

## I PATTERN DI TURING IN RETI CELLULARI: DINAMICHE EMERGENTI IN UN'ARCHITETTURA GENERALE PER LA PERCEZIONE

### 1. INTRODUZIONE

Negli ultimissimi anni, prendendo spunto da numerose evidenze biologiche, ci si è rivolti, con interesse crescente, allo studio e alla progettazione di sistemi cognitivi artificiali (Vernon *et al.* 2008) che possano rendere veramente intelligente l'interazione tra i robot e un ambiente esterno, non strutturato e potenzialmente ostile. In effetti, già nello scorso decennio, la ricerca del campo della percezione per la robotica ha sviluppato un nuovo paradigma, noto come Robotica Orientata al Comportamento (Arkin 1991), nel quale il processo percettivo viene considerato strettamente interconnesso con i bisogni comportamentali dell'agente. In questa nuova prospettiva, il processo cognitivo è visto come un fenomeno complesso ed emergente che fonde una grande quantità di informazioni eterogenee per creare una *rappresentazione interna* compatta ed astratta dell'ambiente esterno, che tiene in considerazione nella sua formazione i bisogni e le motivazioni dell'agente (~~Brooks 1992~~). L'intero processo percettivo è mediato tramite un *stato interno* dipendente dal comportamento, come ipotizzato da Nolfi (2002).

Tale nuova visione dei processi percettivi orientati a sistemi artificiali trova riscontro diretto all'interno degli archetipi propri della teoria dei sistemi complessi, composti da un insieme di elementi dinamici (celle) semplici ma non lineari, connessi tra loro in reti di diversificate topologie. In tali architetture, il ruolo fondamentale delle dinamiche delle variabili di stato di ciascuna cella porta all'emergere di nuove soluzioni per l'intero sistema, mediate dalle condizioni al contorno. Tali soluzioni sono nuove nel senso che non sono riscontrabili nella dinamica delle singole celle. Dal punto di vista biologico, il paradigma per eccellenza dei sistemi dinamici complessi è proprio il cervello, le cui funzioni si basano sulle dinamiche dei singoli neuroni, ma le capacità ad alto livello, quali ad esempio la percezione, sono emergenti e non direttamente riscontrabili nel singolo neurone. È quindi possibile gettare un ponte e formalizzare il processo percettivo all'interno della teoria dei sistemi dinamici complessi.

Sulla base di ciò e delle ultime scoperte in neurobiologia – ad esempio i lavori di W. Freeman (2004) – e del crescente interesse nell’ambito della realizzazione di sistemi cognitivi artificiali abbiamo sviluppato un’architettura generale per implementare il ciclo *sensazione-percezione-azione*. Tale architettura, mutuando dal mondo degli insetti strutture neurali e comportamenti semplici e complessi, è potenzialmente applicabile a varie piattaforme robotiche impegnate nell’adempimento di diverse missioni in ambienti non strutturati. Il processo percettivo elabora un’enorme quantità di informazioni eterogenee per dar vita ad una rappresentazione concisa e astratta dell’ambiente circostante (che noi chiameremo *pattern*) che nello stesso tempo tiene conto delle esigenze e delle motivazioni dell’agente.

I *pattern* del mondo esterno, utilizzati per la selezione dell’azione o del comportamento, sono stati ispirati ai pattern di Turing. Secondo la teoria di Turing (1952), i meccanismi di reazione e diffusione chimica dei morfogeni sono alla base della formazione di pattern in natura (esempio classico è dato dalla spiccata simmetria e regolarità geometrica nei colori e nelle forme dei mantelli degli animali). I pattern di Turing, nel nostro lavoro, vengono usati come stati di equilibrio di sistemi dinamici non lineari. Essi sono attrattivi per particolari insiemi di condizioni dell’ambiente allo scopo di modulare, attraverso un algoritmo di apprendimento con rinforzo, una serie di comportamenti di base paralleli e potenzialmente in competizione l’uno con l’altro. Il *learning* è anche introdotto a livello afferente per modulare i bacini di attrazione dei pattern di Turing in modo da rendere più efficace il processo di classificazione degli eventi sensoriali e formare una rappresentazione interna astratta e flessibile, mediata dall’ambiente e dagli scopi dell’agente. In particolare, la struttura da noi sviluppata, è fondata su vari livelli di controllo costituiti da blocchi funzionali che agiscono sia sullo stesso livello in maniera concorrente, che su livelli gerarchici distinti. Nell’architettura percettiva realizzata, processi percettivi paralleli e gerarchia verticale coesistono insieme consentendo all’agente di mostrare abilità di base così come comportanti complessi. Tra le abilità di base si annoverano comportamenti «geneticamente» prefissati, mentre come comportamenti complessi vanno considerate la capacità di rappresentare «situazioni» del mondo esterno, che l’agente acquisisce tramite il suo apparato sensoriale e processa a livello percettivo, l’ottimizzazione del comportamento in relazione alla missione da compiere e la capacità di pianificare sequenze di azioni per il raggiungimento di specifici obiettivi.

I comportamenti di base, *precognitive behaviors*, sono a tutti gli effetti dei comportamenti basati su riflessi, che in linea di principio non richiedono apprendimento. Prendendo spunto dall’osservazione del comportamento degli insetti, alcuni di questi riflessi che abbiamo preso in considerazione, sono la fonotassi (Webb e Scutt 2000), il riflesso optomotore (Bäohm 1991), la capacità di evitare gli ostacoli sulla base di

sensori di contatto (come le antenne) o, salendo di complessità, stimoli visivi appresi come anticipatori dei precedenti. I riflessi di base agiscono in maniera concorrente e il comportamento finale viene determinato, a partire dalle singole decisioni, ad un livello gerarchico più alto.

Ad un livello superiore vengono inseriti comportamenti, detti *protocognitive behavior*, come, ~~sempre guardando al mondo degli insetti~~, la capacità di tornare al «nido» dopo essere andati in cerca di cibo: per questo vengono utilizzati sia algoritmi di odometria, che altri basati su riferimenti assoluti, tramite i quali l'agente costruisce una mappa dell'ambiente tramite l'apprendimento di *landmark* stabili e successivamente naviga in base ad un confronto tra la mappa memorizzata e le informazioni acquisite da sensori visivi filtrate tramite reti neurali ricorrenti (RNN), come descritto da Arena, Cruse *et al.* (2007). In prossimità della home, algoritmi basati su *optical flow* e riferimenti visivi locali consentono di raggiungere con precisione la posizione desiderata.

Arrivando al *livello di rappresentazione*, il cuore dell'architettura percettiva è costituito da una rete cellulare non lineare (CNN) del secondo ordine. Le CNN sono *array* di semplici circuiti dinamici non lineari (celle) localmente connessi, mutuamente accoppiati ed in grado di influenzarsi in modo locale, utili per processare in maniera analogica una consistente mole di informazioni in tempo reale (Chua e Yang 1988). Le CNN, inizialmente introdotte da Leon O. Chua al fine di operare un'efficiente elaborazione delle immagini, hanno successivamente assunto un ruolo trainante come paradigmi per lo studio e la rappresentazione a basso costo di dinamiche tipiche dei sistemi non lineari complessi, quali i sistemi caotici e talune classi di equazioni differenziali alle derivate parziali discretizzate su lattici spaziali (Manganaro *et al.* 1999). Sulla base di ciò e quanto proposto da Goras e Chua (1995), in relazione alla formalizzazione di CNN di tipo reazione-diffusione (RD-CNN) (Goras e Chua 1995), abbiamo progettato una RD-CNN in modo tale da rappresentare equazioni di reazione-diffusione, formalmente simili a quelle introdotte da Alan Turing per generare i pattern omonimi. Quindi la nostra RD-CNN, sulla base delle informazioni provenienti dagli stimoli sensoriali, è in grado di generare pattern di Turing.

Si noti a tal proposito il parallelismo biologico tra i pattern morfogenetici, che si evolvono come risultato sia delle condizioni strutturali (codice genetico) che delle condizioni al contorno e geometriche (crescita del mantello degli animali), e i pattern astratti che emergono nella nostra RD-CNN, risultato sia della struttura (equazioni e parametri della RD-CNN) che delle condizioni iniziali e/o al contorno, stabilite dai sensori esterni.

Sebbene la nostra sia da considerarsi, in senso stretto un'ipotesi di lavoro, ciò nondimeno esistono dei segnali che alcuni apparati funzionali all'interno del cervello di semplici esseri viventi possano generare dinamiche assimilabili a pattern. Ad esempio il Clione Limacina, un mollusco

acquatico con circa 5.000 neuroni, di cui qualche centinaio solamente ha un ruolo attivo nel controllo della movimentazione, possiede, all'interno dei gangli cerebrali, alcuni neuroni che non emettono *spike*. Questi ultimi, definiti «neuroni di comando» da Grillner in Orlovsky *et al.* (1999) assumono un ruolo attivo nel controllo della velocità dei movimenti del Clione e quindi nel processo comportamentale ~~solo~~ a seguito di stimoli ad alto livello. Tali neuroni riescono a riconfigurare il tipo di locomozione e quindi il comportamento utilizzando quindi dinamiche *non spiking*, assimilabili ai nostri pattern. Ricerche accurate in tale direzione sono tuttora in atto per dare maggiore spessore e riscontro sperimentale a quanto finora solo ipotizzato. Da un altro punto di vista, spesso le ipotesi di lavoro sono indispensabili per poter indagare in ambiti non ancora esplorati, al fine di elaborare nuove metodologie per la realizzazione, in questo caso, di sistemi percettivi artificiali, nell'attesa di poter chiudere il feedback con risultati sperimentali di natura neurobiologica.

Nel nostro modello, i pattern emergenti assumono il ruolo di rappresentazioni astratte delle situazioni dell'ambiente esterno, registrate attraverso i sensori. Un algoritmo di apprendimento non supervisionato rende plastico tale processo a livello afferente permettendo una modulazione dei bacini di attrazione dei pattern di Turing in modo da adeguare la classificazione delle informazioni alle motivazioni dell'agente. In uscita i pattern di Turing vengono associati, tramite una rete di selezione del comportamento, a dei parametri di modulazione dei comportamenti di base che permettono la generazione del comportamento finale come «somma pesata» dei comportamenti di base<sup>1</sup>. I parametri di modulazione sono appresi, per ciascun pattern, tramite un algoritmo di *reinforcement learning*, laddove il segnale di rinforzo è fornito all'agente dal livello motivazionale e rappresenta il livello di soddisfacimento dell'agente in relazione alla situazione in cui si trova e agli obiettivi della sua missione. Dato l'elevato numero di differenti pattern (soluzioni) che persino una RD-CNN di modeste dimensioni è in grado di generare, è possibile immagazzinare varietà comportamentali notevoli e quindi apprendere svariate strategie di controllo dei comportamenti, che legano lo stato dell'ambiente alle esigenze motivazionali del robot, con ciò rispondendo pienamente al significato proprio di percezione orientata all'azione.

In relazione, ad esempio ad un problema di navigazione, i livelli di pianificazione e memoria consentono di memorizzare le coppie rappresentazione-comportamento in catene in una memoria a breve termine e di trasferire le sequenze «vincenti» in una memoria a lungo termine. La presenza di informazioni aggiuntive nell'ambiente, non utilizzate dai

<sup>1</sup> Questa semplice modulazione è intesa a validare il modello complessivo in questa fase iniziale, analizzandone i risultati. Si prevede di ~~rendere tale simulazione, una funzione non lineare~~, sulla base dei risultati sia delle simulazioni che degli esperimenti reali, allo scopo di ottimizzare le capacità complessive del modello percettivo.

comportamenti di base e nella formazione della rappresentazione, può essere utilizzata come una dimensione aggiuntiva della rappresentazione in modo da consentire l'individuazione di sotto-obiettivi che possano servire per rendere più complesse le sequenza di comportamenti.

Nella trattazione seguente i blocchi di pianificazione e memoria non verranno ulteriormente presi in considerazione, in quanto, in una fase iniziale, l'attenzione è stata rivolta alla definizione del cuore della strategia percettiva. I due blocchi suddetti sono stati realizzati con metodologie tradizionali, non richiedendone pertanto ulteriore attenzione in questa sede.

## 2. ARCHITETTURA DI CONTROLLO

In questo lavoro l'attenzione viene focalizzata sull'architettura di controllo riportata in figura 1, con particolare attenzione ai comportamenti di base e il livello di Rappresentazione, il quale racchiude l'essenza del modello poiché costituisce la capacità di astrazione ed è in grado di supportare le funzioni di più alto livello (memoria e pianificazione).

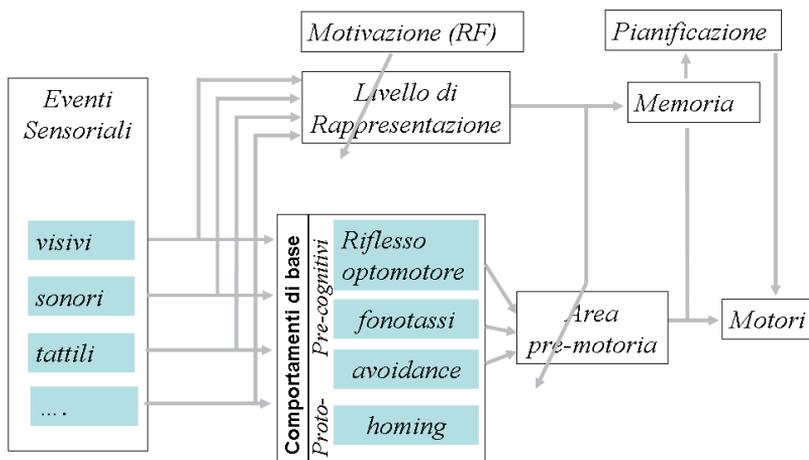


FIG. 1. Schema a blocchi dell'intera architettura percettiva: gli eventi sensoriali attivano i comportamenti di base pre-/proto-cognitivi. Questi ultimi vengono modulati per ottenere il comportamento finale che l'agente attuerà. I parametri di questa modulazione sono determinati dal livello di rappresentazione capace di fondere insieme diversi stimoli sensoriali dando vita ad una rappresentazione astratta dell'ambiente esterno che tiene conto anche della motivazione dell'agente. Tali situazioni vengono conservate in sequenze tramite meccanismi di memorizzazione a breve e a lungo termine, che consentono la pianificazione di catene di comportamenti.

L'implementazione dell'architettura di controllo proposta in questo lavoro è organizzata in livelli di controllo orizzontali concorrenti e in livelli verticali gerarchici.

Il livello più basso è formato dai comportamenti di base «geneticamente» prefissati, i quali, attivati tramite percorsi senso-motori indirizzati da specifici eventi sensoriali, costituiscono il sistema di conoscenze di base dell'agente. Da tali fondamenta è possibile strutturare un comportamento complesso emergente tramite il livello di rappresentazione. Esso è posto ad un livello gerarchico superiore ed agisce come controllo di tipo *feedforward* stabilendo i parametri che generano il comportamento complessivo del robot come funzione delle uscite motorie dei singoli comportamenti di base. L'anello di retroazione è infine chiuso mediante l'interazione fisica tra l'agente e l'ambiente circostante che si realizza da una parte tramite gli attuatori del robot e dall'altra per mezzo dei dati sensoriali acquisiti.

Ad ogni passo dell'agente, il ciclo di controllo prevede la valutazione delle uscite  $A_i(t)$  dei singoli comportamenti di base tramite i relativi ingressi sensoriali  $S_i(t)$ , che al contempo entrano nel livello di rappresentazione che genera di conseguenza un pattern di Turing  $q$ . A quest'ultimo sono associati dei pesi  $K_i^{(q)}$  che modulano i comportamenti di base, dando vita al comportamento finale dell'agente come:

$$A_F(t) = \sum_i K_i^{(q)} \cdot A_i(t)$$

In questo lavoro, l'architettura di controllo è applicata ad un robot dotato di ruote: l'azione finale è realizzata da una rotazione variabile e da una traslazione fissa: i differenti livelli di controllo sono riportati nei paragrafi seguenti.

### 2.1. Blocco sensoriale e comportamenti di base

Per affrontare il problema della navigazione autonoma in un ambiente non strutturato, il robot è fornito di tre sensori di distanza per il rilevamento di ostacoli che coprono il lato sinistro, la parte frontale e il lato destro della sua struttura. Inoltre, il robot riceve informazioni riguardo all'angolo tra il suo orientamento frontale e la direzione del target e, in taluni casi non riportati nel seguito, anche sulla distanza dal target. La figura 2 riporta uno schema esemplificativo dell'intero apparato sensoriale.

In riferimento ai comportamenti di base, ci si è ispirati al mondo degli insetti ed in particolare ai grilli, emulando alcuni dei suoi comportamenti più conosciuti: l'abilità di recuperare la direzione frontale durante il movimento, ovvero il *riflesso optomotore* (Bohm *et. al.* 1991);

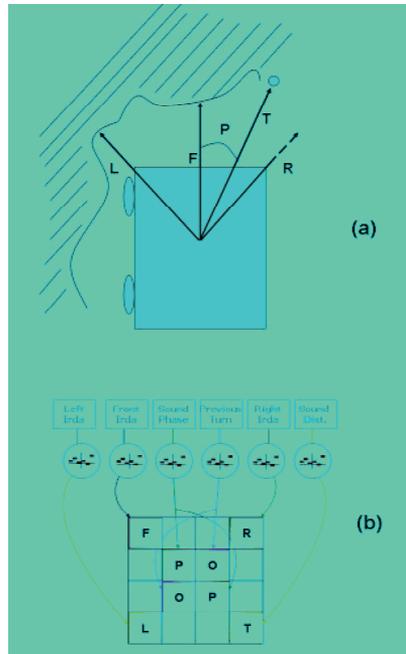


FIG. 2. (a) Il robot è fornito di tre sensori di distanza (F: frontale, L: sinistro, R: destro) e un sensore di target che fornisce informazioni sullo sfasamento (P) e sulla distanza (T) tra il robot e il target. (b) Inizializzazione delle celle del primo layer della CNN: le condizioni iniziali delle celle angolari sono fissate dai sensori di distanza (L, F, R) e dal sensore di distanza dal target (non utilizzato nelle simulazioni e negli esperimenti riportati); le condizioni iniziali sulle celle centrali sono stabilite dalla rotazione effettuata dal robot al passo precedente (O) e dall'angolo tra l'orientamento frontale del robot e il target (P).

l'abilità mostrata dalle femmine di seguire il particolare suono (*chirp*) emesso dai maschi (Webb e Scutt 2000); ~~l'abilità di evitare gli ostacoli rilevati dalle antenne.~~

In particolare, ad ogni passo  $t$ , il riflesso optomotore cerca di compensare la rotazione compiuta al passo precedente come avviene nel grillo, ~~che cerca di compensare l'asimmetria~~ delle sue zampe durante il cammino per mantenere la direzione.

L'*obstacle avoidance* guida il robot a evitare gli ostacoli percepiti dai sensori di distanza ed è qui implementato tramite una versione semplificata del classico *potential field* introdotto da Borenstein e Koren (1992).

Infine, la *fonotassi* qui implementata è una versione semplificata del modello della fonotassi proposto da Horchler *et al.* (2004) e comanda un'azione rivolta verso il target.

## 2.2. Livello di rappresentazione

In questo contesto, come *comportamento complesso* viene considerata l'abilità di interpretare «situazioni» in termini di interazione tra robot e ambiente. Il robot, a livello afferente, percepisce gli stimoli esterni tramite il suo apparato sensoriale. Questi stimoli vengono processati a livello cognitivo per ottimizzare il suo comportamento in relazione alla missione assegnata. Scopo del livello di rappresentazione è quello di trasformare stimoli sensoriali differenti in differenti *rappresentazioni*, che determinano la modulazione dei comportamenti di base secondo meccanismi plasticamente modificati dall'esperienza.

Il livello di rappresentazione è costituito da un *blocco di preprocessing*, dal *cuore percettivo*, dalla *rete di selezione* e dal livello di *motivazione*.

Nel *blocco di preprocessing*, ciascuno stimolo sensoriale diviene ingresso per un *neurone di Sensing* (SN) dotato di funzione di attivazione di tipo lineare a tratti, modulabile sulla base di un processo di apprendimento a livello afferente. Ciascuna uscita dei SN fissa le condizioni iniziali per una cella del sistema dinamico non lineare CNN che realizza il *cuore percettivo* del livello di rappresentazione e che forma la rappresentazione esterna dell'ambiente circostante.

In questo lavoro, le CNN sono state progettate in modo da generare, sulla base delle informazioni provenienti dagli stimoli sensoriali, pattern di Turing. In particolare si usa una CNN a due strati detta *Reaction-Diffusion* (RD)-CNN di dimensione  $4 \times 4$ . Ciascuna cella  $(i,j)$  del sistema ha variabili di stato  $x_{1;i,j}$  per il primo strato e  $x_{2;i,j}$  per il secondo strato ( $i,j = \{1,2,3,4\}$ ) e la sua dinamica è definita da:

$$\begin{aligned} \dot{x}_{1;i,j} &= -x_{1;i,j} + (1 + \mu + \varepsilon)y_{1;i,j} - sy_{2;i,j} + D_1 \nabla^2 x_{1;i,j} \\ \dot{x}_{2;i,j} &= -x_{2;i,j} + sy_{1;i,j} + (1 + \mu - \varepsilon)y_{2;i,j} + D_2 \nabla^2 x_{2;i,j} \\ y_{h;i,j} &= \frac{1}{2} (|x_{h;i,j} + 1| - |x_{h;i,j} - 1|) \quad h = 1, 2 \end{aligned}$$

dove  $y_{h;i,j}$  ( $h = 1,2$ ) è l'uscita dello strato  $h$  della cella  $(i,j)$  e  $D_1$  e  $D_2$ ,  $\mu$ ,  $\varepsilon$  e  $s$  sono parametri del sistema. Per soddisfare le condizioni analitiche per ottenere pattern di Turing, i parametri sono stati fissati a  $D_1 = 0.05$  e  $D_2 = 15$ ,  $\mu = -0.7$ ,  $\varepsilon = 1.1$  e  $s = 0.9$  come spiegato da Arena *et al.* (2007).

Come mostrato dalla figura 2b, l'uscita di ciascun SN fissa le condizioni iniziali per la prima variabile di stato di due celle centrali oppure di una cella angolare: questa disparità è dovuta alla maggiore influenza nella formazione del pattern di Turing mostrata dalle celle angolari (Arena *et al.* 2007). Le condizioni iniziali per le altre celle del primo layer e per tutte le celle del secondo layer sono fissate a zero.

Il sistema evolve verso una condizione di regime in cui tutte le variabili di stato del primo strato si assestano su un valore maggiore di 1 o minore di -1: questa condizione «digitale» rappresenta un pattern di Turing. Ad esso viene attribuito un semplice codice intero, dopo aver numerato le celle da 1 a 16 a partire da quella in alto a sinistra ( $c(1,1) = 0$ ,  $c(1,2) = 1$ , ...):

$$code = \sum_{i=1}^4 \sum_{j=1}^4 \frac{(y_{1;i,j} + 1)}{2} \cdot 2^{c(i,j)}$$

Il codice viene memorizzato in un *pattern vector* (di dimensione limitata) insieme al passo  $t$  della sua ultima occorrenza in modo da consentire un meccanismo di sostituzione dell'elemento più vecchio (LRU) in caso la memoria sia piena.

L'uso dei pattern di Turing come stato stazionario di un sistema dinamico rappresenta una forma di *sensor fusion* ovvero di sintesi di una grande quantità di informazioni eterogenee in un singolo attrattore. Ad ogni passo del robot, infatti, le informazioni provenienti dai sensori e discretizzate dai SN, vengono fuse per formare un'unica e astratta rappresentazione interna dell'ambiente circostante.

I pattern codificati vengono analizzati da una rete di selezione, che associa a ciascun elemento  $q$  del *pattern vector* un set di 3 parametri ( $K_o$ ,  $K_a$  e  $K_p$ ) che vengono usati per modulare i 3 comportamenti di base. Alla prima occorrenza del pattern  $q$ , essi sono scelti in maniera casuale nell'intervallo  $[0, 1]$  con il vincolo che  $K_o^{(q)} + K_a^{(q)} + K_p^{(q)} = 1$ . Successivamente i parametri vengono modificati per mezzo del processo di apprendimento agente a livello efferente (ovvero sull'uscita) del livello di rappresentazione come spiegato nel seguito.

Una volta completato il processo di *learning*, a ciascun passo  $t$ , una volta generato il pattern di Turing  $q(t)$ , i parametri di modulazione ad esso associati vengono utilizzati per dar vita al comportamento complessivo come somma pesata dei comportamenti di base in quell'istante, tramite i suddetti parametri ( $K_o$ ,  $K_a$  e  $K_p$ ).

L'associazione tra i pattern di Turing e i parametri di modulazione è appresa tramite un algoritmo di *reinforcement learning* di tipo *reward-based*, implementato tramite una versione semplificata (Arena *et al.* 2007) della *Motor Map* proposta da Schulten (1992), laddove la funzione di *reward* è calcolata ad ogni passo  $t$  come:

$$RF(t) = RF_o(t) + RF_a(t) + RF_p(t)$$

con:

$$\begin{aligned}
RF_o(t) &= r_o \cdot |A(t-1)| \\
RF_n(t) &= \sum_i r_i \cdot e^{-\gamma d_i(t)} \\
RF_p(t) &= r_o \cdot |p(t)|
\end{aligned}$$

dove  $A(t-1)$  è la rotazione effettuata al passo  $t-1$ ,  $d_i(t)$  è la distanza tra il robot e l'ostacolo rilevato dal sensore  $i$  ( $i = L, F, R$ ) normalizzata nel range  $[0,1]$ ,  $\gamma$  è un parametro e  $p(t)$  è l'angolo tra l'orientamento frontale del robot e il target. Nelle simulazioni descritte in seguito avremo:  $r_o = 1$ ,  $r_p = r_i = 10$ . L'efficacia del comportamento viene dunque valutata in relazione al valore della funzione di *reward* rispetto al valore precedente come  $DRF(t) = RF(t) - RF(t-1)$ : un valore positivo della  $DRF(t)$  indica un comportamento di successo che viene *rinforzato*, come negli esperimenti comportamentali di Skinner (1974), allo scopo di massimizzare la RF. Viceversa, un comportamento negativo porta alla modifica delle funzioni di attivazione (in questo caso delle semplici funzioni a scalini di ampiezza variabile) dei SN a livello afferente, con conseguente modifica indiretta dei bacini di attrazione dei pattern di Turing.

Il risultato di questa procedura è che l'associazione tra gli stimoli sensoriali e i pattern di Turing viene dinamicamente regolata modulando i bacini d'attrazione dei pattern in analogia a quanto proposto in recenti studi di neurobiologia da Freeman (2004), il quale ha mostrato come il processo di apprendimento viene realizzato modulando i bacini di attrazione nello spazio delle fasi nel bulbo olfattivo dei mammiferi.

Maggiori dettagli sia sul modello matematico che sull'algoritmo sono riportati in Arena *et al.* (2007), Arena *et al.* (2008).

### 3. RISULTATI SPERIMENTALI

Le simulazioni sono state realizzate in un ambiente virtuale sviluppato in C++, che permette di generare un'arena limitata da pareti e al cui interno sono presenti il robot con l'apparato sensoriale in figura 2, un certo numero di ostacoli e un target.

L'apprendimento e il test si sono svolti con diverse configurazioni di ostacoli e target in modo da verificare l'effettiva capacità di generalizzazione dell'architettura (fig. 3).

Definendo la *RF*, si fornisce al robot la conoscenza del task da soddisfare, senza peraltro indicargli come realizzarlo: l'algoritmo di apprendimento ha il compito di far imparare al robot il corretto comportamento da seguire per raggiungere i suoi obiettivi.

Sebbene l'architettura sia stata concepita in modo da risultare applicabile a differenti robot impegnati in svariati possibili task, a patto di definire gli specifici comportamenti di base e la *RF*, in questo lavoro,

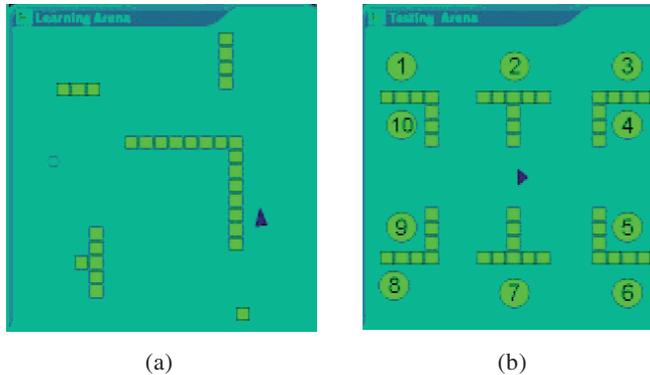


FIG. 3. (a) L'arena utilizzata nel *learning*. (b) L'arena di test: i numeri indicano le diverse posizioni in cui viene posto il target  $\lambda$

come prima applicazione, l'obiettivo del robot consiste nell'evitare gli ostacoli e raggiungere un punto specifico dell'arena (target).

Nel corso del processo di apprendimento nell'arena virtuale, quando un target viene trovato, uno nuovo appare in un punto casuale dell'ambiente. La fase di *learning* termina o quando la varianza nella ricerca dell'azione random associata agli ultimi 1.000 pattern utilizzati scende sotto una predefinita soglia, oppure a seguito del ritrovamento di 5.000 target.

All'inizio del processo di apprendimento il robot modula in maniera casuale i comportamenti di base, poiché i parametri  $K$  sono inizializzati casualmente. Durante l'apprendimento, la *Motor Map* semplificata, descritta approfonditamente in Arena *et al.* (2007), corregge i parametri associati a ciascun pattern, migliorando l'associazione tra rappresentazione interna e comportamento, mentre a livello afferente gli scalini delle funzioni di attivazione dei SN vengono appresi in maniera non supervisionata allo scopo di modulare i bacini di attrazione dei rispettivi pattern emergenti.

La figura 4 mostra l'evoluzione degli scalini del *neurone di sensing* relativo al sensore di distanza frontale durante la fase di *learning*, l'evoluzione dei  $K^{(i)}$  per il pattern più emerso ( $f$  con codice 52274) e i parametri di modulazione utilizzati nelle prime e nelle ultime 30.000 azioni di una simulazione tipica.

Il test dell'apprendimento, effettuato ogni 30.000 passi, consiste in 10 ritrovamenti del target posto in un punto diverso all'interno dell'arena virtuale (fig. 3b).

Per calcolare i benefici apportati dall'algorithm di apprendimento, abbiamo confrontato i risultati di test con quelli di un sistema di controllo che usa parametri di modulazione dei comportamenti di base costanti ( $K_i = 1/3, i = 1,2,3$ ) e con un secondo che utilizza, ad ogni passo,

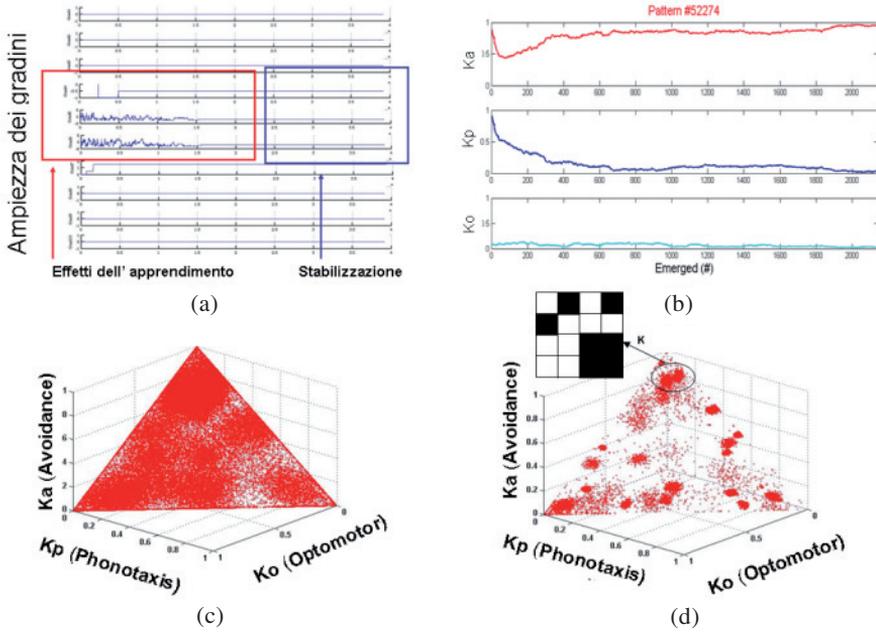


FIG. 4. (a) Ampiezze dei gradini per la funzione di attivazione del *neurone di sensing* associato al sensore di distanza frontale durante l'apprendimento in una simulazione tipica. (b) Evoluzione dei  $K_i^{(j)}$  per il pattern piú frequentemente emerso ( $f$  con codice 52274) nelle prime 8.000 occorrenze. (c) Parametri di modulazione usati nei primi 30.000 movimenti. (d) Parametri di modulazione usati negli ultimi 30.000 movimenti con l'indicazione della regione associata al pattern 52274.

parametri di modulazione random con il vincolo che  $K_o + K_a + K_p = 1$ . I risultati di tale confronto sono riportati in tabella I: l'apprendimento conduce ad una notevole riduzione del numero medio di azioni necessarie a raggiungere il target nelle 10 prove, e del numero medio di collisioni, dimostrando l'efficacia dell'architettura di controllo proposta e la sua capacità di generalizzare tramite le rappresentazioni interne.

TAB. 1. Risultati del test in termini di numero medio di azioni e numero medio di collisioni necessarie per il ritrovamento di un target nel caso di parametri di modulazioni fissi, random e appresi

Parametri	Fissi	Random	Appresi
Numero medio di azioni	166.8	176.3	28
Numero medio di collisioni	95.4	40.7	5

Quest'ultima caratteristica, in particolare, è stata dimostrata effettuando il test in un ambiente diverso da quello di learning. La figura 5

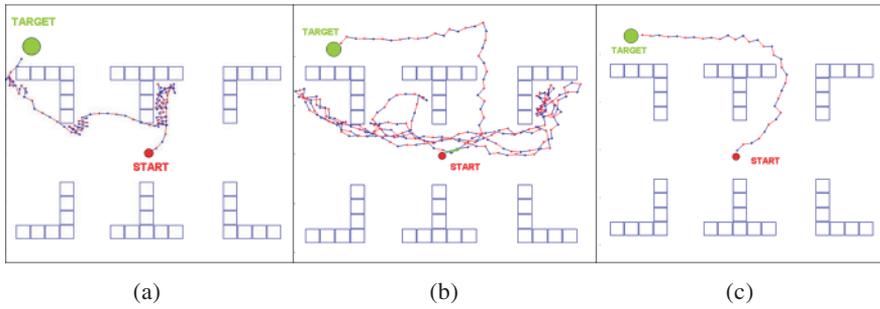


FIG. 5. Traiettorie nell'arena di test in caso di parametri di modulazione dei comportamenti di base costanti (a), scelti casualmente ad ogni passo (b) e appresi (c). Nel caso dei primi due sistemi di controllo (a-b), il robot impiega parecchio tempo per raggiungere il target e collide molte volte con gli ostacoli. Nel caso dell'architettura proposta, l'apprendimento consente al robot di seguire traiettorie verso il target molto dirette ancorché sicure.

mostra esempi di traiettorie seguite durante l'ultima fase di test nei tre casi.

Per verificare se i promettenti risultati ottenuti in simulazione fossero confermati anche su robot reali, abbiamo utilizzato un robot su ruote che si muove in un ambiente reale (fig. 6).

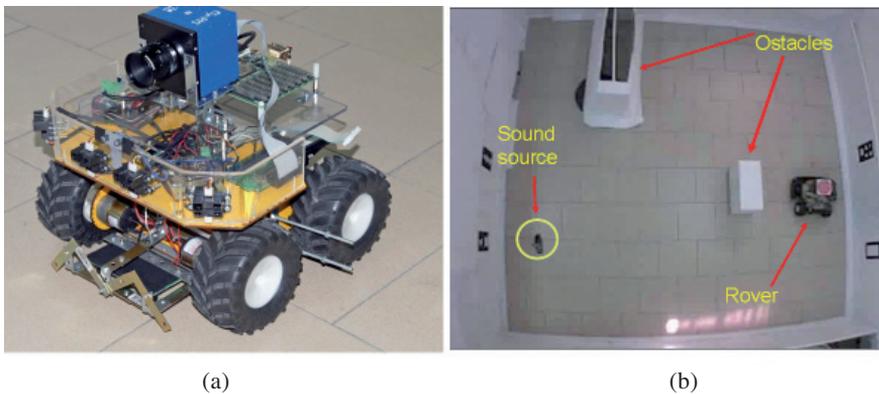


FIG. 6. (a) Robot e (b) arena sperimentale.

Il robot è interfacciato, via Bluetooth e RS232, ad un personal computer, su cui l'architettura di controllo viene eseguita.

Il robot, mostrato in figura 6.a, è un *rover* di dimensioni  $35 \times 35 \text{ cm}^2$  della Lynxmotion 4WD2 con sistema di locomozione di tipo *differential drive*. Ogni sensore di distanza è realizzato tramite una coppia di sensori

infrarossi a breve e lungo raggio che complessivamente coprono l'intervallo 3-80 cm con accuratezza di 3 cm. Sul lato frontale del robot sono posizionate due coppie di sensori e l'unica uscita è data da una media dei dati dei due sensori. Il robot è attratto da una sorgente sonora (target) che emette il suono caratteristico del grillo maschio. La misura dell'angolo tra l'orientamento frontale del robot e la sorgente sonora è ricavato tramite una scheda di processamento analogico, progettata appositamente per riconoscere e amplificare la sillaba sonora emessa dal grillo. Questa elabora i segnali provenienti da due microfoni e restituisce due uscite proporzionali all'intensità dei segnali acquisiti. Il robot è controllato da un microcontrollore STMicroelectronics della famiglia STR7x, il quale acquisisce i segnali analogici dai sensori tramite i canali ADC.

Il microcontrollore controlla in coppia, attraverso due schede Dual Full Bridge Drivers (EVAL6205N), quattro motori in corrente continua e gestisce la comunicazione con il PC.

Gli esperimenti sono stati effettuati all'interno di un'arena di  $3 \times 3 \text{ m}^2$  nella quale sono stati posizionati due ostacoli e una sorgente sonora (fig. 6b). Negli esperimenti sono stati valutati sia l'architettura con parametri di modulazione appresi che il sistema con parametri di modulazione fissi. Quest'ultimo si è infatti mostrato più efficace rispetto a quello con parametri scelti casualmente nel ritrovamento del target nelle simulazioni. In ciascun caso, sono state effettuate tre prove di raggiungimento di un target fisso a partire dalla stessa posizione e orientamento iniziale (fig. 6b).

Alcuni altri esperimenti preliminari effettuati in un ambiente reale con ostacoli, anche mobili, stanno mostrando un'elevata robustezza del modello in ambienti caratterizzati da elevato rumore. Tali risultati sono tuttora in fase di studio per comprendere a fondo, a partire dagli esperimenti in ambienti reali, le potenzialità del modello. Alcuni video in presenza di elevato rumore sono disponibili su web presso il sito ufficiale del progetto SPARK I (2004-2007), riportato nei riferimenti bibliografici.

#### 4. CONCLUSIONI

In questo lavoro è stata descritta e validata attraverso simulazioni ed esperimenti una nuova architettura di controllo del comportamento di piattaforme robotiche che implementa il ciclo sensazione-percezione-azione. L'architettura di controllo si poggia su alcune abilità prefissate e tipiche della piattaforma robotica usata, dette comportamenti di base; questi vengono modulati da una struttura di controllo superiore, il livello di rappresentazione, che apprende come associare insieme di eventi sensoriali a specifici pattern di Turing e questi ultimi ai parametri di modulazione comportamentale.

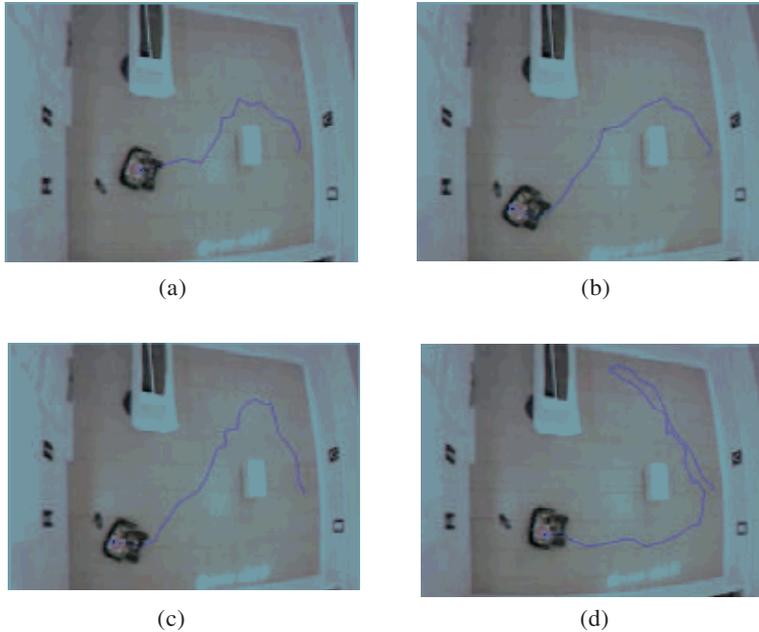


FIG. 7. Caso migliore e caso peggiore in presenza di di parametri di modulazione appresi (a-b) e costanti (c-d).

Il caso di studio affrontato è quello della navigazione di un robot mobile impegnato nel ritrovamento di un target in ambiente non strutturato. Sebbene sia da sottolineare che algoritmi dedicati ad affrontare lo specifico problema possano dare risultati comparabili, la potenzialità del nostro approccio risiede nella sua generalità in quanto può essere applicato a robot diversi impegnati in differenti missioni a patto di ridefinire comportamenti di base e funzione di *reward*.

L'approccio descritto è in via di ulteriore approfondimento per quanto riguarda robot mobili su ruote, e anche in via di applicazione su piattaforme complesse come robot esapodi, in cui il controllo delle azioni è molto più complicato e i comportamenti di base possono includere oltre che l'abilità di evitare gli ostacoli anche quella, ad esempio, di salire sopra un gradino di altezza limitata. Inoltre è già stata dimostrata la possibilità di realizzare le strutture di controllo dei comportamenti di base così come la RD-CNN in un hardware dedicato (analogico o digitale): ciò rende realistica un'implementazione dell'intera architettura di controllo «a bordo» del robot in vista di un'interazione diretta con un ambiente circostante non strutturato che, oltre ad essere «intelligente», sia effettivamente «autonoma».

## RIFERIMENTI BIBLIOGRAFICI

- ARENA P., CRUCITTI P., FORTUNA L., FRASCA M., LOMBARDO D. e PATANÈ L. (2007), *Turing patterns in RD-CNNs for the emergence of perceptual states in roving robots*, «International Journal of Bifurcation and Chaos», 17, 1, pp. 107-127.
- ARENA P., CRUSE H., FORTUNA L., LOMBARDO D., PATANÈ L. e RAPISARDA R. (2007), *Adaptive bio-inspired landmark identification for navigation control*, in *Microtechnologies for the New Millennium*, Gran Canaria, Spain, SPIE Europe.
- ARENA P., DE FIORE S., LOMBARDO D. e PATANÈ L. (2008), *A new CNN-based architecture for action-oriented perception*, in «International Journal of Bifurcation and Chaos», sottoposto per la pubblicazione.
- ARKIN R.C. (1991), *Behaviour Based Robotics*, Cambridge, MIT Press.
- BÄOHN H., SCHILDBERGER K. e HUBER F. (1991), *Visual and acoustic course control in the cricket *Gryllus-bimaculatus**, in «Journal of Experimental Biology», 159, pp. 235-248.
- BROOKS R.A. (1994), *Intelligence without reason*, 12th International Joint Conference on Artificial Intelligence, a cura di J. Mylopoulos e R. Reiter, San Mateo, CA.
- CHUA L.O. e YANG L. (1988), *Cellular neural network: Theory*, in «IEEE Transactions on Circuits and Systems», 35, 10, pp. 1257-1272
- FREEMAN W.J. (2004), *How and why brains create meaning from sensory information*, in «International Journal of Bifurcation and Chaos», 14, pp. 515-530.
- GORAS L. e CHUA L. (1995), *Turing patterns in CNNs*, Part I-II, in «IEEE Trans. Circuits Syst.-I», 42, pp. 602-626.
- MANGANARO G., ARENA P. e FORTUNA L. (1999), *Cellular neural networks: Chaos, complexity and VLSI processing*, New York, Springer-Verlag.
- NOLFI S. (2002), *Power and limits of reactive agents*, in «Neurocomputing», 42, 1, pp. 119-145.
- ORLOVSKY G.N., DELIAGINA T.G. e GRILLNER S. (1999), *Neuronal control of locomotion*, Oxford University Press.
- SKINNER B.F. (1974), *About behaviourism*, New York, Alfred Knopf.
- TURING A.M. (1952), *The chemical basis of morphogenesis*, in «Phil. Trans. of Royal Society», 237, pp. 37-72.
- VERNON D., METTA G. e SANDINI G. (2008), *A survey of artificial cognitive system: Implications for the autonomous development of mental capabilities in computational agents*, in «IEEE Transactions on Evolutionary Computation», in corso di stampa, disponibile presso: [www.eucognition.org/papers/VernonMettaSandini06.pdf](http://www.eucognition.org/papers/VernonMettaSandini06.pdf).
- WEBB B. e SCUTT T. (2000), *A simple latency dependent spiking neuron model of cricket phonotaxis*, in «Biological Cybernetics», 82, 3, pp. 247-269.
- WEBSITE: [www.spark.diees.unict.it](http://www.spark.diees.unict.it) (2004-2007).

Il lavoro è stato possibile grazie al progetto Europeo SPARK (IST 2004-2007). Gli autori ringraziano i partner del progetto (la prof. Barbara Webb, Università di Edinburgo, il prof. Holk Cruse, dell'Università di Bielefeld, il prof. Manuel G. Velarde dell'Università Complutense di Madrid) per la preziosa collaborazione che ha portato all'introduzione del modello percettivo presentato. Gli autori infine ringraziano l'ing. Alessandro Costa per aver curato alcune delle simulazioni.

*Paolo Arena, Davide Lombardo e Luca Patanè, Dipartimento di Ingegneria Elettrica Elettronica e dei Sistemi, Università degli Studi di Catania, Viale A. Doria, 6, 95125 Catania. E-mail: {parena, dlombardo, lpatane}@diees.unict.it*

