

ARCHITETTURE TELEOLOGICHE APERTE AGENTI IN GRADO DI DETERMINARE AUTONOMAMENTE I PROPRI OBIETTIVI

Riccardo Manzotti e Marco Alessandro Villamira

*KTEL (Knowledge and Thinking Engineering Lab), Istituto di Scienze dell'uomo e
dell'ambiente, Università IULM, Via Carlo Bo, 8, 20143 Milano*

riccardo.manzotti@iulm.it, marco.villamira@iulm.it

Il “what problem”: sistemi teleologici aperti e sistemi teleologici chiusi

Fino a oggi la maggior parte degli agenti artificiali è stata realizzata cercando di creare sistemi in grado di risolvere specifici problemi noti a priori. Per questo motivo il massimo sforzo dei progettisti è stato diretto verso la selezione di un insieme di regole per raggiungere obiettivi fissati. Sistemi del genere sono teleologici chiusi in quanto l'insieme dei loro obiettivi (cioè del loro *telos*) è fissato a priori.

I progettisti cercano di creare agenti in grado di individuare come (“how”) risolvere determinati problemi. Secondo molti studiosi di scienze cognitive «L'apprendimento è definibile come una modificazione del comportamento: una modificazione determinata da un obiettivo (goal)» (Newell 1990): ma da dove derivi questo obiettivo da raggiungere è raramente specificato. Nel campo delle reti neurali i paradigmi dominanti sono quelli dell'apprendimento supervisionato o non supervisionato (eventualmente apprendimento per rinforzo). Per esempio, Sutton e Barto, due dei massimi esperti di apprendimento in sistemi artificiali, hanno scritto che «l'apprendimento per rinforzo è un apprendimento su come mappare situazione e azioni in modo da massimizzare un *obiettivo dato*» (Sutton e Barto 1998), p.3. E hanno aggiunto che «l'idea fondamentale [dell'apprendimento] è catturare gli aspetti più importanti di un problema reale nei termini di un obiettivo, dato un certo ambiente e una tipologia di agente [...]. Tutti gli agenti basati sull'apprendimento per rinforzo hanno *goal prefissati*» (Sutton e Barto 1998), p.4-5. In altre parole, secondo questi Autori, che rappresentano l'opinione oggi prevalente nel loro settore, l'apprendimento per rinforzo è pensato per situazioni nelle quali l'agente cerca di trovare “come” ottenere un certo obiettivo: l'obiettivo o gli obiettivi sono fissati a priori dal progettista.

Nei sistemi artificiali del tipo behavior-based, gli obiettivi sono definiti a priori dai progettisti (McFarland e Bosser 1993; Arkin 1999), ma il loro comportamento si modifica in risposta all'interazione con l'ambiente.

D'altra parte, osservando gli organismi biologici e i soggetti umani, è evidente come sia presente la capacità di sviluppare non soltanto nuove strategie, ma anche nuovi obiettivi. I sistemi biologici sono in grado di determinare non soltanto il “come”, ma anche il “che cosa” deve essere ottenuto (Villamira e Manzotti 2004). Nel momento in cui l'obiettivo è impossibile da conoscere al momento del progetto, è necessario che l'agente sia in grado, non soltanto di trovare “come” raggiungere un certo obiettivo, ma anche “che cosa” raggiungere: questo è il “what problem”.

Nel caso degli organismi biologici esistono due situazioni in cui nuovi obiettivi possono essere inglobati negli organismi: la filogenesi e l'ontogenesi (Parisi 1999; Elman, Bates et al. 2001). Nel primo caso i singoli individui non sono sempre in grado di sviluppare nuovi fini, ma nel corso delle generazioni si manifestano nuovi obiettivi che possono tradursi in miglior o peggior adattamento all'ambiente. Nel secondo caso invece ogni singolo individuo è “aperto” a nuovi obiettivi che possono essere individuati. Per esempio, gli esseri umani

sviluppano continuamente nuovi obiettivi diversi da quelli contenuti nel loro patrimonio filogenetico.

Nel caso dei modelli cognitivi artificiali non si è finora riusciti a implementare un livello di plasticità corrispondente al “what problem”. Oggi, una nuova frontiera è rappresentata dalla possibilità, o dal desiderio, di creare macchine (o sistemi artificiali) in grado di individuare autonomamente i propri obiettivi e quindi definire autonomamente la propria ontologia e la propria teleologia.

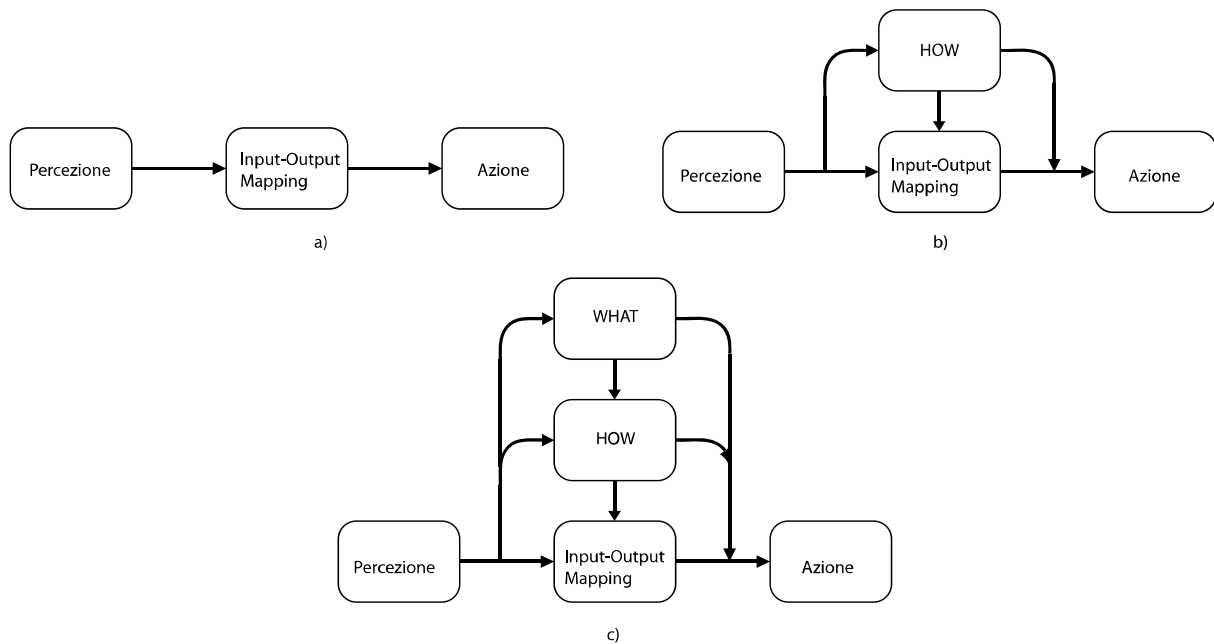


Figura 1. Tre modelli possibili di agente/soggetto corrispondenti a tre gradi distinti di adattabilità all’ambiente. Nel caso a) sia il “what” e il “how” sono definiti a priori; nel caso b) il sistema ha obiettivi fissi, ma è in grado di apprendere “come” raggiungerli; nel caso c) il sistema può apprendere sia “che cosa” cercare di ottenere sia “come” cercare di farlo.

Riassumendo un discorso molto più lungo si potrebbe dire che, attualmente, i sistemi artificiali si dividono in (Figura 1):

- sistemi input-output;
- sistemi in grado di apprendere il “come” realizzare degli obiettivi fissi;
- sistemi in grado di apprendere “che cosa” dover realizzare.

Nel primo caso, l’agente non ha nessuna capacità di modificare i propri scopi e il modo nel quale li persegue. Si tratta di un agente paragonabile a una mappatura stimolo-risposta: una macchina a stati che applica algoritmi prefissati. Non c’è nessuna epigenesi od ontogenesi. Nonostante la complessità comportamentale dell’agente (che potrebbe essere costituito da milioni di istruzioni), tutto avviene perché è stato codificato all’interno della sua struttura. Un artefatto meccanico e un complicato agente software non sono diversi da questo punto di vista: entrambi sono stati pre-programmati sia nel proprio scopo sia nel comportamento. Esempi celebri di questa tipologia sono la pulce artificiale di Tolam (Tolman 1939), le macchine pensanti di Braitenberg (Braitenberg 1984), gli insetti artificiali di Brooks e di Pfeifer (Brooks 1991, 1991; Pfeifer 1999/2001).

Nel caso della seconda tipologia, il sistema è in grado di modificare il modo nel quale persegue uno scopo fissato a priori. Il sistema è in grado di modificare *come* agisce. Il modulo stimolo-risposta è affiancato da un modulo in grado di generare una strategia comportamentale. Il sistema ha un grado maggiore di plasticità comportamentale: è il caso dei sistemi in grado di apprendere. Il modulo “how” modifica le regole comportamentali a priori contenute nel modulo stimolo-risposta. I robot behavior-based rientrano in questa categoria, così come la maggior parte di algoritmi di apprendimento (apprendimento supervisionato o a rinforzo). Questi sistemi determinano come raggiungere un certo risultato sulla base di qualche criterio di ottimizzazione prefissato. Tuttavia questi sistemi sono incapaci di assegnarsi nuovi obiettivi. Alcuni esempi di questi agenti sono il robot sviluppato al LIRA-Lab di Genova (Metta, Sandini et al. 1999; Manzotti, Gasteratos et al. 2001) e Cog dell’MIT (Ferrell e Kemp 1996; Brooks, Breazeal et al. 1999), i software di riconoscimento vocale, gli agenti dotati di reti neurali.

Siamo così giunti al terzo e ultimo caso: sistemi in grado di essere teleologici plastici. Sappiamo che gli organismi biologici hanno questa capacità sia durante la filogenesi che durante l’ontogenesi. E’ possibile implementare questa capacità in un agente artificiale?

Nel seguito descriveremo un’architettura in grado di generare autonomamente i propri obiettivi e che, quindi, dovrebbe poter rientrare nella terza tipologia. Una architettura di questo tipo sarebbe *teleologicamente aperta* nel senso che non avrebbe scopi fissati, ma li genererebbe attraverso l’interazione con l’ambiente. Un’agente del genere avrebbe un *umwelt* (Uexküll 1934) variabile in grado di allargarsi secondo l’accezione proposta da Ziemke ed Emmeche (Emmeche 2001; Ziemke 2001).

Architetture teleologiche aperte

L’idea di una architettura teleologica aperta richiama alcune intuizioni della cibernetica. Quando, per la prima volta, vennero studiati e analizzati i sistemi dotati di feedback, si intuì che essi erano dotati della capacità di avere uno scopo, un fine, e, proprio per questo motivo. Tuttavia la teleologia di questi sistemi era rigida e pre-programmata: non era il frutto dell’interazione con l’ambiente, il risultato di un legame tra la loro esperienza e il loro sviluppo. Ogni sistema cibernetico classico “nasceva” con un certo numero di obiettivi, di fini, da perseguire e per tutta la propria esistenza non si modificava. I sistemi cibernetici classici, dal punto di vista degli obiettivi da perseguire, erano sistemi chiusi. Al contrario, l’architettura che stiamo per descrivere, pur nella sua semplicità, è in grado di darsi nuove obiettivi, è in grado di trasformare la sua esperienza in nuovi obiettivi: il sistema proposto è un sistema teleologico “aperto” che potremmo definire neocibernetico.

L’architettura è costituita da tre moduli (Figura 2): un modulo filogenetico in grado di attivare (*bootstrap*) il sistema; un modulo in grado di immagazzinare nuove categorie con funzione di memoria e un modulo teleologico capace di determinare nuovi obiettivi per il sistema sulla base degli eventi memorizzati nella memoria (una versione di questo agente è stata descritta in (Manzotti e Tagliasco 2005)).

L’idea di base è quella di realizzare un’architettura in grado di modificare, sulla base della propria esperienza, i criteri secondo i quali fa esperienza del mondo esterno. Per esempio, un sistema siffatto diventerà capace di riconoscere visi, poiché i visi hanno fatto parte della sua esperienza sensoriale nel passato. Questo sviluppo ciclico di strutture categoriali e di obiettivi è qui considerato un caso di ontogenesi.

Descriviamo adesso i tre moduli fondamentali di cui deve essere costituita una architettura teleologica aperta.

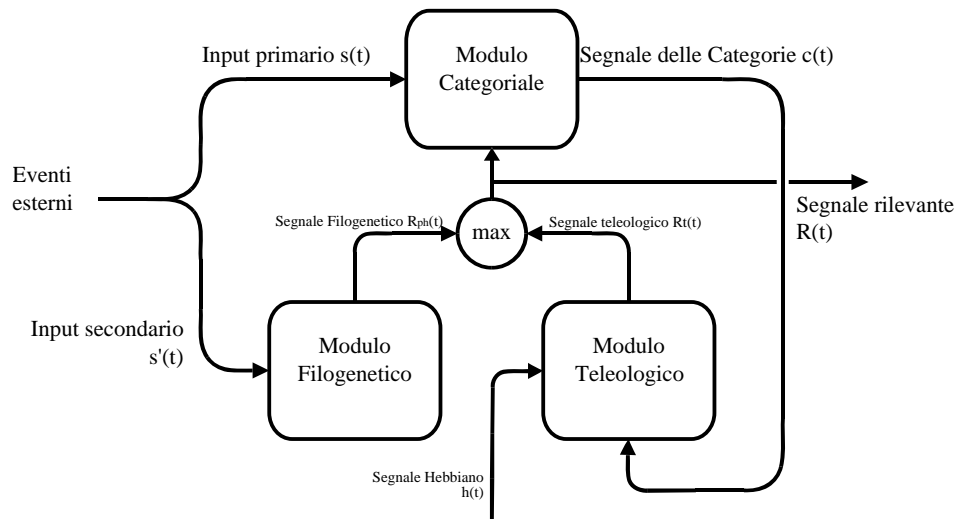


Figura 2: Schema generale di un'architettura teleologica aperta in grado di sviluppare nuovi obiettivi.

Il modulo categoriale

Il modulo categoriale ha due funzioni principali: deve calcolare quanto lo stimolo corrente sia simile alle categorie memorizzate e deve raggruppare gli stimoli in ingresso in categorie.

La prima funzione è implementata calcolando un opportuno vettore delle categorie. Ogni qual volta uno stimolo è ricevuto, per ognuna delle categorie memorizzate, si calcola la distanza corrispondente. L'insieme di queste distanze costituisce il vettore delle categorie che, in pratica, è un istogramma che mostra quale categoria è più simile allo stimolo in ingresso. La dimensione di questo vettore è pertanto variabile con il variare del numero delle categorie memorizzate.

La seconda funzione del modulo delle categorie dipende dal segnale di Sviluppo, che è calcolato come il massimo tra il segnale teleologico e il segnale filogenetico. Sia il segnale teleologico che quello filogenetico sono calcolati sulla base dello stimolo in ingresso. Il segnale filogenetico è calcolato direttamente sulla base dello stimolo in ingresso facendo uso delle categorie innate. Il segnale teleologico è calcolato facendo uso del vettore delle categorie. Dopo che il vettore delle categorie è calcolato, *se e solo se* il segnale di sviluppo è attivo, il modulo delle categorie esegue una delle seguenti azioni facendo uso dello stimolo in ingresso opportunamente memorizzato in un buffer temporaneo: 1) se lo stimolo è troppo simili agli stimoli già memorizzati non fare nulla; 2) se lo stimolo è sufficientemente simili a una delle categorie già memorizzate, lo stimolo è aggiunto a quella categoria estendendone il campo di applicazione; 3) se lo stimolo non è abbastanza simile a nessuno degli stimoli già memorizzati, una nuova categoria è creata.

In questo modo, il sistema non assegna nuove risorse per ogni stimolo in ingresso ma soltanto per quelli che corrispondono ai criteri innati o a quelli acquisiti.

Il modulo filogenetico

Questo modulo è l'unico che ha dei criteri pre-determinati. Da un punto di vista funzionale, ha lo stesso ruolo degli istinti genetici nei sistemi biologici. Per esempio, un bambino guarda con molta più curiosità e interesse a oggetti dai colori brillanti che non a oggetti grigi e privi di colore. Questo comportamento richiede l'esistenza di una funzione pre-determinata che individui delle proprietà rilevanti (per esempio la presenza di colori saturi e brillanti) negli stimoli in ingresso. Questo modulo genera un segnale filogenetico che informa il modulo

categoriale del fatto che lo stimolo in ingresso sia importante dal punto di vista dei criteri innati. Se l'uscita del modulo teleologico è assente o nulla, come accade all'inizio della vita del sistema, l'uscita del modulo filogenetico è sufficiente a costituire il segnale di sviluppo. Se il sistema fosse composto solo del modulo filogenetico e del modulo categoriale il sistema sarebbe equivalente a un sistema basato sull'apprendimento per rinforzo.

Il modulo teleologico

Mentre il modulo filogenetico ha criteri innati per quanto riguarda gli stimoli interessanti, il modulo teleologico seleziona nuovi criteri sulla base dell'esperienza per questo consente l'ontogenesi del sistema (in una versione alternativa dell'esperimento allo stesso modulo era stato dato il nome di modulo ontogenetico o epigenetico (Manzotti e Tagliasco 2005)). Da un punto di vista funzionale, ha lo stesso ruolo delle motivazioni acquisite nei sistemi biologici. L'obiettivo principale del modulo teleologico è quello di trasformare alcune categorie in altrettanti criteri. Non tutte le categorie costruite dal modulo categoriale dovranno diventare criteri.

Il segnale teleologico fornisce una misura di quanto lo stimolo in ingresso sia parte della storia del sistema. Il modo con il quale la storia del sistema è tenuta in considerazione dipende dalle strategie adottate per simulare l'ontogenesi del sistema. Questo sviluppo potrebbe seguire quattro diverse strategie: i) filogenesi; ii) una specie di imprinting; iii) l'occorrenza di particolari eventi esterni e iv) la semplice ripetizione degli stimoli in ingresso.

Queste strategie di sviluppo sono implementate utilizzando dei segnali Hebbiani in quanto utilizzano dei meccanismi di apprendimento Hebbiani. A seconda degli obiettivi della macchina, si possono scegliere diverse strategie di sviluppo. Il modulo teleologico individua la correlazione tra lo stimolo delle categorie corrente e il segnale Hebbiano corrispondente alla strategia voluta. Il valore risultante promuove una particolare categoria a criterio teleologico. Il modulo teleologico indica quanto lo stimolo corrente è parte della storia progressa del sistema.

L'obiettivo principale dell'architettura è quello di creare una struttura che possa essere completamente modificata dalle sue esperienze individuali. Grazie a questa architettura, ogni qual volta un evento è in grado di essere riconosciuto dal modulo categoriale e poi selezionati dal modulo teleologico, diventa parte della storia ontogenetica del sistema in corso di sviluppo. L'evento è responsabile dell'occorrenza di un processo e del fatto che quel processo possa ripetersi. Ogni qual volta un processo diventa parte della struttura del sistema, la sua probabilità di ripetersi è aumentata. Gli eventi che diventano il contenuto del sistema diventano le motivazioni del sistema. Sono processi che si intrecciano con la storia del sistema: si perpetuano grazie al modo in cui il sistema si lascia modificare.

I processi creati dall'esistenza del sistema sono logicamente e casualmente continui con l'ambiente. Inoltre essi modellano e modificano ciò che l'ambiente è. Essi creano nuovi oggetti e un nuovo ambiente modificando le condizioni attraverso le quali l'ambiente può determinare degli effetti.

L'architettura ha solo un obiettivo: se un processo accade e "passa attraverso" il sistema, la probabilità che quel processo accada nuovamente aumenta considerevolmente.

Risultati sperimentali

Per verificare l'efficacia dell'architettura abbiamo implementato i moduli descritti in un sistema robotico di grande semplicità (Figura 3). Allo scopo un agente che contiene l'architettura descritta deve sviluppare una nuova motivazione sulla base della sua esperienza. Nell'esperimento noi abbiamo simulato una testa robotica che guardava a diversi stimoli in un periodo di tempo fissato e abbiamo poi misurato il tempo trascorso a fissare ciascuno degli stimoli.

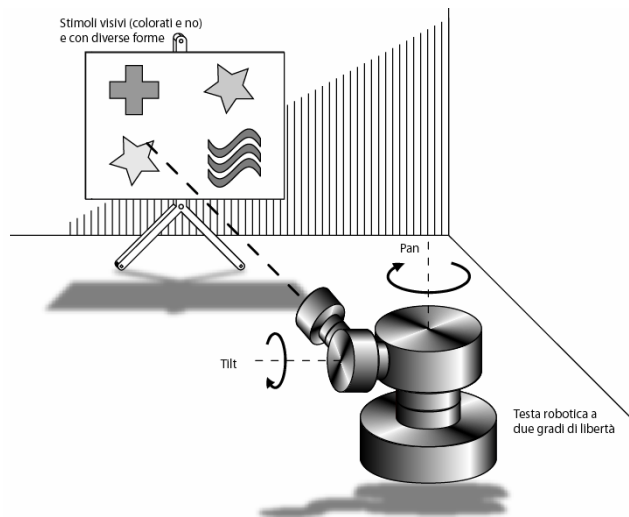


Figura 3 L'agente utilizzato era semplice: simulava una testa robotizzata con due gradi di libertà e telecamere logpolari in grado di fissare stimoli diversi.

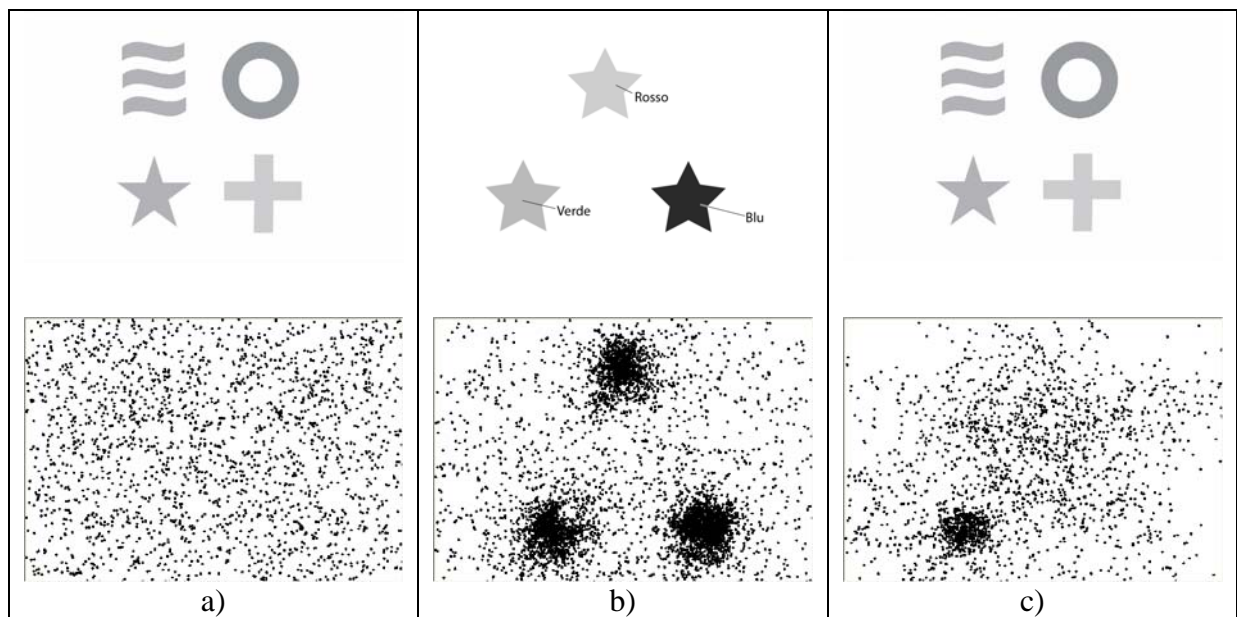


Figura 4. Tre fasi dell'esperimento. In una prima fase il robot vede oggetti che non determinano nessuna risposta da parte del modulo di *bootstrap* (per esempio oggetti bianchi). In una seconda fase il sistema fa esperienza di oggetti colorati con certe forme. Nella terza fase il sistema è reattivo a certe forme anche in assenza delle proprietà che stimolavano il modulo di *bootstrap* (i colori).

Nell'esperimento, una classe di stimoli in ingresso (non codificati a priori nell'architettura del sistema) è in grado di produrre una modifica nel comportamento del sistema che cambia non soltanto *come* (il comportamento) il sistema agisce ma anche *che cosa* il sistema cerca di ottenere (le motivazioni). Qualcosa, che accade nell'ambiente (per esempio la comparsa di una certa classe di stimoli) diventa costitutivo del comportamento dell'agente. Il sistema ha, in questo esperimento preliminare, un solo comportamento possibile: guardare o non guardare altri oggetti.

Un campo visivo simulato fornisce l'ambiente nel quale l'architettura può agire; il sistema utilizza telecamere con immagini logpolare controllato con due gradi di libertà (tilt e pan). Lo stimolo in ingresso è costituito da immagini log polari a colori acquisite da una telecamera log polare simulata (Sandini, Questa et al. 2000).

L'utilizza di una telecamera dotata di fovea (come nel caso delle telecamere log polari) è importante per la creazione di un soggetto artificiale. Il motivo è che le immagini con fovea incorporano l'idea di una direzione precisa dello sguardo del soggetto (una specie di direzionalità implicita dello sguardo). La parte centrale in una immagine log polare, detta fovea, ha una risoluzione spaziale molto più alta e, conseguentemente, è la parte più importante dell'immagine; la parte verso cui è diretto lo sguardo.

Il sistema di controllo motorio è programmato per fare movimenti saccadici casuali. Il modulo motorio genera movimenti saccadici sulla base di un parametro di controllo λ che controlla la densità di probabilità dell'ampiezza del movimento saccadico. L'angolo è completamente casuale. Se λ è piccolo (vicino a 0) la densità di probabilità è quasi costante e, come risultato, lo sguardo si muoverà in modo tendenzialmente uniforme su tutto il campo visivo. Se λ è grande (vicino a 1) i movimenti piccoli diventano molto più probabili. In questo modo quando il sistema sta guardando qualcosa di interessante (valore del parametro grande), si muoverà di poco e resterà sull'oggetto di interesse. Non appena l'immagine non sarà più interessante il sistema avrà un parametro di controllo piccolo e inizierà a muoversi con grandi movimenti saccadici su tutto il campo visivo.

Abbiamo presentato diverse classi di stimoli visivi al sistema. Un primo insieme consisteva in una serie di figure geometriche prive di colore. Abbiamo misurato il tempo trascorso dal sistema su ciascuna parte dell'immagine. Come si vede nella Figura 4a, il sistema non manifestava nessun interesse particolare per nessuno stimolo. I grandi movimenti saccadici facevano muovere lo sguardo in modo uniforme su tutte le parti dell'immagine. In seguito (Figura 4b) abbiamo mostrato al sistema figure colorate aventi tutte la stessa forma e colorate. Il risultato è stato che il sistema era molto più interessato alle parti del campo visivo che contenevano stimoli colorati. In seguito, le figure grigie iniziali sono state ripresentate al sistema. Come si vede in Figura 4c, la figura (il triangolo) che era stata associata al colore, è diventata interessante a prescindere dal fatto che fosse colorata o meno.

L'esperimento, pur nella sua semplicità, dimostra che una architettura artificiale può essere teleologica aperta, ovvero in grado di generare nuovi obiettivi. La maggior parte dei sistemi artificiali attuali (a differenza dei sistemi biologici o di alcuni sistemi studiati nel contesto della vita artificiale) non ha la capacità di istanziare nuove relazioni causali, ma sono bloccati una configurazione causale pre-determinata. Per esempio, un termometro implementa una relazione casuale fissa tra la temperatura prossimale e l'altezza di una colonna di mercurio. Si può rappresentare la relazione tra la temperatura e l'altezza della colonna nel seguente modo:

$$S_{temp} \rightarrow R_{mercury}$$

dove S e R rappresentano lo stimolo e la risposta suscitata dalla relazione causale. Nel caso dell'architettura appena descritta la struttura causale è diversa: l'evento, che è responsabile per la creazione della relazione causale è anche il contenuto di quella relazione causale. Possiamo rappresentare la nuova situazione in questo modo

$$S_{pattern} \rightarrow (S_{pattern} \rightarrow R)$$

La forma della stella, che non aveva alcuna rilevanza causale per l'agente al suo inizio, è diventata uno degli scopi che ne guidano il comportamento: è stato l'accadere dell'evento stella che ha determinato l'istanziamento di una relazione causale che aveva la stella come oggetto. L'esperienza ha determinato una modifica della struttura teleologica dell'agente.

Riferimenti bibliografici

- R. C. Arkin (1999). *Behavior-Based Robotics*. Cambridge (Mass), MIT Press.
- V. Braitenberg (1984). *Vehicles: Experiments in Synthetic Psychology*. Cambridge (Mass), MIT Press.
- R. A. Brooks (1991). "Intelligence Without Representations." *Artificial Intelligence Journal* 47: 139-159.
- R. A. Brooks (1991). "New Approaches to Robotics." *Science* 253: 1227-1232.
- R. A. Brooks, C. Breazeal, M. Marjanovic, B. Scassellati e M. Williamson (1999). The Cog Project: Building a Humanoid Robot. *Computation for Metaphors, Analogy, and Agents*. Nehaniv. Berlin, Springer-Verlag. 1562.
- J. L. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi e K. Plunkett (2001). *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge (Mass), MIT Press.
- C. Emmeche (2001). "Does a robot have an Umwelt? Reflections on the qualitative biosemiotics of Jakob von Uexküll." *Semiotica* 1(4): 653-693.
- C. Ferrell e C. Kemp (1996). *An Ontogenetic Perspective to Scaling Sensorimotor Intelligence*. 1996 AAAI Fall Symposium on Embodied Cognition and Action.
- R. Manzotti, A. Gasteratos, G. Metta e G. Sandini (2001). "Disparity estimation in log polar images and vergence control." *Computer Vision and Image Understanding* 83: 97-117.
- R. Manzotti e V. Tagliasco (2005). "From "behaviour-based" robots to "motivations-based" robots." *Robotics and Autonomous Systems* In press.
- D. McFarland e T. Bossert (1993). *Intelligent Behavior in Animals and Robots*. Cambridge (Mass), MIT Press.
- G. Metta, G. Sandini e J. Konczak (1999). "A developmental approach to visually guided reaching in artificial systems." *Neural Networks* 12: 1413-1427.
- A. Newell (1990). *Unified Theories of Cognition*. Cambridge (Mass), Harvard University Press.
- D. Parisi (1999). *La mente*. Bologna, Il Mulino.
- R. Pfeifer (1999/2001). *Understanding Intelligence*. Cambridge (Mass), MIT Press.
- G. Sandini, P. Questa, D. Scheffer e A. Mannucci (2000). *A Retina-like CMOS Sensor and its Applications*. SAM-2000, Cambridge, USA, IEEE.
- R. S. Sutton e A. G. Barto (1998). *Reinforcement Learning*. Cambridge (Mass), MIT Press.
- E. C. Tolman (1939). "Prediction of the vicarious trial and error by means of the schematic sowbug." *Psychological Review* 23: 318-336.
- J. v. Uexküll (1934). *A Stroll through the Worlds on Animals and Men: A Picture Book of Invisible Worlds*. New York, International University Press.
- M. A. Villamira e R. Manzotti (2004). *Comunicazione e sistemi*. Milano, FrancoAngeli.
- T. Ziemke (2001). "The Construction of 'Reality' in the Robot: Constructivist Perspectives on Situated Artificial Intelligence and Adaptive Robotics." *Foundations of Science* 6(1-3): 163-233.