., HOFFMAN H.S. e GRIFFITH B.C. (1957), *The Discrimination of Speech Sounds within and across Phoneme Boundaries*, in «Journal of Experimental Psychology», 54, pp. 358-368.
- MAROCCO D., CANGELOSI A. e NOLFI S. (2003), *The Emergence of Communication in Evolutionary Robots*, in «Philosophical Transactions of the Royal Society - A», 361, pp. 2397-2421.
- NAKISA R.C. e PLUNKETT K. (1998), *Evolution of a Rapidly Learned Representation for Speech*, in «Language and Cognitive Processes», 13, pp. 105-127.
- PALMER S. (1978), *Fundamental Aspects of Cognitive Representation*, in E. Rosch e B. Lloyd (a cura di), *Cognition and Categorization*, Hillsdale, NJ, Lawrence Erlbaum, pp. 259-302.
- PARISI D. (2001), *La simulazione*, Bologna, Il Mulino.
- PEIRCE C.S. (1978), *Collected Papers*, vol. II: *Elements of Logic*, in C. Hartshorne e P. Weiss (a cura di), Cambridge, Mass., Belknap.
- PFEIFER R. e SCHEIER C. (1999), *Understanding Intelligence*, Cambridge, Mass., The MIT Press.
- PYLYSHYN Z.W. (1981), *The Imagery Debate: Analog Media versus Tacit Knowledge*, in N. Block (a cura di), *Imagery*, Cambridge, Mass., The MIT Press, pp. 151-206.
- SEARLE J.R. (1980), *Minds, Brains, and Programs*, in «Behavioral and Brain Sciences», 3, pp. 417-424, trad. it. *Menti, cervelli e programmi*, Milano, Clup-Clued, 1984.
- SHANKS D.R. (1997), *Distributed Representations and Implicit Knowledge: A Brief Introduction*, in K. Lamberts e D. Shanks (a cura di), *Knowledge, Concepts, and Categories*, Hove, UK, Psychology Press, pp. 197-214.

- SHARKEY N. e ZIEMKE T. (2000), *Life, Mind and Robots: The Ins and Outs of Embodiment*, in S. Wermter e R. Sun (a cura di), *Hybrid Neural Systems*, Heidelberg, Springer Verlag, pp. 313-332.
- SMOLENSKY P. (1988), *On the Proper Treatment of Connectionism*, in «Behavioral and Brain Sciences», 11, trad. it. *Il connessionismo tra simboli e neuroni*, a cura di M. Frixione, Genova, Marietti, 1992.
- STEELS L. (2002), *Grounding Symbols through Evolutionary Language Games*, in A. Cangelosi e D. Parisi (a cura di), *Simulating the Evolution of Language*, London, Springer-Verlag, pp. 211-226.
- VOGT P. (2002), *The Physical Symbol Grounding Problem*, in «Cognitive Systems Research», 3, 3, pp. 429-457.
- WINOGRAD T. (1972), *Understanding Natural Language*, New York, Academic Press.

*Alberto Greco, Laboratorio di Psicologia e Scienze Cognitive, Dipartimento di Scienze Antropologiche, Università di Genova, Via Balbi 4, 16126 Genova. E-mail: greco@nous.unige.it*

*Angelo Cangelosi, School of Computing, Communication and Electronics, University of Plymouth, Plymouth PL4 8AA (UK). E-mail: acangelosi@plymouth.ac.uk*

*Thomas Riga, Salita Pollaiuoli 3/3A, 16123 Genova. E-mail: thomasriga@yahoo.com*