

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Scuola di Ingegneria e Architettura
Corso di Laurea in Ingegneria Informatica, Elettronica e
Telecomunicazioni

APPRENDIMENTO HEBBIANO IN
ROBOTICA: TEORIA E APPLICAZIONE

Elaborata nel corso di: Fondamenti di Informatica B

Tesi di Laurea di:
MATTEO FRANCIA

Relatore:
Prof. ANDREA ROLI

ANNO ACCADEMICO 2013–2014
SESSIONE II

PAROLE CHIAVE

Apprendimento Hebbiano

Robotica

Value system

Distributed adaptive control

*A nonna Ione che in questo giorno non è
presente ma che è sempre con me.
Ai miei genitori, Luigi e Mirella, che hanno
saputo crescermi, aiutarmi e sostenermi con
entusiasmo.
A Chiara, per condividere ogni giorno le fatiche, i
sorrisi e l'amore della vita.
Ai miei cugini, Laura, Luca e Martina, per me
come fratelli.
Ai nonni e gli zii, per esserci sempre stati.
A Antonella, Roberta, Gualtiero e Roberto, per
l'affetto e l'allegria che ogni giorno mi
trasmettono.
Agli amici Alessandro, Benedetta, Luigi, Simone
e Thomas, per essere stati i compagni di viaggio
di questi anni.
A Mauro che mi ha insegnato a non arrendermi
mai.
Alla mia famiglia, per essere così unica e speciale.
A Neve e all'aereo invisibile.*

Indice

Introduzione	ix
1 Apprendimento Hebbiano	1
1.1 Preambolo	1
1.2 La teoria di Hebb	2
1.3 Formulazione matematica dell'apprendimento Hebbiano	8
1.3.1 Ottimizzazione dell'apprendimento Hebbiano per reti neurali a singolo livello	12
1.3.2 Regola di Oja	13
1.3.3 L'importanza dei vincoli nell'apprendimento Hebbiano	14
2 Value system	17
2.1 Apprendimento value-based in robotica	18
2.2 Il principio valore	19
2.3 Apprendimento Hebbiano aumentato	20
2.4 Confronto tra regole di apprendimento	21
3 Distributed adaptive control	23
3.1 Teoria della connessione tra mente, corpo e cervello . .	25
3.2 Architettura DAC	27
3.3 Comportamenti orientati al goal	30
4 WABOT	33
4.1 Requisiti	34
4.2 Analisi dei requisiti	35
4.2.1 Modello del dominio	38
4.2.2 Casi d'uso	39

4.2.3	Scenari	40
4.3	Analisi del problema	42
4.3.1	WABOT e la nicchia ecologica	42
4.3.2	Value system	44
4.3.3	Architettura della rete neuronale	49
4.3.4	Debug e problemi relativi all'analisi dei risultati	52
4.3.5	Stile architeturale	52
4.3.6	Abstraction gap	55
4.4	Piano di lavoro	55
4.5	Osservazioni empiriche	56
	Conclusioni	59

Introduzione

Intelligenza è un termine descrittivo. Definire se un organismo sia intelligente o no non è semplice. Significati diversi sono dati in riferimento al *behaviour* di un agente, al contesto di tale comportamento e all'ambiente in cui l'agente stesso è immerso [22]; nessuno di questi è totalmente soddisfacente.

Secondo la definizione data dal dizionario Treccani, l'intelligenza è un *complesso di facoltà psichiche e mentali che consentono all'uomo di pensare, comprendere o spiegare i fatti o le azioni, elaborare modelli astratti della realtà, [...] giudicare, e lo rendono insieme capace di adattarsi a situazioni nuove e di modificare la situazione stessa quando questa presenta ostacoli all'adattamento; propria dell'uomo [...] è riconosciuta anche, entro certi limiti (memoria associativa, capacità di reagire a stimoli interni ed esterni, ...), agli animali [...]*.

Il libro *Understanding Intelligence* [33] riporta le definizioni di alcuni studiosi. Pinter evidenzia l'abilità di adattarsi dell'individuo: *intelligence is the ability to adapt oneself adequately to relatively new situations in life*. Peterson mette in luce la capacità di correlare stimoli: *intelligence is a biological mechanism by which the effects of a complexity of stimuli are brought together and given a somewhat unified effect in behaviour*. Dearborn sottolinea invece l'importanza dell'apprendimento e dell'esperienza: *intelligence is the capacity to learn or to profit by experience*.

Lo psicologo canadese Hebb, nel libro *The Organization of Behavior: A Neuropsychological Theory* [17], introduce due tipologie di intelligenza: A e B. Il primo tipo, noto come *the innate potential*, rappresenta la capacità dell'individuo di svilupparsi in base alle proprietà innate di possedere un cervello e un buon metabolismo neuronale. B indica invece il funzionamento del cervello a seguito del

percorso di sviluppo. Se A determina B, l'intelligenza è un carattere puramente ereditario e indipendente dall'esperienza del singolo. Un potenziale innato non garantisce che il processo di sviluppo si verifichi, l'esperienza dell'individuo è essenziale per l'evoluzione del comportamento. Il tipo A non può quindi essere considerato come l'unica causa dell'intelligenza B, esso è una condizione necessaria ma non sufficiente.

La scienza cognitiva (*cognitive science*) [38] propone uno studio interdisciplinare di tale fenomeno e dei processi della mente. Tali argomenti sono oggetto di molte discipline: intelligenza artificiale, fisiologia, psicologia, filosofia e tante altre.

La mancanza di un'ontologia formale non è di aiuto. In senso comune la capacità di pensiero, prendere decisioni e apprendere sono punti chiave nell'interpretare l'intelligenza [32]. Gardner introduce la teoria delle intelligenze multiple in [12], secondo cui non esiste una versione unica universale di tale fenomeno, bensì forme diverse indipendenti tra loro.

Massimiliano Versace indica che la mente, motore principale di behaviour intelligenti, sia strettamente legata al cervello e al corpo. Quest'ultimo apre a una nuova dimensione: l'interazione con l'ambiente [30].

Brooks, in [5], spiega che non è semplice tracciare una distinzione formale tra intelligenza e nicchia ecologica, probabilmente tale discriminazione non è nemmeno rilevante se l'agente è *embodied* nell'ambiente: il robot fa esperienza diretta del mondo, le azioni intraprese hanno un feedback immediato sulle sue percezioni. L'interazione è il punto fermo dell'architettura di sussunzione proposta dallo studioso [6], l'intelligenza dell'agente è espressa tramite coordinazione e collaborazione di processi asincroni.

Questa architettura mette in luce un problema noto come *the frame of reference*: la prospettiva di un osservatore esterno che percepisce l'esistenza dell'agente differisce dal punto di vista dell'agente stesso. Il behaviour di un organismo spesso non può essere spiegato solamente in funzione dei meccanismi interni, bensì è necessario considerare la dimensione dell'interazione. Un comportamento intelligente può emergere dalla collaborazione di processi semplici, le formiche ne sono un esempio [16].

Negli ultimi decenni sono stati fatti progressi negli studi sul funzio-

namento del cervello [10]. La produzione di modelli sintetici consente di eseguire un'analisi impossibile da effettuare su organismi reali.

Il sistema nervoso centrale non agisce in isolamento: il cervello opera tramite un corpo che si trova in un ambiente. L'adattamento dell'individuo è quindi fondamentale ai fini della sopravvivenza, esso porta al miglioramento genetico della specie aumentando la complessità interna dell'individuo. In termini di teoria selezionista, l'evoluzione si mostra anche nella perdita di caratteri precedentemente acquisiti.

La teoria Hebbiana propone una spiegazione dell'adattamento delle cellule neuronali cerebrali durante la fase di apprendimento. Come riportato in [36]: *an agent is learning if it improves its performance on future tasks after making observations about the world.*

Questa introduzione accenna a una piccola parte dei problemi noti nello studio dell'intelligenza. L'interazione di cervello, mente e corpo è fondamentale ai fini dell'apprendimento. Risultati notevoli sono ottenuti nello studio di tale fenomeno mediante l'osservazione di esperimenti termodinamici ed entropia, vedi [52], [53] e [54]. Questo approccio matematico si fonda sull'osservazione dell'apprendimento da un punto di vista esterno all'agente, lo scopo di questa trattazione è invece quello di fornirne una formulazione più naturale.

Apprendimento Hebbiano in robotica: teoria ed applicazione cattura l'essenza dell'elaborato mettendone in luce le parole chiave. Partendo dall'introduzione della teoria necessaria per trattare l'apprendimento di un agente, si prosegue con l'approfondimento di tematiche utili per la costruzione di un robot in grado di imparare.

Nel 1949 Hebb sviluppa una teoria sull'apprendimento associativo espresso in termini di modulazione dell'efficienza sinaptica tra cellule neuronali. Questo studio approfondisce il problema dal punto di vista neurofisiologico dell'organismo.

Il primo capitolo si apre con una breve biografia dello psicologo canadese, utile per contestualizzare le sue ricerche. Si fornisce in seguito una descrizione della teoria Hebbiana, soffermandosi particolarmente sulle nozioni di plasticità sinaptica, *assembly* di neuroni e processo del pensiero. Essendo non supervisionato, l'apprendimento Hebbiano presenta pro e contro. Dopo averne ricavato la formulazione matema-

tica, si introduce l'importanza dell'uso di vincoli al fine di indirizzare la convergenza dell'apprendimento.

Il libro *Understanding Intelligence* lega i concetti di intelligenza e apprendimento a quello di agente autonomo. Ai fini dell'adattamento, il singolo deve essere in grado di percepire cosa sia giusto o sbagliato. La soluzione è proposta in forma di *value system*.

Il secondo capitolo si presta alla definizione di tale elemento e ne descrive le caratteristiche salienti. Il value system è responsabile dei bisogni, emozioni e apprendimento dell'agente.

I primi due capitoli sono fortemente legati tra loro: il sistema-valore è qui inteso come punto chiave per lo sviluppo degli studi di Hebb all'interno delle reti neurali. Inoltre, il value system consente di irrobustire l'apprendimento estendendo la formulazione proposta dallo psicologo.

Il terzo capitolo propone un'architettura per la modellazione del processo adattativo in funzione del condizionamento Pavloviano. DAC fornisce un punto di partenza per lo sviluppo di robot autonomi governati da reti neurali, e rappresenta l'intersezione tra teoria del comportamento e sperimentazione scientifica.

Il modello proposto da Rolf Pfeifer and Paul Verschure si basa sulla nozione di value system e mette in luce l'importanza del fenomeno di emergenza: a partire dall'interazione con l'ambiente, il robot è in grado di imparare comportamenti complessi non previsti in fase di design.

Dopo tanta teoria il percorso si conclude con un po' di pratica.

WABOT è un robot mobile in grado di imparare esplorando l'ambiente. Il suo compito è quello di raggiungere la fonte di cibo evitando di urtare gli oggetti presenti nella nicchia ecologica.

Questo caso studio consente di applicare le nozioni introdotte nei capitoli precedenti. A partire dall'architettura DAC, si esplorano le funzionalità del value system e dell'apprendimento Hebbiano con la possibilità di trarre osservazioni empiriche fondate sulla sperimentazione.

L'italiano, lingua per eccellenza di scrittori e poeti, non si presta all'e-

sposizione di temi informatici. Come si può osservare nell'introduzione, questo testo è ricco di inglesismi che non trovano una traduzione diretta. È bene fare attenzione anche ai *false friends*, ovvero quei vocaboli che, se tradotti, assumono un significato differente dall'originale (situato e *situated*). Termini stranieri e citazioni, laddove appaiono per la prima volta, sono riportati in stile *corsivo*.

Pfeifer e Verschure sono figure di riferimento del percorso seguito. Gran parte del materiale informativo è preso da articoli scientifici di loro proprietà. Si riporta con particolare interesse il libro *Understanding Intelligence*, caposaldo di questo elaborato.

Capitolo 1

Apprendimento Hebbiano

1.1 Preambolo

Donald Olding Hebb (Chester, Luglio 22, 1904 - Agosto 20, 1985), nato in Canada, fu uno psicologo influente nell'area della neuropsicologia. Egli studiò la modalità con cui il funzionamento delle cellule neuronali contribuisce allo sviluppo di processi psicologici, approfondendo il legame tra sistema nervoso e behaviour.

Hebb diede il nome a una teoria conosciuta come apprendimento Hebbiano, che egli stipulò nel libro *The Organization of Behavior: A Neuropsychological Theory* [17]. Tale teoria fonda le basi matematiche per lo sviluppo di un nuovo algoritmo di apprendimento nelle reti neuronali.

Hebb fu cresciuto da genitori entrambi medici. Gli scritti di Freud e Watson stimolarono nel ragazzo un forte interesse verso la psicologia. Nel 1937 accettò una borsa di studio presso il *Montreal Neurological Institute* (MNI) dove studiò le conseguenze provocate da lesioni al cervello su intelligenza e comportamenti di uomini e animali. Tramite questi esperimenti Hebb concluse che il substrato neurale responsabile delle funzioni di memoria e apprendimento è distribuito e non strettamente localizzato in una singola area del cervello. Nel 1942, presso *Yerkes Laboratory*, egli iniziò a esplorare i processi emozionali nei primati e la loro influenza sull'apprendimento. Stimolato dal clima presente nel centro di ricerca, Hebb iniziò a fondare le basi del proprio pensiero articolandole in un elaborato scritto. La motivazio-

ne principale sorse dalla necessità di dare un fondamento scientifico a fenomeni di natura visiva, e non solo, non spiegabili con le conoscenze contemporanee.

The Organization of Behavior: A Neuropsychological Theory fu pubblicato nel 1949 ed in breve tempo riscosse particolare successo. In tale libro è presentata la teoria che secondo lo psicologo canadese lega il comportamento di un soggetto con elementi di natura fisiologica e psicologica: *the problem of understanding behavior is the problem of understanding the total action of the nervous system and vice versa*. Conoscere il comportamento delle singole cellule e le modalità con cui esse interagiscono è necessario, ma non sufficiente, per trarre conclusioni sul funzionamento del cervello mammifero. In questo modo Hebb cercò di colmare il *gap* che fino a quel momento aveva separato psicologia e neurofisiologia.

1.2 La teoria di Hebb

The Organization of Behavior: A Neuropsychological Theory fornisce una spiegazione prettamente teorica riguardo ai processi mentali (tra i quali percezione, apprendimento, memoria ed emozioni), riconducendoli alle proprietà fisiologiche del cervello. Capire le funzionalità delle aree di quest'ultimo e studiarne l'interazione è importante al fine di correlare tali attività con la teoria del comportamento. Uno dei primi problemi affrontati dallo psicologo fu la coesistenza di discipline che fino a quel momento avevano seguito percorsi divergenti. Hebb riuscì a fondare la sua teoria sull'unione di elementi di neurofisiologia e psicologia.

L'attività relativa al pensiero può essere vista sotto luci diverse. Secondo il punto di vista psicologico, il pensiero è un processo non completamente controllato dall'ambiente. In prospettiva fisiologica, il pensiero è una trasmissione di natura complessa tra il sistema dei sensori e quello motore.

Le sperimentazioni su mammiferi tenute da Hebb dimostrarono che è impossibile descrivere un behaviour come una sequenza di interazioni *stimulus-response* secondo la riflessologia Pavloviana. Qualcosa di complesso come il pensiero interviene, nonostante la natura scientifica

di tale fenomeno rimanga sconosciuta allo psicologo. Hebb supera il limite del comportamentismo, il behaviour di un organismo non è imputabile unicamente a interazioni osservabili con l'ambiente. Questa assunzione è in contrasto con il principio di attenzione, ovvero un behaviour basato su un insieme di risposte selettive agli eventi esterni. Così come non è possibile asserire che un comportamento sia deducibile unicamente dall'interazione con l'ambiente esterno, non è possibile negare l'influenza di quest'ultimo sull'insieme delle azioni intraprese da un individuo.

La psicologia moderna assume che behaviour e funzionalità delle cellule neuronali siano legati tra loro, è negata l'esistenza di un'entità superiore che moduli il comportamento dei neuroni a dispetto di ciò che essi farebbero naturalmente.

Prima di proseguire con lo sviluppo della teoria Hebbiana è utile fornire una descrizione, prettamente discorsiva, riguardo alla struttura e al funzionamento della cellula neuronale.

Un neurone è l'unità cellulare che costituisce il tessuto nervoso. Esso è una cellula biologica, eccitabile elettricamente e chimicamente, in grado di processare informazione. I segnali in ingresso sono propagati al soma mediante dendriti. Il corpo del neurone modula il proprio livello di attivazione in base agli input ricevuti. Nel caso in cui tale valore sia sufficientemente elevato all'interno di un certo intervallo di tempo, il soma trasmette una sequenza di impulsi elettrici (*spike*) alle cellule neuronali cui è connesso mediante una fibra nervosa chiamata assone [23]. Lo scambio informativo tra neuroni è regolato da sinapsi, esse sono strutture in grado di abilitare o inibire il transito di segnali.

L'informazione è trasmessa in forma di treni di impulsi. Quando uno di questi attraversa una giunzione sinaptica, il sistema nervoso rilascia neurotrasmettitori al fine di migliorare o inibire il canale informativo. La modulazione di tale efficacia porta a forme di memoria e apprendimento ed è nota come plasticità sinaptica. Questa proprietà neurobiologica fa parte dei fenomeni di plasticità neuronale, l'aumento della forza di trasmissione è detto potenziamento a lungo termine (*Long term potentiation, LTP*), il processo inverso è quello di depressione a lungo termine (*Long term depressione, LTD*).

Il fattore tempo è fondamentale per la trasmissione, la direzione con cui un treno di impulsi è condotto può dipendere dal *timing* di

altri segnali. *Spike-timing-dependent plasticity* (STDP) è il processo biologico in grado di modulare l'efficacia sinaptica in base agli istanti in cui i neuroni si scambiano segnali, gli intervalli di trasmissione influiscono sulle fasi di LTP e LTD.

L'interconnessione di cellule neuronali porta alla formazione di una struttura estremamente complessa. Il sistema nervoso centrale appare come una rete neuronale, esso realizza una serie di collegamenti che portano alla connessione di elementi ricettori e attuatori.

Studi nel campo elettrofisiologico indicano che ogni parte del cervello sia costantemente attiva, le eccitazioni di neuroni sono imposte su quelle pre-esistenti. Questa assunzione indica che la conseguenza di uno stimolo sia condizionata dall'attività neuronale precedente.

Il processo di *firing* tra cellule può essere spontaneo, l'attività sensoriale è essenziale per la sua regolazione ma non è necessaria per iniziarlo. Esso può essere attivato a fronte del verificarsi di un evento e può persistere al termine dello stesso. Tale persistenza può essere interpretata come una forma di memoria.

La trasmissione di informazione tra cellule è correlata a variazioni del campo di forza misurato dall'elettroencefalogramma (EEG). Alti livelli di potenziale indicano un sincronismo locale (*hyper-synchrony*) nel firing e producono una fase di incoscienza inibendo gli stimoli sensoriali. Al contrario, l'attività sensoriale interrompe la fase di *hyper-synchrony* e induce bassi livelli di potenziale.

La ripetuta stimolazione di specifici ricettori porta allo sviluppo di un assembly di cellule neuronali in seguito a un processo associativo. Tale insieme è in grado di agire come un sistema chiuso anche al termine della stimolazione. Sistemi di cellule attivi nello stesso momento tendono ad associarsi influenzandosi reciprocamente.

Il processo di apprendimento associativo costituisce la base della teoria di Hebb. Un organismo apprende correlando tra loro stimoli (condizionamento classico) oppure stimolo e risposta (apprendimento per rinforzo o condizionamento operante). La prima forma di apprendimento, nota anche come condizionamento Pavloviano [29], consiste nell'associare uno stimolo condizionato (*conditioned stimulus*, CS) artificiale a uno incondizionato (*unconditioned stimulus*, US) che naturalmente evoca un comportamento riflesso (*unconditioned response*, UR). Solitamente CS è uno stimolo neutro per l'organismo, tuttavia

in seguito al condizionamento esso è in grado di provocare una risposta condizionata (*conditioned response*, CR). CR e UR possono coincidere, tuttavia la prima è acquisita mediante l'esperienza legata al condizionamento del soggetto. CS e US sembrano essere correlati alla funzione dell'amigdala. La stimolazione di quest'area porta a ottenere risposte non condizionate, se soggetta a lesioni le associazioni tra CS e US possono essere danneggiate.

Il condizionamento classico è realizzato mediante la modulazione della plasticità sinaptica riducendo la resistenza al passaggio di informazione. Questi cambi strutturali permettono lo sviluppo di due tipi di memoria: istantanea ed evanescente, istantanea e permanente.

Il punto chiave della teoria dello psicologo canadese è riportato testualmente e fornisce le basi per la formulazione matematica dell'apprendimento Hebbiano in reti neuronali:

When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.

La forza sinaptica aumenta in efficacia in modo proporzionale alla correlazione dell'attività pre e post-sinaptica. In neuroscienza questa assunzione è riferita come *Hebb Synapse* e fornisce spunto per diversi studi.

Esperimenti elettrofisiologici confermano ciò che è stato riportato sopra: stimoli ripetuti provocano il cambiamento della risposta post-sinaptica a seguito, ad esempio, di fasi LTP. Tali esperimenti sono effettuati eccitando neuroni *target* con specifici stimoli. Quando la cellula neuronale post-sinaptica si attiva, la connessione con la componente pre-sinaptica aumenta di efficacia e i segnali in input sono potenziati in modo temporaneo o permanente.

Una protuberanza sinaptica (*synaptic knob*) si sviluppa aumentando le dimensioni di contatto tra assone e dendrite. Questo processo, che si verifica in conseguenza di un'elevata attività neurale tra due cellule, porta alla diminuzione della resistenza sinaptica. L'incremento dell'area di contatto indica che il processo di firing della cellula afferente A porta a un firing predicibile della cellula efferente B. A acquisisce controllo su B, le due cellule neuronali saranno attive contemporaneamente con probabilità direttamente proporzionale alle dimensioni

dell'area di contatto.

In questo modo si sviluppano veri e propri assembly di neuroni che agiscono come sistemi chiusi e dinamici. La fase di *recruitment* porta all'inserimento di nuove cellule all'interno dell'assembly. Il meccanismo di frazionamento esclude le componenti che cessano di essere sincronizzate con l'attività del sistema. L'associazione tra cellule e assembly può variare nel corso del tempo. Questa variabilità porta allo sviluppo di proprietà differenti della percezione a seconda dello stadio di integrazione in corso.

L'attività simultanea di due assembly, denotati per brevità di annotazione con A e B, stabilisce una facilitazione tra gli stessi tramite le connessioni anatomiche e lo sviluppo di knob sinaptici. Il sistema risultante è *più della somma delle parti che lo compongono*. Al fine di sottolineare il principio, si preferisce denotarlo con un nuovo simbolo T e di non utilizzare il termine AB.

La funzionalità del nuovo sistema è determinata dalla successione ordinata e ripetuta delle attività dei sottosistemi. Nei primi stadi di sviluppo della percezione T è eccitato dalla successione delle attivazioni in A e B, con il passare del tempo lo stesso effetto può scaturire dall'attivazione di uno solo dei sottosistemi.

Un assembly è, quindi, una rete di cellule neuronali che tendono ad attivarsi insieme in seguito all'apprendimento per associazione. Durante un processo mentale, la forza sinaptica delle connessioni che legano i componenti del sistema è aumentata in seguito alle ripetute attivazioni simultanee. All'interno dell'assembly la modulazione dell'efficienza sinaptica segue un preciso ordine di attivazione. L'attività del sistema è fortemente legata all'architettura spazio-temporale dei propri membri. In seguito è riportato un ulteriore passo della teoria di Hebb che spiega come il formarsi di agglomerati di assembly sia legato alla formulazione del pensiero:

Any frequently repeated, particular stimulation will lead to a slow development of a cell assembly [...] capable of acting briefly as a closed system, delivering facilitation to other such systems and usually having a specific motor facilitation. A series of such events constitutes a phase sequence - the thought process. Each assembly action may be aroused by a preceding assembly, by sensor event or by both. The central facilitation from one of these activities on the next is the



Figura 1.1: *Gestalt law of closure* (principio di completamento): la figura è percepita nel suo intero nonostante non sia completa.

propose of attention. [...] The theory is a form of connectionism [...], it does not make any single nerve cell or pathway essential to an habit or perception.

Immaginando A, B e C come i vertici di un triangolo, la figura è percepita nel suo intero solo al termine di una fase di apprendimento che porti alla percezione delle singole parti dell'oggetto in modo separato. Il processo percettivo è costituito da una successione temporale di attività della struttura associativa sviluppata mediante le sinapsi. Se la percezione delle parti porta a quella dell'intero, allora il processo di completamento diviene un corollario di quello associativo [51], vedi figura (1.1).

Il pensiero è il risultato dell'attivazione sequenziale di assembly tra loro indipendenti. Tale teoria soffre della mancanza di dettagli scientifici ma fornisce una linea guida che si è dimostrata valida nel corso degli anni.

I concetti descritti in precedenza non sono aggiornati agli studi più recenti. Lo scopo di questa sezione è quello di fornire uno spaccato, seppur breve, degli studi dello psicologo canadese. Per ragioni legate a fattori di tempo, si è deciso di non approfondire gli effetti

delle emozioni e delle lesioni cerebrali nella fase di apprendimento e percezione.

1.3 Formulazione matematica dell'apprendimento Hebbiano

La teoria di Hebb fornisce le basi per un algoritmo di apprendimento all'interno delle reti neuronali. Tale regola deriva dalla modulazione della plasticità sinaptica di due neuroni attivi simultaneamente, il principio è anche indicato come *neurons which fire together, wire together*.

Prima di procedere con un approfondimento di natura matematica è necessario definire un piccolo glossario dei termini utilizzati più frequentemente.

- Forza sinaptica: parametro che può essere modulato al fine di ottimizzare l'architettura della rete neuronale per compiere un *task* specifico. Tale parametro è anche riferito come peso della connessione, peso sinaptico ed efficienza sinaptica;
- Apprendimento: modulazione della forza sinaptica;
- Regola di apprendimento: procedura che regola il processo di apprendimento. Esistono diverse classi di regole, modulazioni sinaptiche guidate dalla correlazione dell'attività pre e post-sinaptica sono dette apprendimento Hebbiano (*Hebbian learning*).

La formulazione è ricavata da articoli scientifici che utilizzano notazioni diverse, quando possibile è mantenuto invariato il significato delle annotazioni matematiche.

Dove ritenuto opportuno la spiegazione dei parametri è riportata in forma di elenco puntato al fine di mettere in luce il loro significato.

La teoria dell'apprendimento si basa su principi di località e cooperatività [14]. La prima assunzione indica che la modulazione della forza

sinaptica dipende unicamente da variabili locali quali lo stato della sinapsi ed attività pre e post-sinaptiche:

$$\frac{dw_{ij}}{dt} = \Delta w_{ij} = F(w_{ij}, f_i(a_i), f_j(a_j)) \quad (1.1)$$

- Δw_{ij} : variazione della forza sinaptica;
- w_{ij} : forza sinaptica corrente;
- f_i : funzione che regola l'attività post-sinaptica;
- f_j : funzione che regola l'attività pre-sinaptica;
- a_i : attività del neurone post-sinaptico;
- a_j : attività del neurone pre-sinaptico.

Considerando f_i e f_j come funzioni identità si ottiene:

$$\frac{d}{dt} = \Delta w_{ij} = F(w_{ij}, a_i, a_j) \quad (1.2)$$

Il principio di cooperatività indica che l'attività delle cellule neuronali pre e post-sinaptiche deve essere simultanea affinché il processo di apprendimento regoli il peso della connessione. La funzione F può essere sviluppata in serie di Taylor:

$$\Delta w_{ij} = F(w_{ij}, a_i, a_j) = c_0(w_{ij}) + c_1^{post}(w_{ij})a_i + c_1^{pre}(w_{ij})a_j + c_2^{post}(w_{ij})a_i^2 + c_2^{pre}(w_{ij})a_j^2 + c_2^{corr}(w_{ij})a_i a_j + O(a^3) \quad (1.3)$$

Il coefficiente $c_2^{corr}(w_{ij})a_i a_j$ realizza la funzione AND del principio di cooperatività.

Una prima versione di F è ottenuta eguagliando tutti i coefficienti, eccetto $c_2^{corr}(w_{ij})$, a 0. $c_2^{corr}(w_{ij})$ rappresenta il ratio η che determina la velocità di convergenza del processo di apprendimento. Se $c_2^{corr}(w_{ij}) < 0$ tale processo è detto anti-Hebbiano: la forza sinaptica è diminuita ogni volta che due neuroni connessi tra loro sono attivi simultaneamente (LTD), si ottiene così un comportamento opposto a quello descritto nella teoria di Hebb. Considerando unicamente il coefficiente $c_2^{corr}(w_{ij})$ la funzione F diviene quindi:

$$\Delta w_{ij} = F(w_{ij}, a_i, a_j) = c_2^{corr}(w_{ij})a_i a_j \quad (1.4)$$

Essa è riferita come *plain Hebbian learning*, sostituendo η in (1.4) si ottiene:

$$\Delta w_{ij} = \eta a_i a_j \quad (1.5)$$

Il modello proposto tuttavia omette il fattore temporale descritto dalla teoria STDP. Se due neuroni si attivano nello stesso istante è difficile dimostrare che l'attività post-sinaptica sia conseguenza di quella pre-sinaptica. Quando la cellula neuronale pre-sinaptica si attiva, un certo intervallo di tempo intercorre prima che il treno di impulsi sia propagato ai neuroni efferenti. È possibile estendere il modello precedente (1.5):

$$\Delta w_{ij} = \eta \bar{a}_i \bar{a}_j \quad (1.6)$$

\bar{a}_i e \bar{a}_j rappresentano la media dell'attività pre e post-sinaptica all'interno di un intervallo di tempo dT . Il peso della connessione a seguito del processo di apprendimento è definito come:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t) \quad (1.7)$$

Trascurando la media temporale delle attivazioni e sostituendo (1.5) in (1.7) si ottiene:

$$w_{ij}(t+1) = w_{ij}(t) + \eta a_i(t) a_j(t) \quad (1.8)$$

Questa formulazione matematica non pone vincoli al valore di w_{ij} , ne consegue che questo modello soffre di problematiche legate a *over-generalization*. Al fine di evitare un *overflow* del peso sinaptico è necessario che:

$$\lim_{w_{ij} \rightarrow w^{max}} c_2^{corr}(w_{ij}) = 0 \quad (1.9)$$

w^{max} (*upper bound*) è il valore massimo consentito del peso sinaptico. Ad esempio è possibile assumere:

$$c_2^{corr}(w_{ij}) = k(w^{max} - w_{ij}) \quad (1.10)$$

k è una costante positiva. A fronte di stimolazioni ripetute le sinapsi tendono saturare a w^{max} . Un *forgetting* ratio ϵ è necessario nonostante non sia previsto nella teoria dello psicologo canadese:

$$\Delta w_{ij} = k(w^{max} - w_{ij}) a_i a_j - \epsilon w_{ij} \quad (1.11)$$

Un aspetto interessante delle regole di apprendimento è il principio di competizione: se alcune sinapsi sono rinforzate, le altre legate al medesimo neurone post-sinaptico sono indebolite, vedi 1.3.3.

La funzione F di apprendimento può essere implementata in modalità diverse; vedi [14], 1.3.1 e 1.3.2.

Considerazioni interessanti riguardano anche la funzione pre-sinaptica f_j :

$$f_j(a_j) = a_j - w_{ij} \quad (1.12)$$

Questa equazione realizza un fenomeno conosciuto come *heterosynaptic LTD*, la ripetizione di un treno di impulsi tende a diminuire il peso della connessione sinaptica.

Il termine heterosynaptic indica un modello di plasticità in cui l'attività di un certo neurone provoca la modulazione del peso delle connessioni con altri neuroni. Si supponga di studiare il comportamento di tre cellule neuronali n_1 , n_2 , n_3 connesse allo stesso neurone n_T . n_1 e n_2 si attivano simultaneamente mentre n_3 è attivo quando le altre unità non lo sono. L'algoritmo di plain Hebbian learning (1.5) aumenta la forza sinaptica di n_1 e n_2 in maniera maggiore rispetto al peso della connessione tra n_T e n_3 .

Tale forma di apprendimento rafforza i *pattern* di risposta a una specifica stimolazione. Se quest'ultima è desiderata allora le performance della rete neuronale rispetto al task assegnato migliorano, in caso contrario possono accadere conseguenze imprevedibili come, ad esempio, lo sviluppo di fobie e distonia [24].

La formulazione matematica qui proposta prevede una forma di apprendimento non supervisionato che tuttavia non è in grado di ricoprire l'intero processo di apprendimento umano e animale.

In [28] è introdotto un meccanismo di feedback in grado di riconoscere errori durante la fase di apprendimento. L'approccio *neocortical learning* combina il plain Hebbian learning con meccanismi *error-driven* portando alla creazione di LEABRA (*Local, Error-driven and Associative, Biologically Realistic Algorithm*).

1.3.1 Ottimizzazione dell'apprendimento Hebbiano per reti neurali a singolo livello

L'apprendimento Hebbiano precedentemente proposto può portare a tempi di convergenza lunghi. Zurada, in [55], propone un meccanismo di apprendimento supervisionato per reti neurali a singolo livello.

(1.5) è ridefinita in forma vettoriale come:

$$\Delta w = \eta x o \tag{1.13}$$

- Δw : vettore delle variazioni dei pesi sinaptici;
- η : costante di apprendimento positiva;
- x : vettore degli input di una cellula neuronale;
- o : valore ritornato dalla funzione di attivazione del neurone.

Si supponga di eseguire una fase di *training* su una cellula neuronale con un insieme di p pattern x (x_1, \dots, x_p), l'output desiderato è indicato con d (d_1, \dots, d_p). Al fine di eseguire un addestramento mirato (1.13) è specializzata nella seguente formulazione:

$$\Delta w = \eta x d \tag{1.14}$$

d indica il vettore degli output desiderati, la fase di apprendimento consiste in:

$$\begin{aligned} & w^0 \\ w^1 &= w^0 + \eta x_1 d_1 \\ & \dots \\ w^p &= w^0 + \eta \sum_{i=1}^p x_i d_i \end{aligned} \tag{1.15}$$

Al termine dell' n -esimo ciclo il risultato è:

$$w^{np} = w^0 + n\eta \sum_{i=1}^p x_i d_i \tag{1.16}$$

Al fine di evitare un overflow della forza sinaptica il parametro η è diminuito a ogni step:

$$\begin{cases} \eta^n = \alpha^{n-1}\eta^0 \\ \alpha = \{k | k \in (0, 1)\} \end{cases} \quad (1.17)$$

La regola di apprendimento ottimizzata prevede una convergenza più veloce:

$$\Delta w = \frac{p-1}{p}w + \frac{1}{p}\eta xd \quad (1.18)$$

La fase di apprendimento ora consiste in:

$$\begin{aligned} & w^0 \\ w^1 &= \frac{p-1}{p}w^0 + \frac{1}{p}\eta x_1 d_1 \\ & \dots \\ w^p &= \left(\frac{p-1}{p}\right)^{p-1}w^0 + \frac{\eta}{p} \sum_{i=0}^{p-1} \left(\frac{p-1}{p}\right)^i x_{p-i} d_{p-i} \end{aligned} \quad (1.19)$$

Al termine dell' n -esimo ciclo di apprendimento il risultato è:

$$w^{np} = \left(\frac{p-1}{p}\right)^{p-1}w^0 + \sum_{j=0}^{n-1} \left(\frac{p-1}{p}\right)^{jp} \frac{\eta}{p} \sum_{i=0}^{p-1} \left(\frac{p-1}{p}\right)^i x_{p-i} d_{p-i} \quad (1.20)$$

Se la fase di training è terminata durante un numero di *step* $b < n$ i pesi ottenuti non rappresentano valori corretti. La forza sinaptica oscilla attorno alla soluzione finale e rappresenta un risultato valido solo al termine dell' n -esimo ciclo.

L'algoritmo proposto prevede una complessità computazionale ridotta rispetto al plain Hebbian learning. I risultati empirici osservano una fase di convergenza ottimizzata in caso di rete neuronale con elevato numero di input.

1.3.2 Regola di Oja

Erkki Oja, in [27], introduce un nuovo modello di apprendimento che estende la formulazione matematica ricavata dalla teoria di Hebb (1.5).

$$\Delta w_{ij} = \eta(a_i o_j - o_j^2 w_{ij}) \quad (1.21)$$

Δw_{ij} è la variazione della forza sinaptica tra il neurone post-sinaptico j e quello pre-sinaptico i , a_i è il valore di input del neurone (ad esempio il livello di attivazione) potenziato dal parametro w_{ij} , o_j è il segnale di output e η è la costante di apprendimento positiva. Il forgetting ratio modula il cambiamento del peso tramite o_j^2 .

Scritta in termini di vettori (1.21) diventa:

$$\Delta w = \eta(ao - o^2w) \quad (1.22)$$

Assumendo che la funzione di attivazione sia la funzione identità:

$$o = w^T a = a^T w \quad (1.23)$$

Sostituendo (1.23) in (1.22):

$$\Delta w = \eta(aa^T w - w^T aa^T w w) \quad (1.24)$$

L'equazione è formulata in [27] da una versione estesa dell'apprendimento Hebbiano formulato in (1.7) a cui un fattore di normalizzazione è aggiunto:

$$w_{ij}(t+1) = \frac{w_{ij}(t) + \eta a_i o_j}{\{\sum_{k=0}^n [w_{kj} + \eta a_k o_j]^2\}^{1/2}} \quad (1.25)$$

1.3.3 L'importanza dei vincoli nell'apprendimento Hebbiano

Il contenuto di questa sezione è prevalentemente tratto da [26].

A causa dell'apprendimento Hebbiano classico il peso delle sinapsi tende a saturare al valore di upper bound. Il processo di apprendimento guidato da (1.5) può portare all'instabilità di una rete neuronale e a problemi di over-generalization.

L'introduzione di vincoli per regolare il comportamento sinaptico può arginare queste problematiche. Data una semplice equazione per descrivere la plasticità sinaptica:

$$\Delta w = F(w) \quad (1.26)$$

Esistono due categorie principali di vincoli:

1. Moltiplicativi: la forza sinaptica decade in modo proporzionale al proprio stato attuale;

$$\Delta w_{ij} = F(w) - \epsilon w_{ij} \quad (1.27)$$

2. Sottrattivi: la forza sinaptica decade con un valore prefissato moltiplicato per la costante di decadimento k .

$$\Delta w_{ij} = F(w) - \epsilon k \quad (1.28)$$

Considerando una cellula neuronale post-sinaptica, w è il vettore di n pesi sinaptici inizialmente distribuiti in maniera random con valori compresi tra w^{min} e w^{max} . Un possibile vincolo è quello di conservare il valore totale dei pesi sinaptici in maniera sottrattiva o moltiplicativa. In questo modo si realizza il principio di competitività introdotto in precedenza:

$$const = w_{tot} = \sum_{i=0}^n w_i \quad (1.29)$$

w_{tot} rappresenta il valore totale conservato dei pesi sinaptici.

Si consideri F come la funzione Gaussiana e $w^{min} = 0$. Senza vincoli le sinapsi tendono a saturare a w^{max} , a causa di questo andamento la selettività nella risposta della cellula si perde (figura 1.2, colonna uno). Applicare un vincolo moltiplicativo a un gruppo di sinapsi diminuisce la loro forza in modo graduale dal centro alla periferia (figura 1.2, colonna due). Sotto vincoli di natura sottrattiva alcuni pesi sinaptici (approssimativamente $\frac{w_{tot}}{w^{max}}$) tendono a saturare a w^{max} mentre altri decrescono fino a raggiungere il valore w^{min} (figura 1.2, colonne tre e quattro). Questo fenomeno di *sharpening* produce risultati interessanti nelle correlazione di due mappe spaziali.

Diminuire w^{min} sotto il valore 0 non influenza vincoli moltiplicativi o dinamiche non vincolate. Tale modifica accentua lo sharpening della parte positiva del campo ricettivo sottoposto a vincoli sottrattivi, si ottiene lo stesso effetto a seguito di un incremento di w^{max} o decremento di w_{tot} .

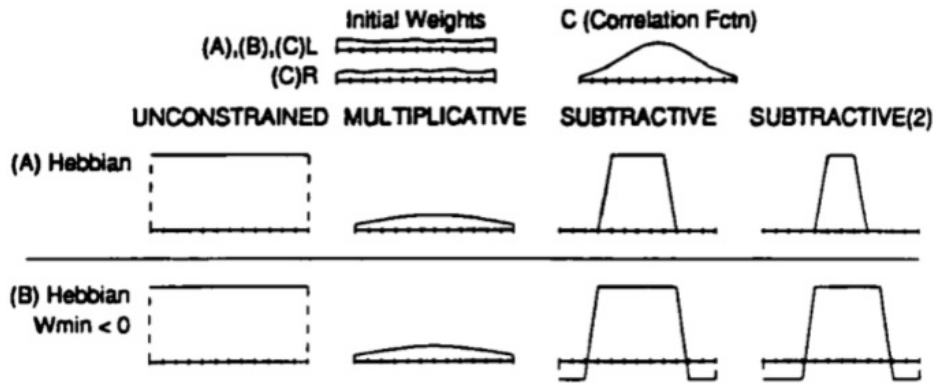


Figura 1.2: Illustrazione presa da [26]. Risultati del processo di apprendimento *libero* e sottoposto a vincoli moltiplicativi e sottrattivi.

In breve, l'apprendimento Hebbiano non condizionato tende a distruggere la selettività della risposta di una cellula neuronale. Vincoli moltiplicativi sviluppano una forza sinaptica graduata in risposta a un input specifico. Vincoli di natura sottrattiva portano a risultati differenti nell'apprendimento lineare in quanto le sinapsi saturano a w^{max} o w^{min} .

Capitolo 2

Value system

Diverse forme di apprendimento sono legate all'abilità di un organismo di percepire il valore adattativo del proprio behaviour [11].

Un value system fornisce vincoli con l'obiettivo di selezionare un comportamento in grado di ottimizzare l'adeguamento alla nicchia ecologica. Esso è un sistema complesso e dinamico, che può evolvere ed estendersi tramite l'esperienza acquisita dall'individuo. Un value system non opera in maniera predefinita: infatti presenta caratteri innati (*bias*) descritti in strutture complesse all'interno del fenotipo dell'organismo.

Nel sistema nervoso sono eseguiti processi di auto-organizzazione per sopperire all'assenza di segnali di *reward* e di un modello esplicito dell'ambiente. Tali meccanismi sono legati all'idea di valore, o segnale valore, e presentano similitudini con la nozione di *fitness* descritta nella teoria evolutiva. Il percorso di adattamento ottimizza tale concetto rispetto all'ambiente in cui l'organismo è *embodied*.

Il valore di un evento è definito in funzione dell'attività che lo stesso provoca all'interno del sistema, e si riflette nella probabilità che il pattern di risposta agli stimoli percepiti si ripresenti nelle medesime condizioni ambientali. In un value system esplicito, quest'ultimo è espresso come un segnale in grado di modulare la forza dei legami sinaptici, al fine di migliorare e indirizzare il processo di adattamento del singolo. A seguito di tale cambiamento, un nuovo comportamento può emergere, provocando a sua volta variazioni di attività nel segnale valore.

In [11] è mostrata una possibile implementazione di value system, la cui attività è calcolata rispetto a segnali provenienti dall'amigdala.

Il sistema-valore, così come le strutture responsabili della fase di apprendimento e memoria, è distribuito e può essere costituito da più sottosistemi. Esso fonda le basi per una forma di apprendimento detta *value-based*. Il value system è ora inteso come una struttura neuronale responsabile della mediazione del processo di apprendimento. Una plasticità sinaptica dipendente dal segnale valore consente l'emergenza di behaviour robusti rispetto a over-generalization e rumore.

2.1 Apprendimento value-based in robotica

L'apprendimento value-based è soggetto di interesse in robotica *brain-based* (BB) e reti neurali artificiali (ANN). Un sistema-valore non determina completamente il comportamento dell'organismo, bensì regola il processo selettivista di auto-organizzazione.

Darwin IV [9] implementa due value system: il primo governa l'apprendimento tra sensori visivi e sistema motore; il secondo influenza i comportamenti riflessi del robot. L'agente è in grado di adattarsi basandosi su esperienza acquisita e *value-scheme*.

Pfeifer, in [31], definisce in modo discorsivo il task di un sistema-valore: *a value scheme (or system) encodes what matters to the robot*. Questa definizione è totalmente generica e si presta a svariati scenari di applicazione.

In [31] è fornito lo spunto per correlare gli studi di Hebb sulle emozioni e l'apprendimento guidato dal value system. Questo argomento è importante nell'interazione tra essere umano e robot. Per motivi di tempo tale tematica non è approfondita, per visionare studi recenti vedi [3].

Il concetto di valore può non essere esplicito. Questo è implementato nella forma di riflessi, architettura di rete, regola di apprendimento e proprietà di sensori e attuatori. In questo caso il sistema-valore è definito dal designer del sistema. Un approccio incrementale può essere seguito per evitare problemi di *over-design* legati alla complessità del mondo reale.

Alcuni modelli di dispositivi brain-based fanno uso pervasivo di value system. Gli stimoli percepiti subiscono un processo di generalizzazione e in seguito, tramite value-based learning, è selezionato il behaviour da attuare. I robot sviluppati secondo tali principi implementano riflessi e comportamenti innati, che ne garantiscono la sopravvivenza iniziale.

2.2 Il principio valore

Il libro Understanding Intelligence [32] dedica un intero capitolo a questo principio:

An autonomous agent has to be equipped with a value system and with mechanisms for self-supervised, incremental learning employing principles of self organization.

Un agente autonomo deve essere in grado di percepire ciò che è giusto o sbagliato.

In un value system esplicito, il segnale valore è generato come conseguenza di un behaviour e regola in modo diretto l'apprendimento. In caso contrario, il processo auto-organizzativo è determinato in termini di riflessi, i quali controllano l'interazione con l'ambiente. Anche la motivazione interna di un robot è realizzata in funzione del segnale valore. Esso imprime una *driving force* necessaria al fine di compiere il task assegnato (ad esempio esplorare l'ambiente per acquisire esperienza) [15].

Driving force e bias codificano in termini informatici la nozione di valore, intendendolo come *a simple, relative low-level valences that can initiate the development process*.

In seguito all'imprevedibilità della nicchia ecologica, è necessario un meccanismo per la generazione di behaviour e uno per la loro selezione. Dato un insieme di comportamenti, è importante individuare quelli in grado di massimizzare l'attività del value system. L'approccio ricorrente consiste nell'esecuzione di behaviour random valutati al loro termine. Questo meccanismo richiede tempi lunghi e spesso non è nemmeno possibile esplorare tutte le possibilità. Una soluzione parziale prevede l'introduzione di bias in forma di riflessi. Biologicamente

quest'ultima è introdotta dalla fase evolutiva [40], mentre in un robot è specificata dal designer.

Nel caso in cui il value system sia troppo specifico, il sistema non è flessibile e non è in grado di generare diversità; in caso contrario può non convergere.

2.3 Apprendimento Hebbiano aumentato

L'auto-organizzazione può essere affrontata con due metodologie diverse: selettivismo e costruttivismo.

- Essa è un processo selettivo dal punto di vista del Darwinismo. *When we think we are learning something, we are only discovering what has already been built into our brains* [13]. Le cellule neuronali sono interconnesse tra loro con grande ridondanza ma solo i collegamenti ottimali sono selezionati e migliorati dal value system.
- Dal punto di vista costruttivista, l'auto-organizzazione è un processo di creazione di nuove strutture a partire da quelle pre-esistenti. I collegamenti sinaptici tra neuroni sono stabiliti se necessario.

Così come il value system, l'apprendimento Hebbiano segue la linea selezionista. Una formula di apprendimento Hebbiano value-based è sviluppata estendendo (1.5) con il segnale valore V :

$$\Delta w_{ij} = V \eta a_i a_j \quad (2.1)$$

V costituisce l'attività esplicita del value system e introduce il concetto di bias all'interno del sistema. Come riportato in [40], l'apprendimento è conseguente all'esecuzione di un'azione da parte dell'individuo.

Un ritardo è costituito dal tempo che intercorre dalla selezione del comportamento, attivazione della rete neuronale, alla fase di valutazione finale. Una forma di memoria, vedi (1.6), è richiesta per memorizzare le attivazioni provocate dal behaviour selezionato.

(2.1) realizza una forma di learning non supervisionato che non prevede segnali di errore o reward. L'apprendimento è guidato dalla prospettiva dell'individuo, condizione necessaria per lo sviluppo di agenti autonomi.

2.4 Confronto tra regole di apprendimento

Verschure, in [50], compara i risultati di 3 diverse regole di apprendimento (denotate con il termine *condizione*):

1. Condizione V:

$$\begin{cases} \Delta w_{ij}(t) = \eta v(t) a_i(t) a_j(t) - \epsilon v(t) w_{ij}(t) a_i(t), \\ v(t) \in [0, 1] \end{cases} \quad (2.2)$$

2. Condizione HA:

$$\Delta w_{ij}(t) = \eta a_i(t) a_j(t) - \epsilon w_{ij}(t) a_i(t) \quad (2.3)$$

3. Condizione HP:

$$\Delta w_{ij}(t) = \eta a_i(t) a_j(t) - \epsilon w_{ij}(t) \quad (2.4)$$

η è la costante di apprendimento, $v(t)$ è l'attivazione media del sistema-valore, $a_i(t)$ è il livello di attivazione del neurone post-sinaptico, $a_j(t)$ è il livello di attivazione del neurone pre-sinaptico, ϵ è il forgetting ratio.

V (2.2) è una forma estesa di value-based learning (2.1) in cui la modulazione della forza sinaptica è legata al segnale valore. HA (2.3) non presenta l'attivazione media del sistema-valore, $v(t)$ è uguale a 1. HP (2.4) estende (1.5). Le tre equazioni presentano w_{ij} nel termine di depressione. Come indicato in 1.3.3, tale vincolo moltiplicativo limita il variare della plasticità sinaptica.

Il task del robot, chiamato NOMAD, consiste nel raccogliere certi blocchi ed evitarne altri.

Il processo di apprendimento è analizzato in termini di pattern di firing, adattabilità del behaviour e pesi sinaptici.

Le regole di apprendimento V e HA portano ad un'elevata probabilità di osservare stabilmente risposte condizionate. HA mostra una frequenza maggiore di CR rispetto a V, il ratio $\frac{CR}{UR}$ è elevato. Differenze di performance sono osservabili studiando la plasticità delle connessioni sinaptiche.

Utilizzando HP, la forza sinaptica cresce rapidamente, tuttavia è presente un'oscillazione tra fasi di depressione e potenziamento. Per questo motivo la possibilità di osservare CR è bassa. Quando un riflesso condizionato è attivato, la depressione sinaptica è maggiore del potenziamento; solamente uno stimolo incondizionato può attivare il comportamento riflesso. Quando ciò accade, il valore di potenziamento è maggiore rispetto a quello di depressione e l'agente è di nuovo in grado di rispondere a stimoli condizionati. In HA il fenomeno di oscillazione è ancora presente.

Con la regola di apprendimento V, il peso delle connessioni cresce in maniera controllata. A dispetto delle altre formulazioni, l'andamento della plasticità sinaptica si mantiene stabile. In HA e HP il cambio di peso è legato solamente alle attività pre e post-sinaptiche.

Il sistema nervoso modifica la propria struttura associativa durante un periodo di tempo mediamente lungo. Esso tende infatti ad aggiustarsi in relazione a pattern di firing che possono indurre ad auto-potenziamento. Tale fenomeno può portare a over-generalization e l'apprendimento avviene anche in caso di rumore.

Rispetto a HA e HP, V è meno sensibile a piccole correlazioni ed è più lento nell'acquisizione di risposte condizionate. La differenza non dipende dai parametri η o ϵ . Il segnale $v(t)$ assume valori compresi nell'intervallo $[0, 1]$ quindi non può amplificare la modulazione della forza sinaptica. Per questo motivo la regola di apprendimento V è robusta: lo sviluppo delle associazioni è ottimizzato limitando l'intervallo di tempo in cui la variazione può avvenire. I value system generano una risposta solamente a fronte del verificarsi di eventi ritenuti utili.

Attuando la fase di apprendimento solamente quando necessario, il value-based learning rappresenta un'estensione valida del plain Hebbian learning.

Capitolo 3

Distributed adaptive control

Il contenuto di questa sezione è prevalentemente tratto da [44].

Secondo Thorndike, gli organismi biologici eccellono nella capacità di adattarsi all'impredicibilità dell'ambiente [41].

Il paradigma *distributed adaptive control* (DAC) fu introdotto da Pfeifer, Verschure e Krose nel 1992 [47]. Proposto nella sua prima versione, cerca di modellare il processo adattativo in funzione del condizionamento Pavloviano. A seguito dell'apprendimento, l'individuo modifica la propria interazione con l'ambiente in modo permanente o evanescente. DAC fornisce un punto di incontro tra teoria del comportamento e sperimentazione scientifica.

La nozione di imprevedibilità può essere applicata in campo somatico e ambientale. Il primo è il risultato della variabilità delle strutture di cervello e corpo (*body*) dell'individuo. Il sistema nervoso si adatta a tale diversità tramite un percorso di auto-organizzazione. L'imprevedibilità ambientale emerge dalla natura del mondo reale, la risposta adattativa del singolo è proprio l'apprendimento. Tale processo si riflette nella probabilità che una risposta (non) condizionata sia causata da uno stimolo condizionato.

Il paradigma DAC si interessa di capire come gli organismi biologici acquisiscano, elaborino ed esprimano la conoscenza appresa. Esso è inizialmente modellato con un'architettura costituita da tre livelli tra loro dipendenti: *reactive control structure*, *adaptive controller* e *reflective control structure*. Reactive control structure implementa i comportamenti riflessi esprimendoli come relazioni pre-cablate tra

stimoli incondizionati (US) e risposte incondizionate (UR). Adaptive controller si occupa dello sviluppo dei componenti del condizionamento classico. Reflective control structure fornisce una rappresentazione degli stimoli condizionati e delle relative risposte. Meccanismi di pianificazione possono essere realizzati sviluppando una successione temporale di tali rappresentazioni.

L'insieme delle tre componenti definisce un sistema di apprendimento completo (*complete learning system*).

Inizialmente DAC è sviluppato per realizzare un modello di apprendimento per rinforzo in grado di correlare CS e UR tramite segnale di reward legato alla risposta non condizionata. Basandosi sul principio di località dell'apprendimento, l'accoppiamento CS-UR è in seguito ottenuto mediante condizionamento classico. Queste nozioni sono codificate nel modello DACI.

Il paradigma è implementato ed esportato all'interno di un robot reale. Quest'ultimo è dotato di tre sensori: distanza (CS), collisione (US-) e target (US+). Reflective controller structure mappa gli stimoli incondizionati US+ e US- sull'insieme CS, portando alla formazione di comportamenti riflessi responsabili della risposta ad una collisione e dell'inseguimento del target.

Gli stimoli non condizionati scatenano attività all'interno di assembly neuronali. L'attivazione di questi ultimi riflette lo stato interno (*internal state*, IS) del robot. Un meccanismo *winner-take-all* garantisce, in base a tale valore, che il robot non esegua contemporaneamente behaviour incompatibili tra loro. Nel caso in cui nessuno degli assembly sia attivo, l'agente esegue un comportamento di default per proseguire l'esplorazione della nicchia ecologica.

Il processo Hebbiano guida l'associazione IS-CS. I primi esperimenti mostrano problemi di over-generalization, l'utilizzo del plain Hebbian learning è così esteso con un termine di depressione.

Nella seconda versione del modello (DACII) è introdotta un'unità inibitoria.

DACIII è sviluppato per esplorare le proprietà della reflective control structure. Il controllore introdotto nell'ultima versione estende quello implementato in DACII aggiungendo nuovi componenti. Esso comprende un'unità di memoria a breve termine (*short term memory*, STM) in cui è memorizzata l'associazione CS-CR. Tale forma di

memoria è realizzata tramite un *buffer* dinamico. In seguito alla percezione di un evento, una o più coppie sono spostate nell'unità di memoria a lungo termine (*long term memory*, LTM). Un meccanismo winner-take-all preleva un elemento da LTM, lo esegue ed in seguito lo reinserisce in STM.

I risultati della sperimentazione dimostrano che DACIII sia un sistema di apprendimento completo. Esso è in grado di acquisire stimoli CS, accoppiarli con le opportune risposte ed elaborarli al fine di adattarsi all'ambiente [44].

3.1 Teoria della connessione tra mente, corpo e cervello

Verschure, in [45], interpreta il cervello come mediatore tra eventi interni, corpo e ambiente in cui l'organismo vive. Tale teoria, denominata *mind, brain, body nexus* (MBBN), è guidata dalle nozioni di behaviour, anatomia e psicologia. Essa cerca di esplorare le proprietà fisiologiche del cervello e di capire come esse influenzino i comportamenti emergenti.

Il modello ottenuto è validato mediante l'implementazione in robot reali, in quanto per agire il cervello ha bisogno di un corpo.

L'approccio proposto è detto H4W: un agente deve rispondere a quattro domande per specificare come (*How?*) sopravvivere. *Why?*: motivazione espressa in termini di driving force e bias. *Where?*: definire la nicchia ecologica. *When?*: timing delle azioni. *What?*: definire i componenti *embodied* nell'ambiente.

MBBN cerca di dare risposta ad ognuno degli interrogativi e propone un'architettura descritta in termini di natura dei processi interni all'agente, della loro interazione, dei protocolli comunicativi e della struttura di controllo.

Il modello proposto estende quello descritto all'inizio del capitolo. Esso comprende quattro livelli: *soma*, *reactive*, *adaptive* e *contextual*; vedi figura (3.1).

Soma descrive il corpo dell'agente, definisce i suoi bisogni (*needs*), le sue sensazioni e i meccanismi per il controllo di attuatori e sensori. Reactive implementa i comportamenti riflessi mediante *sensory-*

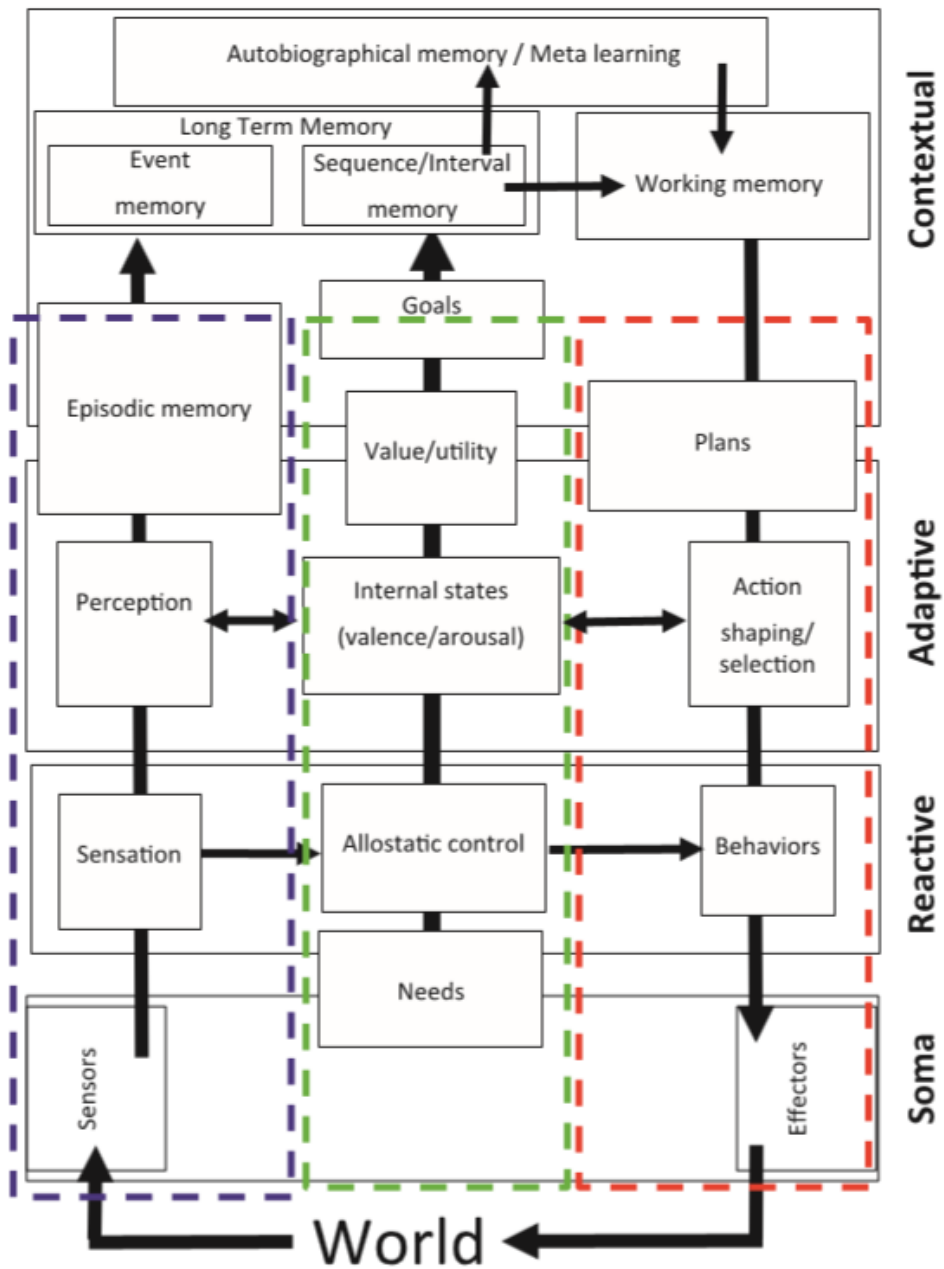


Figura 3.1: Illustrazione presa da [45]. Diagramma dell'architettura DACIII. Colonna blu: sensazione e percezione del mondo, colonna verde: percezioni di se stessi (*self*), colonna rossa: interfaccia con l'ambiente esterno.

motor loop. Questo componente è simile all'architettura di sussunzione proposta da Brooks [5] e prende spunto dai robot di Braitenberg [2]. Reactive comprende anche l'unità di inibizione responsabile dei meccanismi winner-take-all. Adaptive estende il livello precedente occupandosi della fase di apprendimento. Quest'ultima è tipicamente gestita tramite un'estensione del plain Hebbian learning. DACIII utilizza una regola detta *Correlative Subspace Learning* (CSL) basata su una forma di apprendimento Hebbiano predittivo, vedi [45] per ulteriori dettagli. L'ultimo livello, contextual, estende la componente adaptive ed è in grado di pianificare successioni di azioni per realizzare le politiche interne al robot. Tali azioni sono espresse come coppie CS-CR atomiche.

Inizialmente il behaviour dell'agente è governato dalla componente reattiva del livello reactive, con il passare del tempo il controllo passa al contextual layer.

3.2 Architettura DAC

Negli anni passati il campo dell'intelligenza artificiale è stato criticato per la mancanza di processi di apprendimento generici [34].

Le reti neurali consentono di sviluppare proprietà interessanti come meccanismi di generalizzazione e *fault tolerance*. Questi sistemi, se implementati a seguito della definizione di un'ontologia predefinita dal designer, soffrono di problemi quali mancanza di *situatedness* e *(symbol) grounding*.

Come proposto da Brooks, tali problematiche possono essere affrontate costruendo agenti autonomi. Studi approfonditi dimostrano che realizzare robot basati sul modello *sense-think-act* non sia sufficiente, occorrono ulteriori meccanismi per l'acquisizione e l'elaborazione di conoscenza.

Il paradigma DAC combina la nozione di embodiment e la flessibilità delle reti neurali. Questo approccio porta allo sviluppo di sistemi adattativi e si basa su quattro assunzioni:

1. Il mondo reale è noto parzialmente ed è parzialmente predicibile. Tale assunzione indica perché sia necessario apprendere. È

possibile fare previsioni su condizioni future anche se in maniera limitata;

2. Il mondo non consiste in una collezione di oggetti ed eventi a sé stanti, essi sono definiti in relazione all'interazione tra agente e ambiente;
3. Gli input del sistema consistono in uno spazio continuo di stati sensoriali;
4. L'agente è sottoposto a vincoli di natura real time quindi deve agire in maniera tempestiva.

Il paradigma è definito in quattro passi e l'intero processo di sviluppo deve essere centrato sul robot.

- È necessario definire gli elementi che costituiscono il corpo del robot e l'ambiente. Questo setup fisico è descritto in termini di sensori e attuatori e determina le modalità di interazione con la nicchia ecologica.
- Come riportato in 2.2 un agente autonomo deve essere in grado di valutare i propri stimoli e behaviour [31]. Definire il sistema-valore è fondamentale.

Nella prima versione dell'architettura DAC è utilizzata una nozione di valore implicito. Il value system è predefinito e individua i vincoli regolatori dell'auto-organizzazione. Esso definisce i riflessi sense-act, i parametri interni del robot e le proprietà dei sensori e del sistema motore.

Il sistema-valore rappresenta il supervisore del robot. Sebbene il processo di apprendimento sia di tipo Hebbiano, quest'ultimo è supervisionato dalla risposta del value system al verificarsi di eventi esterni. In questo modo il robot impara unicamente dal proprio punto di vista.

Inizialmente i riflessi determinano completamente il comportamento del robot. L'integrazione di sensori più sofisticati porta ad un processo di adattamento, tale sviluppo può essere visto come categorizzazione dell'interazione.

- Definire l'architettura della rete neuronale. Il modello utilizzato deriva dal condizionamento Pavloviano. Le assunzioni relative al value system sono codificate in una particolare struttura. Ogni sensore è connesso al nodo del layer neuronale appropriato.

È importante definire la tipologia del processo di apprendimento tramite cui il controllo del sistema motore passa dai riflessi a sensori più sofisticati. Acquisendo categorie basate sulle azioni, l'agente impara ogni volta che un'azione è portata a termine. Non esiste distinzione tra fase di apprendimento e di esecuzione. Se tale discriminazione esistesse, il robot dovrebbe decidere quando agire e quando apprendere. Tale assunzione non tiene conto dell'imprevedibilità ambientale e della possibile tempestività dei cambiamenti della nicchia ecologica. Un apprendimento continuo è richiesto.

Le strutture neuronali, che consentono di associare informazioni sensoriali e azioni, rappresentano meccanismi generici per la raccolta di conoscenza. Qualsiasi elemento acquisito sarà fondato (grounded) sull'esperienza dell'agente.

DAC consente di effettuare passi avanti negli studi legati al processo di grounding (il termine symbol è volontariamente omissso).

- Analizzare l'interazione tra agente e ambiente secondo diversi punti di vista [8]. In questo modo è possibile correlare il comportamento del sistema alle dinamiche correnti nella rete neuronale.

Definita la nicchia ecologica, predire il comportamento del robot non è semplice. A seguito del percorso di adattamento sono introdotti elementi di controllo inesistenti a tempo di design. Questi ultimi emergono dall'interazione con l'ambiente.

Se lo sviluppo segue una linea robot-centrica allora l'agente fa propria la nozione di situatedness. Esso è in grado di agire sfruttando la propria esperienza.

Il robot è soggetto a vincoli di natura real time, l'interazione diretta minimizza la capacità di computazione richiesta e quindi i tempi di calcolo.

Il suo comportamento è reattivo ad eventi esterni sconosciuti e sviluppa caratteri pro-attivi in grado di anticipare scenari noti. L'esperienza può portare a maggiore robustezza aumentando la probabilità che il robot agisca correttamente in situazioni non previste dal designer.

La prima versione del modello, presentata nel 1992, mostra diverse limitazioni se messa a confronto con le evoluzioni DACII e DACIII. Nella versione riportata ogni nodo è legato ad attività sensoriali o ad azioni del sistema motore. Meccanismi di alto livello come STM, LTM sono assenti. Il sistema valore non considera i bisogni dell'agente, inoltre non vi è traccia di emozioni.

L'apprendimento si basa unicamente su un meccanismo associativo Hebbiano. Esso rappresenta implicitamente l'interazione sistema-ambiente in termini di coppia stimulus-response.

3.3 Comportamenti orientati al goal

Questa sezione si riferisce all'articolo [35].

Codificare il concetto di task richiede solitamente un processo guidato dalla pianificazione. Le azioni di un robot sono organizzate attorno al compito assegnato.

Secondo Newell: *intelligent behaving systems require goals*. Il goal stesso è esplicitamente rappresentato all'interno del sistema. Al fine di elaborare piani d'azione è necessario possedere a priori un modello dettagliato dell'ambiente.

DAC fornisce un approccio alternativo. Pfeifer e Verschure realizzano un robot in grado di imparare ad evitare ostacoli e localizzare il target presente nell'ambiente. Ulteriori dettagli saranno forniti nel capitolo successivo.

Il robot è equipaggiato con sensori di prossimità, collisione e target. Esso può eseguire cinque diverse azioni: gira a sinistra, gira a destra, gira a sinistra in senso opposto, gira a destra in senso opposto, avanza. Quest'ultima è attiva quando nessuna delle precedenti è in esecuzione.

La struttura di controllo prevede quattro layer neuronali e un'unità inibitoria. Il layer RF (*range finder*) è composto da 37 nodi, ognuno dei quali è connesso ad un sensore di prossimità. Anche il layer di colli-

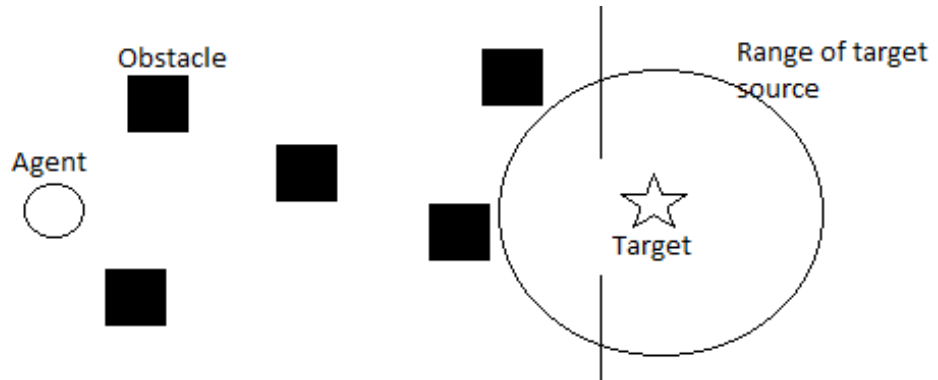


Figura 3.2: Esempio di ambiente descritto in [35].

sione (C) contiene 37 unità connesse ad altrettanti sensori di contatto. Il Target layer (T) comprende due nodi attivati dai sensori di target. L'ultimo gruppo, M, codifica le azioni che il robot può eseguire. Una collisione sulla sinistra provoca la risposta *gira a destra in direzione opposta*. Se il target è individuato sulla sinistra del robot è attuata l'azione *gira a destra*. Un nodo in C (e T) può eccitare una sola unità in M. L'unità inibitoria (I) regola l'attività di C e T.

Quando il livello di attivazione nel layer di collisione supera una soglia prefissata, l'output da T a M è inibito. In questo modo è implementato un istinto di sopravvivenza primordiale: evitare un ostacolo è prioritario rispetto ad inseguire l'obiettivo. L'attivazione di ogni unità all'interno del sistema prevede unicamente valori non negativi. L'apprendimento, influenzato dal value system, è espresso tramite potenziamento dei legami RF-C e RF-T.

L'ambiente utilizzato nell'esperimento è riportato in figura (3.2).

Inizialmente il robot collide con gli ostacoli con un'alta probabilità. I riflessi *gira a sinistra in senso opposto* o *gira a destra in senso opposto* sono attuati di conseguenza. A fronte del verificarsi della collisione, l'apprendimento Hebbiano associa le unità interessate dei layer C e RF. Con il passare del tempo l'agente tenderà ad anticipare gli ostacoli evitandoli senza avervi prima urtato.

Quando l'agente percepisce la presenza del target inizia ad avvicinarsi. Se quest'ultimo è localizzato dietro ad una fessura del muro, nel robot emerge un comportamento che lo porta a fiancheggiare la

parete.

Un osservatore esterno al sistema può considerare che l'agente stia perseguendo il goal di seguire il perimetro della stanza. Tale fenomeno è legato ad emergenza e auto-organizzazione: comportamenti orientati al goal possono quindi emergere dall'interazione con l'ambiente.

Capitolo 4

WABOT

Il cammino si conclude con questo capitolo, descrivendo il processo produttivo per la realizzazione del caso studio proposto da Pfeifer e Verschure in [35].

Nel corso degli anni, l'approfondimento degli studi in campo psicologico e neurofisiologico ha portato ad un *transfer* di conoscenza nelle reti neurali artificiali. La semplicità di applicazione dell'apprendimento Hebbiano ha consentito di sviluppare sistemi ispirati al modello biologico. Ciò mi ha permesso di implementare una versione, seppur semplificata, dell'architettura DAC.

Pfeifer, Verschure, Versace e Hebb sono figure di riferimento che hanno guidato lo sviluppo di questo progetto. Parte del materiale informativo e delle illustrazioni è stato preso dai seguenti articoli scientifici: [49], [34], [45], [50], [43], [17], [48] e [37]. Di particolare importanza è il libro *Understanding Intelligence* [33]. Qualsiasi forma di errore interpretativo è da rimandare unicamente al sottoscritto.

Il processo produttivo cerca di seguire il modello di sviluppo, approfondito nel corso di ingegneria del software, noto come *Capability Maturity Model* (CMM). In questo capitolo, tuttavia, non è riportata la documentazione tecnica del progetto. In seguito sono enfatizzate le fasi di analisi guidate dai principi DAC. Parte fondamentale del progetto è la stesura dei requisiti: formularli in modo chiaro e preciso è essenziale per mettere a fuoco il problema vincolandone i gradi di libertà. Una successiva fase di analisi porta alla definizione del modello del dominio, *use case* e scenari. La formulazione di tali artefatti

fornisce la linea guida per implementare le funzionalità dell'agente. Questo è il punto di partenza per un'ulteriore analisi che porta alla produzione di un'architettura logica in termini di struttura, interazione e comportamento dei componenti. Individuato il problema, sono indicati *abstraction gap* e piano di lavoro.

Una volta implementata, la rete neuronale sarà esportata all'interno di WABOT, un robot costruito con tecnologia Lego Mindstorm.

Gli esperimenti riportati in [35] sono parzialmente descritti in 3.3. Al fine di scrivere un capitolo auto-descrittivo, alcune nozioni sono riprese e approfondite.

Articoli scientifici e il libro *Understanding Intelligence* fanno talvolta uso di notazioni differenti. Laddove necessario questa differenza è resa nota al lettore.

4.1 Requisiti

WABOT è un robot mobile equipaggiato con quattro sensori passivi: due di collisione, uno di prossimità e uno di target. Se interrogato, il sensore di collisione restituisce 1 nel caso in cui sia premuto, altrimenti 0. Il sensore di prossimità, utilizzando tecnologia ad ultrasuoni, fornisce una misura compresa tra 0 (ostacolo vicino) e 254 (ostacolo lontano). Il valore 255 indica che nessun oggetto sia rilevato entro la distanza massima. Il sensore di target ritorna un segnale compreso tra 0 (target non rilevato) e 100 (target rilevato con intensità massima).

WABOT dispone di due motori indipendenti, A e B, internamente dotati di giroscopio. A controlla la ruota di sinistra, B quella di destra. Essi sono pilotabili secondo due modalità diverse: indicando la velocità di crociera o i gradi di rotazione.

Come indicato in 3.3, l'agente è in grado di compiere cinque azioni: gira a destra ($v_A > v_B = 0$), gira a sinistra ($v_B > v_A = 0$), gira a destra in senso opposto ($v_B < v_A = 0$), gira a sinistra in senso opposto ($v_A < v_B = 0$), avanza ($v_B = v_A > 0$). Quest'ultima è eseguita quando nessuna delle precedenti è attiva.

All'interno dell'ambiente in cui il robot si muove sono posizionati casualmente oggetti di natura e forma diversa, vedi figura (3.2). Al verificarsi di una collisione sul lato sinistro, WABOT attua la risposta non condizionata gira a destra in senso opposto. Caso simmetrico è quello di un urto rilevato nella parte destra del robot. Quando percepisce il target, il robot vi si avvicina.

Il compito dell'agente è quello di raggiungere la fonte di cibo, target individuato da una un'entità luminosa, cercando di colpire il minor numero di ostacoli.

Il robot è realizzato con tecnologia Lego Mindstorm, implementando all'interno dell'unità computazionale (NXT) l'architettura DAC.

4.2 Analisi dei requisiti

DAC controlla il robot per mezzo di una rete neuronale. Prima di procedere è bene capirne struttura e funzionamento.

Nel campo dell'intelligenza artificiale una rete neuronale è un modello di computazione ispirato dal sistema nervoso centrale [18]. Problemi come analisi di mammografie [39], previsioni dei mercati finanziari [1], studio della struttura delle proteine [20], sono stati affrontati in modo efficace grazie alla natura di queste reti. Parallelismo massivo, apprendimento, fault tolerance e bassi consumi energetici, sono caratteristiche chiave delle ANN.

A causa della complessità, questo campo richiede il contributo di discipline tra cui neurofisiologia, psicologia e statistica. Inoltre è difficile eseguire manualmente il *tuning* dei parametri sinaptici (se non in caso di reti di dimensioni ridotte). L'apprendimento è fondamentale per l'addestramento e auto-organizzazione del sistema.

McCulloch e Warren implementarono un modello di intelligenza artificiale [25] [36]. Il loro elaborato, fondato sulla teoria della computazione di Turing e fisiologia della cellula neuronale, propone un modello di neurone con soglia binaria; vedi figura (4.1).

Le reti neurali possono essere viste come alberi o grafi pesati: le cellule rappresentano i nodi mentre i loro legami sinaptici identificano gli archi orientati.

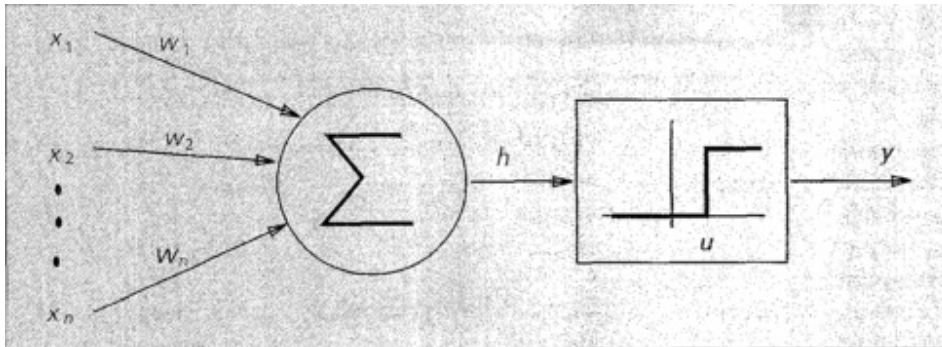


Figura 4.1: Illustrazione presa da [18]. Modello di neurone formulato dalla coppia McCulloch-Pitts. Quando la sommatoria dei segnali in ingresso supera una soglia prefissata il neurone si attiva.

Questo caso studio riduce la complessità di tali sistemi effettuando alcune semplificazioni:

1. L'attivazione del neurone non tiene conto dei tempi di arrivo dei segnali in ingresso. Quando l'input della cellula varia, il valore di output è aggiornato in funzione della sommatoria pesata degli ingressi. Tale segnale è espresso come un'unità informativa atomica e non come un treno di impulsi;
2. Lo scambio informativo tra neuroni è istantaneo;
3. La plasticità sinaptica non è influenzata dal timing dei segnali pre e post-sinaptici, sono trascurati processi STDP;
4. Si trascura l'esistenza di neuroni diversi e di diverse forme di interconnessione. L'informazione è trasmessa solamente in forma di spike;
5. Il valore di output può essere ricondotto direttamente al livello di attivazione della cellula ($o_j = a_j$).

Per definire una rete neuronale artificiale è opportuno seguire cinque step [33]:

1. Caratteristica del neurone: descrivere la funzione di attivazione e come il risultato, da essa restituito, sia correlato al valore di output;

2. Connettività: come i neuroni sono interconnessi tra loro e la direzionalità del collegamento;
3. Regola di propagazione: come l'output della cellula è trasmesso;
4. Regola di apprendimento: vedi Capitolo 1;
5. *Embedding*: come la rete neuronale è esportata all'interno del robot.

È quindi riportato il modello del dominio, vedi figura (4.2). In tale artefatto UML sono riportate le interfacce dei componenti del sistema. È possibile individuare tre raggruppamenti:

1. Struttura della rete: insieme dei componenti che costituiscono cellula, livello e rete neuronale. Neuroni con le stesse funzionalità formano un livello, il sistema è costruito a partire dall'interconnessione di più layer. L'intera architettura si sviluppa attorno al pattern *observer*;
2. Informazione: l'informazione è rappresentata in forma di segnali fisici. Livello di attivazione e spike specializzano un segnale generico;
3. Strutture connettive: interfacce responsabili dello scambio informativo. Ogni sinapsi presenta un input (assone) e un output (dendrite), la plasticità è modellata con *IWeight*. La regola di apprendimento è legata alla funzione *learning()*, è possibile modificarla specializzando *ISynapse*.

4.2.1 Modello del dominio

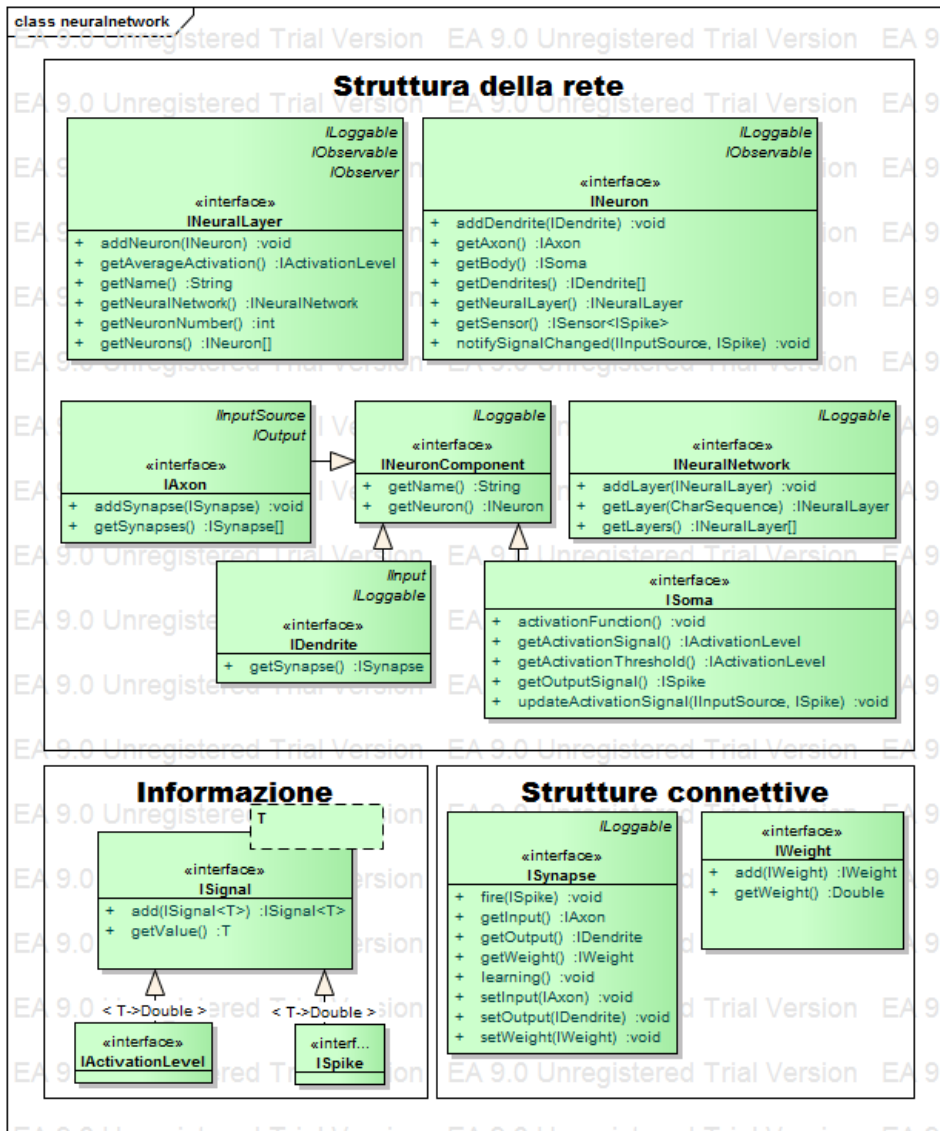
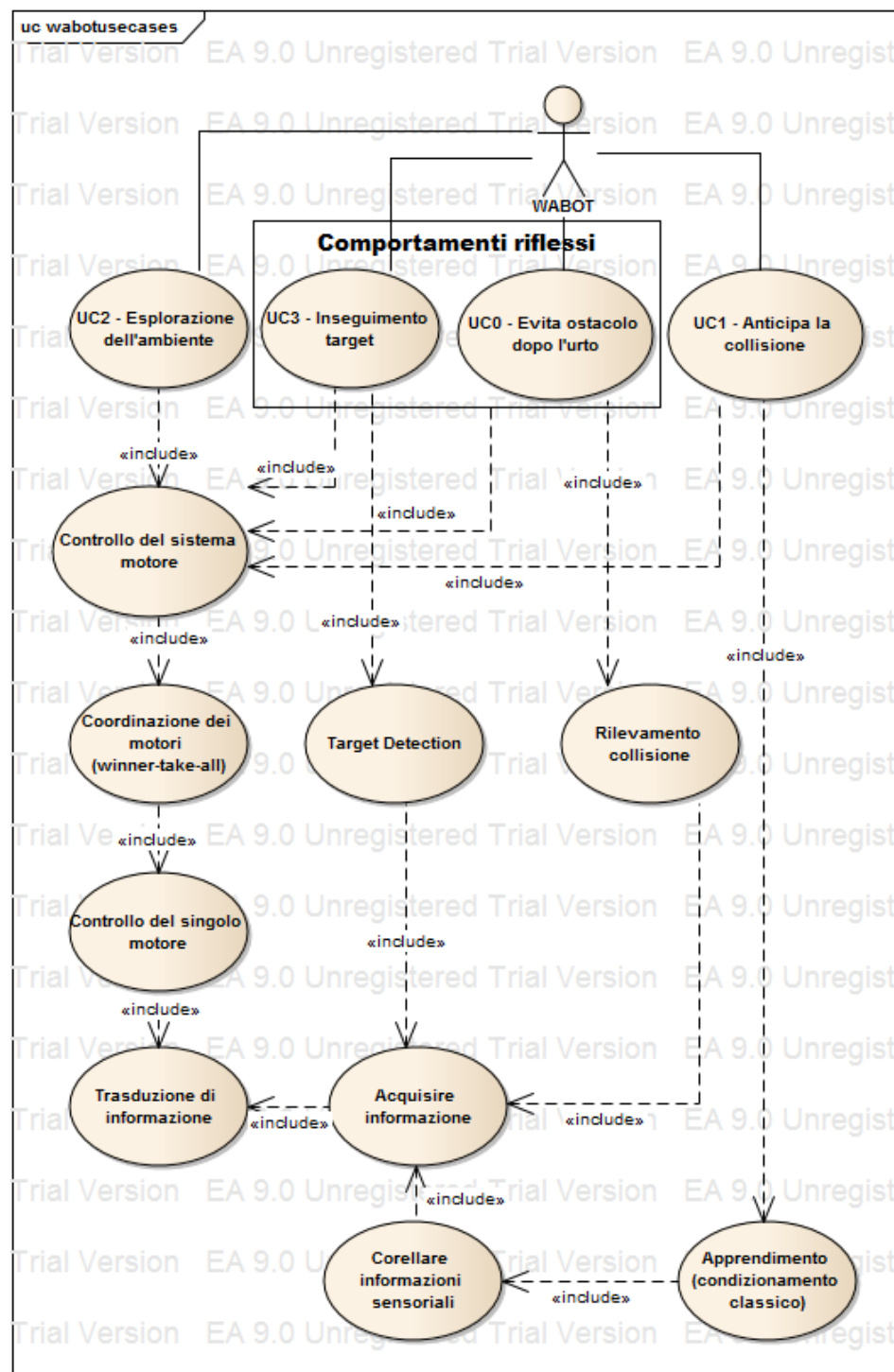


Figura 4.2: Modello del dominio semplificato di una rete neuronale biologica.

4.2.2 Casi d'uso



4.2.3 Scenari

Ogni Use Case riguarda uno o più scenari che indicano come il robot debba interagire con l'ambiente al fine di perseguire specifici obiettivi. Uno scenario descrive brevemente le funzionalità dell'agente e un suo ipotetico behaviour.

In seguito sono mostrati brevi spaccati del funzionamento del robot, osservazioni empiriche sono riportate al termine del capitolo.

UC0 - Collisione con ostacolo

Descrizione: Nel corso dell'esplorazione della nicchia ecologica, WABOT collide con gli ostacoli presenti all'interno dell'ambiente.

Attori principali: WABOT.

Attori secondari: Ostacoli presenti nell'ambiente.

Condizioni precorse: nessuna fase di apprendimento è preceduta in WABOT.

Corso principale: WABOT collide con un ostacolo, se l'ostacolo si trova alla sua sinistra esso esegue l'azione gira a destra in senso opposto, altrimenti gira a sinistra in senso opposto. Se l'ostacolo è sufficientemente grande da essere percepito dal sensore di prossimità, un processo di apprendimento Hebbiano rafforza l'efficienza della connessione che lega i neuroni pre e post-sinaptici interessati dalla collisione.

Condizione di successo: WABOT evita l'ostacolo.

UC1 - Anticipo di collisione con ostacolo

Descrizione: Nel corso dell'esplorazione della nicchia ecologica, WABOT anticipa la collisione con gli ostacoli presenti all'interno dell'ambiente.

Attori principali: WABOT.

Attori secondari: ostacoli presenti nell'ambiente.

Condizioni precorse: un processo di apprendimento è precorso in WABOT.

Corso principale: dopo aver ripetutamente colliso con gli ostacoli, WABOT impara ad anticiparli ed evitarli.

Condizione di successo: l'interazione tra robot e ambiente cambia, ora WABOT è in grado di anticipare gli oggetti.

Condizione di fallimento: se un ostacolo non è rilevato dal sensore di prossimità, WABOT non è in grado di anticipare la collisione.

UC2 - Esplorazione della nicchia ecologica

Descrizione: WABOT avanza per esplorare la nicchia ecologica alla ricerca di cibo.

Attori principali: WABOT.

Condizioni precorse: WABOT è in funzione.

Corso principale: WABOT esplora la nicchia ecologica alla ricerca di cibo fino a quando un ostacolo è colpito o è percepito il gradiente di intensità del target.

UC3 - Inseguimento del target

Descrizione: nel corso dell'esplorazione della nicchia ecologica, WABOT percepisce la presenza del target.

Attori principali: WABOT.

Attori secondari: target presente nell'ambiente.

Condizioni precorse: nessuna fase di apprendimento è preceduta in WABOT.

Corso principale: WABOT percepisce il target all'interno dell'ambiente. Se quest'ultimo si trova alla sua sinistra, il robot gira a sinistra, altrimenti a destra. Quando il sensore di prossimità rileva un ostacolo davanti al robot, il processo di apprendimento Hebbiano aumenta la forza sinaptica che lega i neuroni pre e post-sinaptici interessati.

Condizione di successo: l'interazione tra robot e ambiente cambia, WABOT tende ad inseguire il target.

4.3 Analisi del problema

Per affrontare al meglio la fase di analisi, è utile definire un percorso che consenta di aggredire il problema minimizzando *l'abstraction gap* il più possibile.

Il paradigma DAC è funzionale all'evasione dei requisiti. In riferimento a questo caso studio occorre definire i quattro passi proposti dall'architettura. Tre di essi sono sviluppati all'interno di questa sezione. Le osservazioni sull'esperimento seguono la fase di implementazione, per questo motivo sono riportate in 4.5.

4.3.1 WABOT e la nicchia ecologica

WABOT è un robot cingolato provvisto di due motori indipendenti che gli consentono di esplorare la nicchia ecologica. La tecnologia Lego Mindstorm consente di fare uso di un numero totale di quattro sensori e tre motori. I dettagli tecnici dei componenti sono rimandati al sito ufficiale (<http://mindstorms.lego.com>).

L'agente presenta due sensori di collisione: uno alla sua sinistra e uno alla sua destra. I sensori di luminosità e di prossimità sono posizionati sull'asse di simmetria del robot. La localizzazione del target è legata alla percezione del gradiente di intensità emesso dallo stesso: tanto più il robot è vicino, tanto più la luce è intensa.

La costruzione del robot segue i principi descritti da Pfeifer in [31]. La linea guida proposta, indicata come *Understanding by building*, segue un percorso di sintetizzazione: partendo dai singoli componenti si procede con la costruzione di WABOT. In questo modo, il designer

conosce tutti gli elementi del sistema ed è in grado di sperimentare alternative a sua discrezione. Fattore principale è quello del costo (*principle of cheap design*): se esistono più vie per creare il robot, a parità di performance è bene scegliere quella con il prezzo inferiore. Utilizzare sensori complessi e costosi non è necessario nel caso in cui l'agente non sia in grado di sfruttarli.

La nozione *cheap* si applica anche in termini di programmazione: né tutti gli scenari devono essere codificati, né un sistema di apprendimento universale è richiesto. Se il compito del robot è quello di evitare gli ostacoli, apprendere anche la funzione XOR può non essere necessario.

Modello differenziale

Implementare un modello differenziale, come descritto in [19], non è possibile se si intende utilizzare sia il sensore di prossimità che quello di target, vedi figura (4.3). La spiegazione è semplice: il numero di ingressi non è sufficiente per ospitare un ulteriore misuratore di luminosità. Per ovviare a questo problema, si è cercato di approssimare il modello differenziale utilizzando un solo sensore. Ipotizzando una direzione iniziale del robot (ad esempio sinistra), si campiona il segnale *input*. Se $input(t + 1) \geq input(t)$ la direzione logica del robot è mantenuta invariata, viceversa è invertita.

In teoria, non c'è differenza tra teoria e pratica; in pratica la differenza esiste e si vede (Yogi Berra). Tale soluzione è fortemente limitativa: se la direzione logica coincide con quella reale, il robot si avvicina al target; in caso opposto accade il contrario.

Mentre questa scappatoia è parzialmente valida per il sensore di target, applicare lo stesso meccanismo al sensore di prossimità è irrazionale. Quest'ultimo controlla l'attività di entrambi i neuroni legati ai rilevatori di collisione. La struttura del robot deve quindi rispettare un vincolo preciso: i due sensori devono essere vicini al fine di essere entrambi coperti dall'angolo di apertura degli ultrasuoni.

Queste limitazioni influenzano il behaviour di WABOT ma sono accettabili a seguito ai vantaggi che Lego Mindstorm comporta. Tale tecnologia consente di sviluppare in modo rapido il robot e si presta ad altissima flessibilità, componibilità e modellabilità. Inoltre, il firmware

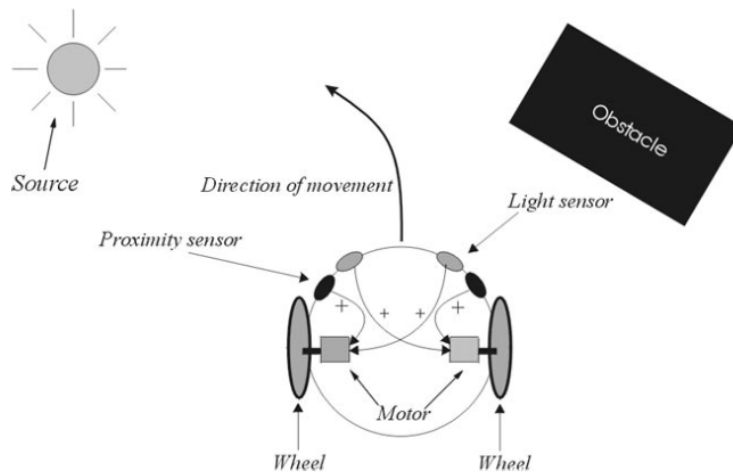


Figura 4.3: Illustrazione presa da [33]. Schema del modello differenziale di un veicolo di Braitenberg.

LeJos, fornendo procedure di alto livello, consente di astrarre dalle dinamiche dei protocolli comunicativi di sensori e attuatori.

4.3.2 Value system

Il value system pervade implicitamente la struttura della rete neurale. Esso costituisce la base di conoscenza di WABOT e ne definisce i parametri di funzionamento.

Sistema motore

I motori sono pilotabili tramite funzioni logiche:

1. *rotate(degrees)*: il motore esegue una rotazione del valore di *degrees* e si ferma. *degrees* $>$ 0 il motore ruota in senso *forward*, altrimenti *backward*.
2. *setSpeed(speed)*: consente di pilotare il motore specificando la velocità di crociera. *speed* $>$ 0 il motore ruota in senso *forward*, altrimenti *backward*. *speed* = 0 ferma il motore se in movimento.

3. *stop*: ferma il motore se in movimento, equivale a *setSpeed(0)*.

La gestione degli attuatori richiede, in seguito al parallelismo delle reti neurali, un meccanismo winner-take-all per evitare che comportamenti in conflitto tra loro entrino in corsa critica (il robot non può andare avanti e indietro contemporaneamente). A tal fine, un componente *monitor* garantisce che in un certo istante di tempo t solo un behaviour sia attivo. Se due comportamenti competono per ottenere il controllo dei motori, non è possibile predire in modo deterministico quale dei due sia attuato.

Si è deciso di non implementare uno *stack* che accodi i behaviour non eseguiti. Si consideri, ad esempio, uno scontro frontale in cui entrambi i sensori di collisione sono premuti. Si assuma per ipotesi che essi attivino simultaneamente i comportamenti riflessi. Al termine di una delle due risposte non condizionate, è possibile che il robot non stia più collidendo con l'ostacolo. Ne risulta che attuare il comportamento riflesso non eseguito in precedenza sia errato: se così fosse, il robot starebbe rispondendo ad uno stimolo che non è più percepito. Ogni richiesta di esecuzione è scartata fino al completamento della risposta in corso.

Comportamenti riflessi e motivazione

I riflessi previsti sono quattro e possono essere riassunti in due categorie: collisione e inseguimento.

Non avendo forma circolare, se al verificarsi di una collisione frontale WABOT attuasse la risposta gira a destra in senso opposto, è probabile che il sensore sinistro rimanga premuto. A seguito di tale pressione, il robot risponde girando a sinistra in senso opposto trovandosi nuovamente di fronte all'oggetto (con la possibilità di un nuovo urto frontale). Per questo motivo le risposte alla collisione sono estese in una sequenza di azioni descritte con il seguente pseudo-codice.

Gira a destra in senso opposto:

1. *A.rotate(-180)*
2. *B.rotate(-180)* //WABOT retrocede distanziandosi dall'ostacolo...

3. *B.rotate(-180)* //... e poi gira a destra in senso opposto.

Gira a sinistra in senso opposto:

1. *A.rotate(-180)*

2. *B.rotate(-180)* //WABOT retrocede distanziandosi dall'ostacolo...

3. *A.rotate(-180)* //... e poi gira a sinistra in senso opposto.

I comportamenti riflessi legati all'approccio del target rimangono invece invariati. Segue una descrizione in pseudo codice.

Gira a destra:

1. *B.stop*

2. *A.rotate(180)*

Gira a sinistra:

1. *A.stop*

2. *B.rotate(180)*

Sistema dei sensori

I segnali in ingresso al sistema sono tradotti in valori logici. Questa funzione, implementata da una famiglia di *transducer*, è essenziale per gestire l'eterogeneità degli input. Come descritto nei requisiti, i tre gruppi di sensori generano range di valori diversi e di diversa natura. I sensori di collisione sono descritti da un modello binario, mentre quelli di target e prossimità discretizzano i loro stati continui in un intervallo di valori interi. In fase implementativa tale *mapping* non è a carico del programmatore, LeJos consente di leggere in modo diretto la misura discretizzata.

Per uniformare il range dei valori in ingresso, gli input sono normalizzati:

$$\begin{cases} \text{valoreNormalizzato} = \frac{\text{input} - \text{input}_{\min}}{\text{input}_{\max} - \text{input}_{\min}} \\ \text{input}_{\min} \geq 0 \\ \text{input}_{\max} > \text{input}_{\min} \end{cases} \quad (4.1)$$

$input$ è il risultato del campionamento del sensore, $input_{max}$ è il valore di upper bound, $input_{min}$ è il valore di lower bound. A seguito di tali trasformazioni, il sensore di contatto continua a fornire un valore binario, mentre quello di target emette un segnale $t \in [0, 1]$. t tende a 0 se il target non è rilevato o è distante, $t = 1$ se il gradiente è percepito alla massima intensità. Il valore normalizzato d della distanza rilevata dal sensore di prossimità è a sua volta compreso in $[0, 1]$. Tale segnale individua la lontananza dell'ostacolo dal robot: tanto più il valore è alto, tanto più l'oggetto è distante.

Il value system deve attivarsi in seguito alla presenza di ostacoli vicini e trascurare quelli più lontani.

$$probabilita' = \frac{1}{e^{kd}} \quad (4.2)$$

k individua il coefficiente con cui il valore d è considerato. Il risultato, compreso in $(0, 1]$, rappresenta una distribuzione di probabilità. Se il valore è alto, quindi il robot è vicino ad un oggetto, la possibilità di urto è elevata.

In seguito alla trasduzione, la natura delle informazioni non è più distinguibile. In questo modo le unità neuronali non devono preoccuparsi di gestire l'eterogeneità sensoriale.

I sensori di WABOT sono di natura passiva, il sistema si deve occupare di interrogarli periodicamente. Seguendo il principio *divide et impera*, non vi è intenzione di trattare le problematiche legate al polling all'interno delle rete neuronale. Un componente *poller* si occupa di raccogliere ciclicamente le informazioni sensoriali e di notificarle alle cellule con periodicità di (circa) 300ms.

Parametri di funzionamento

Prima di passare alla descrizione della struttura neuronale è utile fare il punto sui parametri che influenzano l'apprendimento: η amplifica l'intensità della fase di potenziamento, ϵ riveste il ruolo opposto modulando la depressione sinaptica.

Se il termine di potenziamento è molto maggiore a quello di depressione, a lungo andare si perde la selettività della risposta della cellula. In caso contrario WABOT non è in grado di apprendere. I

due devono essere bilanciati, è necessario eseguire una stima anche se approssimativa.

Le unità del sistema sono binarie, la loro attivazione massima è pari a 1. Considerando che il layer di collisione contenga due nodi, uno per sensore, e assumendo per ipotesi che mediamente la collisione riguardi un solo lato del robot, il livello medio di attivazione di C, indicato con \bar{a} , è pari a 0.5. Utilizzando la seguente regola di apprendimento Hebbiano $\Delta w_{ij} = \eta a_i a_j - \epsilon \bar{a} w_{ij}$, $w^{min} = 0$ e $w^{max} = 1$, sono poste le seguenti considerazioni:

1. Nessuna collisione rilevata. $a_i = \bar{a} = 0$, non vi è apprendimento.
2. Collisione rilevata in un lato del robot. $a_i = 1$ e $\bar{a} = 0.5$, due scenari sono plausibili:
 - (a) $a_j = 1$: ostacolo rilevato dal sensore di prossimità. $\Delta w_{ij} = \eta * 1 * 1 - \epsilon * 0.5 * w_{ij}$, sostituendo $w_{ij} = w^{max}$ si ottiene $\Delta w_{ij} = \eta - \epsilon * 0.5$. Se $\epsilon > 2\eta$ la forza sinaptica diminuisce e WABOT dimentica parte della conoscenza appresa. $\epsilon \ll 2\eta$ porta a over-generalization. Ai fini dell'apprendimento occorre che $\epsilon < 2\eta$ facendo si che tra i due non ci sia uno scarto elevato. w_{ij} è inizialmente inizializzato ad un valore random nell'intervallo $[0, 0.2]$. Considerando $w_{init} = 0.2$ si ottiene $0.2\epsilon < 2\eta$. Il vincolo moltiplicativo, essendo legato al peso corrente del legame, influenza fortemente ϵ . I valori scelti empiricamente sono $\eta = 0.1$ e $\epsilon = 0.3$.
 - (b) $a_j = 0$: ostacolo non rilevato dal sensore di prossimità. $\Delta w_{ij} = \eta * 1 * 0 - \epsilon * 0.5 * w_{ij}$, la forza sinaptica decresce con velocità direttamente proporzionale a ϵ .
3. Se l'apprendimento ha associato gli stimoli percepiti dai sensori di prossimità e collisione, la presenza di un oggetto vicino provoca attività anche nelle cellule neuronali del livello C. $a_j = 1$ implica quindi $a_i = 1$, la forza sinaptica oscilla tra fasi di depressione e potenziamento. Ne consegue che anche a seguito dell'apprendimento c'è possibilità che il robot collida con un oggetto.

La soglia Θ della cellula è convenzionalmente posta a 0.5. Tale valore non influenza i neuroni del layer di contatto né quelli del livello T (in quanto la funzione di normalizzazione segue un andamento lineare). Occorre invece stimare il parametro k della funzione (4.2).

Sapendo che il sensore di prossimità presenta $input_{min} = 0$ e $input_{max} = 255$, l'attivazione del neurone è determinata come:

$$p_j = \begin{cases} 1 : \frac{1}{e^{kd}} > \Theta \\ 0 : \frac{1}{e^{kd}} \leq \Theta \end{cases} \quad (4.3)$$

Considerando il caso in cui il nodo sia eccitato:

$$attivazione = \frac{1}{e^{kd}} > \Theta \quad (4.4)$$

Sviluppando i calcoli si ottiene:

$$k < \frac{1}{d} \ln\left(\frac{1}{\Theta}\right) \quad (4.5)$$

Assumendo che il neurone si attivi quando $input > \frac{input_{max} - input_{min}}{2}$, ponendo il valore normalizzato $d = 0.5$:

$$k < 2 \ln 2 = 1.38 \quad (4.6)$$

Empiricamente si osservano buoni risultati attivando il neurone ad un quinto del range del sensore. Essendo in questo caso $k < 3.47$, arrotondando all'intero inferiore si ottiene $k = 3$.

4.3.3 Architettura della rete neuronale

La struttura della rete neuronale si basa su quella proposta in [46]. Sono previsti quattro livelli, vedi (4.4). Le connessioni C-M e T-M sono predefinite e fissate [48], l'interconnessione è statica e la plasticità sinaptica si mantiene costante al valore unitario. Un nodo appartenente all'insieme C (o T) può attivare una sola azione in M. Ogni unità di tale insieme implementa uno specifico riflesso, il livello contiene in tutto quattro nodi.

L'unità inibitrice realizza l'istinto primordiale di sopravvivenza: evitare gli ostacoli è prioritario rispetto all'avvicinarsi al target. La soglia di inibizione è convenzionalmente fissata a 0.15. Quando l'attività media di C supera tale valore, l'output di T è interdetto.

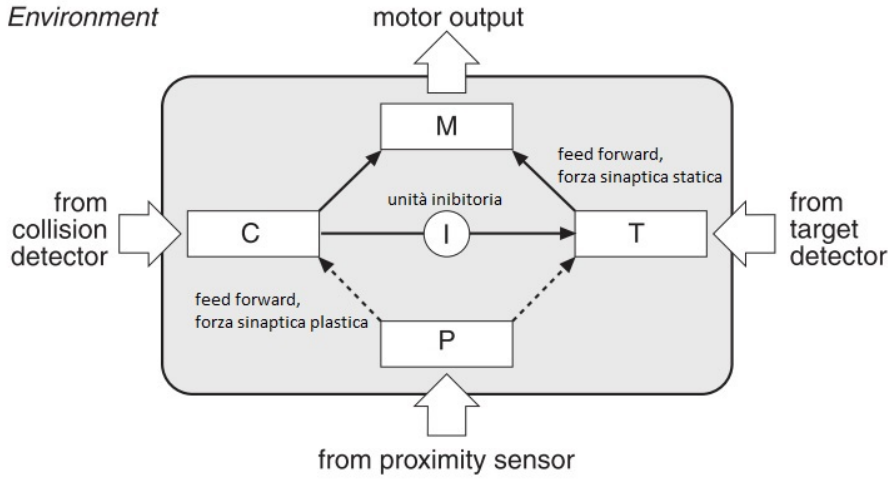


Figura 4.4: Illustrazione presa da [32]. I quattro livelli della rete neurale: prossimità (P), collisione (C), target (T), sistema motore (M). Nello schema è compresa anche l'unità di inibizione (I). Linee tratteggiate indicano collegamenti plastici, linee continue indicano collegamenti statici. In questo capitolo, il layer P è individuato con RF al fine di mantenersi coerenti con l'annotazione utilizzata negli articoli iniziali del paradigma DAC.

L'input del nodo i -esimo del layer C è:

$$h_i = c_i + \sum_{j=1}^N w_{ij} p_j \quad (4.7)$$

- p_j : output del nodo j del layer di prossimità;
- c_i : valore rilevato dal sensore di collisione;
- w_{ij} : forza sinaptica che lega i due nodi;
- h_i : somma pesata gli input del nodo i .

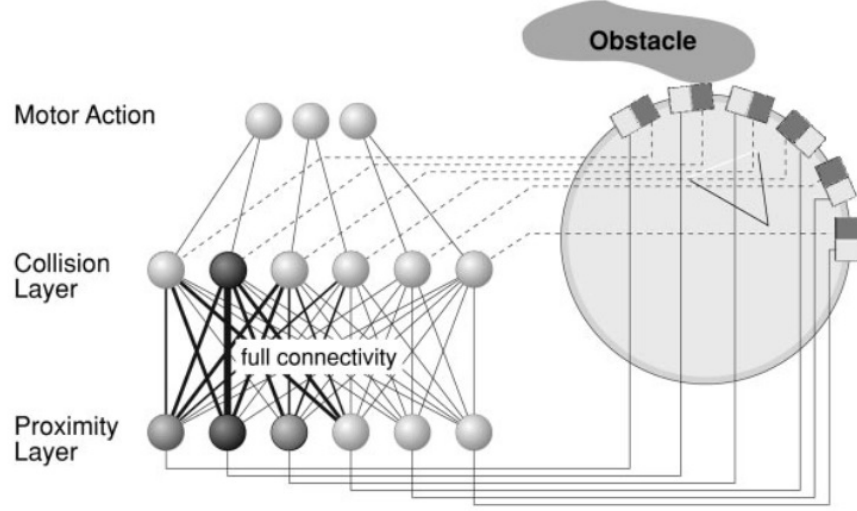


Figura 4.5: Illustrazione presa [32]. Schema dell'interconnessione dei livelli neuronali.

- N : numero dei nodi del layer di prossimità.

Il livello di attivazione del nodo i è definito come:

$$a_i = f(h_i) = \begin{cases} 0 & : h_i \leq \Theta, \\ 1 & : h_i > \Theta \end{cases} \quad (4.8)$$

I layer RF-C e RF-T sono completamente connessi, vedi figura (4.5). L'apprendimento è formulato come segue:

$$\Delta w_{ij} = \frac{1}{N}(\eta a_i p_j - \epsilon \bar{a} w_{ij}) \quad (4.9)$$

N ha il compito di diminuire il modulo della variazione, a_i è l'attivazione post-sinaptica, p_j è l'output del nodo pre-sinaptico, \bar{a} è l'attivazione media del livello di collisione. Assumendo $p_j = a_j$ si ottiene:

$$\Delta w_{ij} = \frac{1}{N}(\eta a_i a_j - \epsilon \bar{a} w_{ij}).$$

L'apprendimento avviene ogni volta che WABOT urta un oggetto o ne rileva la presenza durante l'approccio al target.

Immaginando la rete neuronale come una matrice di pesi, il behaviour emergente è il risultato della trasformazione di tale matrice all'interno di uno spazio multidimensionale.

Il robot non implementa una nozione esplicita di goal. In questo caso studio, all'agente è assegnato un singolo task. Tale condizione è vincolante. Se all'interno dell'ambiente, oltre al cibo, fosse introdotta una fonte d'acqua, il value system dovrebbe modulare gli stimoli percepiti in modo differente nel caso in cui il robot sia affamato o assetato [7].

4.3.4 Debug e problemi relativi all'analisi dei risultati

La complessità dell'analisi delle reti neurali è direttamente proporzionale al numero di cellule. Nonostante questo caso studio comprenda un totale di otto nodi, analizzare il comportamento del robot al variare dei parametri del value system non è semplice per due motivi:

1. Le condizioni ambientali (disposizione e forma degli oggetti, intensità del target, ...) determinano il comportamento osservabile del robot. WABOT è in grado di reagire agli stimoli in alcuni scenari meglio di altri (ad esempio cambiando la posizione di partenza del robot).
2. La tecnologia Lego Mindstorm è limitante in termini di programmazione e debug. Per osservare il log di sistema real time è necessario connettere WABOT ad un PC via USB. A seguito di tale cablaggio la mobilità del robot è limitata. Salvare un log da visionare al termine della sperimentazione è pressoché inutile: per correlare le dinamiche del sistema agli eventi ambientali è necessario un monitoraggio in tempo reale. Inoltre, le eccezioni lanciate in fase di esecuzione sono di difficile interpretazione a causa delle ridotte dimensioni del display LCD.

4.3.5 Stile architetturale

Il modello neuronale proposto in questo caso studio fa riferimento alla figura (4.1). Il diagramma UML (4.6) illustra la rete in termini di

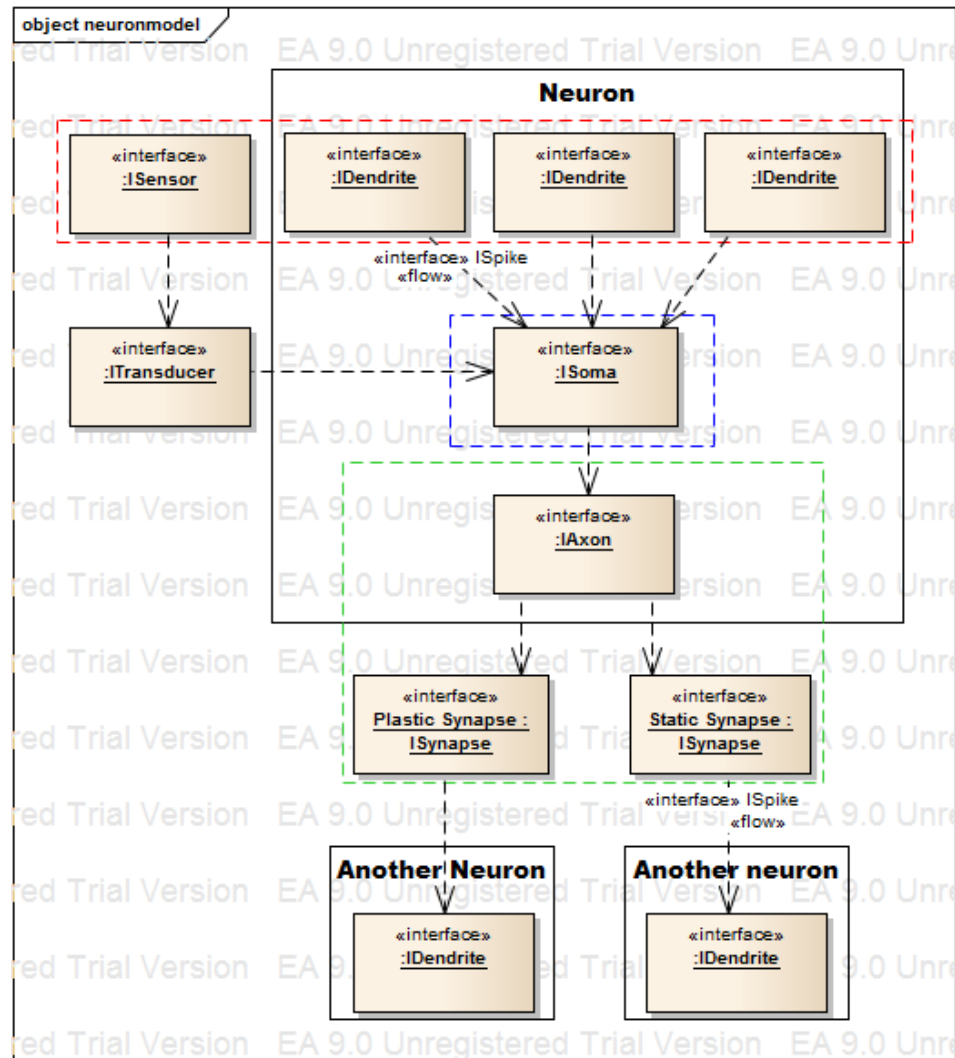


Figura 4.6: Traduzione del modello (4.1) in termini degli elementi descritti in (4.2). Riquadro rosso: input della cellula. Riquadro blu: unità di elaborazione. Riquadro verde: strutture per la trasmissione del segnale di output attraverso legami sinaptici statici e plastici.

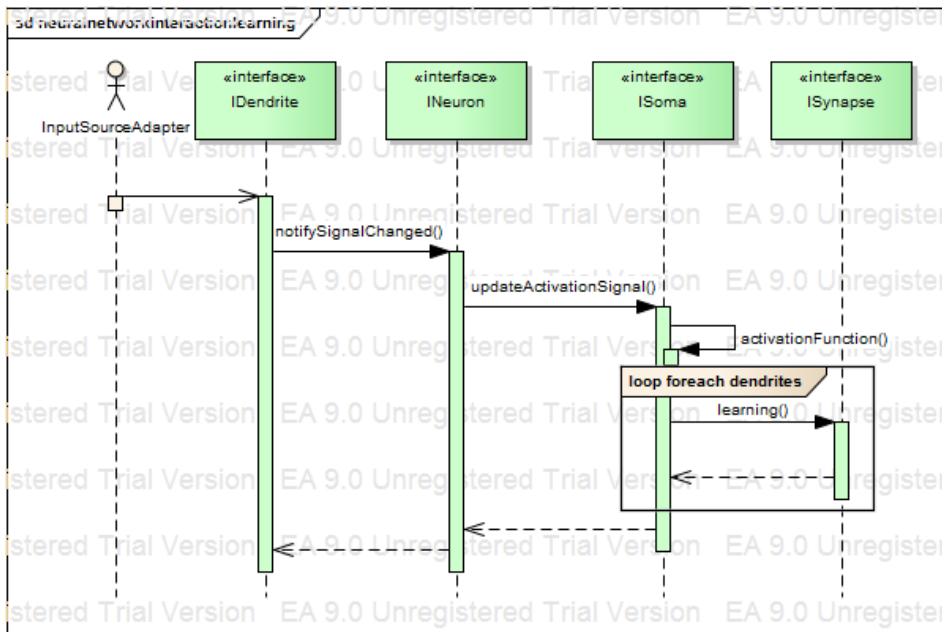


Figura 4.7: Flusso di controllo della fase di apprendimento.

componenti e flusso dati. I dendriti costituiscono i canali di input della cellula neuronale. L'informazione trasmessa è rappresentata dal segnale ISpike. ISoma è il corpo del neurone e adempie alle funzioni di calcolo dell'attivazione, apprendimento e trasmissione del segnale di output lungo l'assone. Tale fibra nervosa è realizzata dall'interfaccia IAxon e trasmette informazione attraverso ISynapse. Eccetto i quattro nodi del layer M, relativo al controllo del sistema motore, ogni altro componente della rete neuronale è collegato ad un sensore. RF e T comprendono entrambi una sola unità, C ne possiede due.

Il diagramma di sequenza, figura (4.7), mostra lo sviluppo della fase di apprendimento etero-sinaptica. A fronte della notifica di variazione di un valore di input, il neurone procede con l'elaborazione dell'informazione e attua il processo di apprendimento sulle sinapsi dei propri dendriti.

4.3.6 Abstraction gap

Il gap che sussiste tra i requisiti e la tecnologia di riferimento è prettamente teorico. Senza troppi sforzi è possibile implementare un modello di rete neurale con una tecnologia OO come Java. Allo stesso modo, Lego Mindstorm fornisce gli strumenti necessari per la creazione di WABOT. Alcune problematiche possono emergere a seguito della ridotta capacità di computazione del *brick* NXT e dalla versione ridotta di JVM prevista da LeJos.

La distanza da colmare è determinata dal background teorico richiesto per trattare reti neurali, processi di apprendimento e DAC. Si prevede che nello sviluppo di questo caso studio un numero importante di ore sia speso nell'approfondimento di tali argomenti.

4.4 Piano di lavoro

Completata la fase di analisi è descritto il processo di sviluppo. Inizialmente il funzionamento della rete neurale è testato in un ambiente simulato. La successione dei dati sensoriali è impostata manualmente al fine di stabilire se il processo di apprendimento stia seguendo il corso desiderato. La rete subisce una fase di training mediante *benchmark* di dati appositamente studiati. Segue un esempio (riportato in linguaggio naturale) utilizzato per testare l'apprendimento tra le unità dei layer RF e C:

```
Sensore di prossimita' rileva ostacolo vicino;  
Collisione sul lato sinistro;  
Fine collisione;  
Sensore di prossimita' non rileva alcun ostacolo.
```

Una volta verificato il funzionamento, la rete neurale è esportata all'interno di WABOT. Il behaviour del sistema è osservato in un log trasmesso ad un PC via USB. Essendo la rete neurale fortemente parallelizzata, questo step intermedio causa problemi di corse critiche per l'accesso al bus seriale. Un meccanismo di coordinazione è implementato per garantire il corretto funzionamento del sistema.

Accertato che il robot operi secondo le dinamiche descritte in fase di analisi, è possibile sconnetterlo dal PC in modo che possa esplorare liberamente l'ambiente. Seguono le osservazioni empiriche.

4.5 Osservazioni empiriche

Questa sezione riporta le osservazioni relative alla fase di sperimentazione, evidenziando i (mancati) riscontri con il caso studio proposto da Pfeifer e Verschure.

1. WABOT è fortemente limitato dal numero di sensori. L'associazione del sensore di prossimità a entrambi in nodi del layer di contatto fa sì che un unico stimolo CS scateni sempre due risposte UR in conflitto tra loro;
2. Il modello differenziale soffre delle problematiche introdotte in 4.3.1. Se direzione logica e fisica coincidono, quando il robot individua il target vi si avvicina con andamento asintotico girando verso destra o sinistra;
3. In seguito ai primi due punti è difficile osservare l'emergere del comportamento che induce il robot a seguire il perimetro della stanza;
4. L'NXT ha capacità computazionali ridotte. Per ottenere performance accettabili sono stati implementati solamente due flussi di controllo: uno per il polling sequenziale dei sensori e uno per l'esecuzione del behaviour. A causa del primo, è possibile predire in modo deterministico il comportamento di WABOT a seguito, ad esempio, di una collisione frontale: il sensore di collisione sinistro è interrogato prima di quello destro, il riflesso attuato pertanto è gira a destra in senso opposto. Utilizzando inoltre un unico sensore di prossimità, la percezione di un ostacolo provoca sempre la stessa risposta.

Il robot non genera diversità.

Gli stimoli condizionati (CS) sono legati all'attivazione dei sensori di prossimità mentre quelli non condizionati (US) ai sensori di

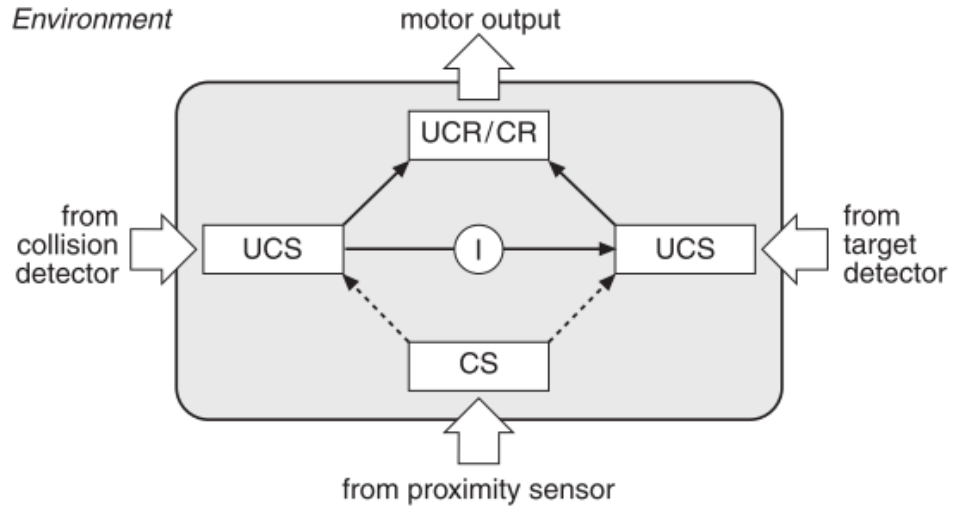


Figura 4.8: Illustrazione presa da [33]. Architettura DAC descritta in termini di condizionamento Pavloviano. I termini UR e US sono qui denotati come UCR e UCS.

target e collisione. Risposte condizionate (CR) e non condizionate (UR) coincidono e rappresentano lo spazio in cui i behaviour sono memorizzati.

Una volta liberato, WABOT esplora l'ambiente. Inizialmente il robot urta con gli oggetti incontrati. La pressione dei sensori di collisione attiva i neuroni all'interno del sistema e il robot evita l'ostacolo. Con il ripetersi degli urti, l'attivazione nello spazio CS diviene sufficientemente forte per attuare autonomamente la risposta non condizionata. Il controllo passa dai riflessi primitivi a comportamenti evoluti correlati agli stimoli condizionati: WABOT ha imparato ad anticipare gli oggetti. Tale fenomeno emerge dall'apprendimento.

A seguito dell'oscillazione tra fasi di depressione e potenziamento, periodicamente il robot dimentica la correlazione CR-UR. L'anticipo dell'ostacolo può essere anche acquisito dal condizionamento RF-T.

In merito a tali osservazioni, i parametri del value system sono settati per garantire una rapida convergenza dell'apprendimento: WABOT anticipa gli oggetti mediamente dopo quattro o cinque collisioni. Tale parametrizzazione porta ad over-generalization, il potenziamento è molto maggiore della depressione sinaptica. Con i mezzi e il tem-

po a disposizione tale meccanismo rappresenta un metodo efficace per verificare l'associazione CS-UR in tempi brevi.

Durante la fase di sperimentazione, è difficile indicare se WABOT stia evitando un ostacolo girando a destra o sinistra a seguito del condizionamento o per raggiungere il target. Per verificare tale comportamento è sufficiente sconnettere il sensore di luminosità: se WABOT mostra ancora tali risposte l'apprendimento RF-T è andato a buon fine. Questo fenomeno è riconducibile agli studi di Hebb sulle lesioni cerebrali: la rimozione di un componente può degradare le performance del sistema senza eliminarne completamente la conoscenza appresa.

Conclusioni

Hebbian learning, value system e DAC sono argomenti strettamente legati tra loro. Osservazioni e analisi delle metodologie di apprendimento sono sviluppate all'interno di ogni capitolo. Le seguenti conclusioni descrivono come il percorso affrontato fornisca spunti futuri per lo studio di fenomeni complessi quali processi mentali e intelligenza.

La teoria Hebbiana, così presentata, è lontana dal definire una soluzione universale al problema dell'apprendimento. Nonostante la semplicità di tale formulazione, possono emergere comportamenti complessi a seguito dell'associazione di cellule neuronali. La modulazione della forza sinaptica, guidata dai principi di località e di cooperatività, permette all'individuo di imparare correlando gli stimoli percepiti dall'ambiente e consente di essere estesa per catturare fenomeni come i disturbi emotivi (ad esempio la paura).

Tale semplicità è allo stesso tempo una limitazione. La teoria Hebbiana risente dell'assenza di un supervisore: anche se estesa dal value system, la mancanza della nozione di errore impedisce di attuare correzioni durante l'apprendimento. Senza possibilità di rettificare i comportamenti acquisiti, al più quest'ultimi possono essere dimenticati.

La robotica *quotidiana* è ancora lontana dall'idea di agenti senzienti. Come riportato da Versace: *robots are everywhere... but where?* [42]. Le motivazioni sono legate, in parte, alle problematiche dell'integrazione con l'essere umano [4], e sono trattabili, ad esempio, con metodi di apprendimento per dimostrazione. Il *Reinforcement learning* è una formulazione altrettanto valida, e l'integrazione di tali processi alla teoria di Hebb può portare a risultati interessanti.

DAC, introdotto nel 1992, è un paradigma adatto per lo sviluppo di agenti autonomi. L'architettura implementata in WABOT è co-

struita ad hoc per la persecuzione di un task specifico e non consente di attuare behaviour complessi (ad esempio legati alla pianificazione). Nella sua prima versione, DAC mette in luce molteplici limitazioni.

Le conclusioni non vanno tratte in ottica di performance: un robot specializzato nell'evitare urti è più efficace di WABOT. Invece, è opportuno porre attenzione al fine che tali studi perseguono, cioè capire fenomeni complessi come intelligenza e apprendimento.

Le versioni successive del paradigma estendono in maniera incrementale le funzionalità dell'agente. Articolando la struttura organizzativa, esse forniscono mezzi per lo sviluppo di un sistema di apprendimento completo. Sfruttando la flessibilità delle reti neurali, l'architettura può essere estesa attraverso modelli di rete in grado di esprimere bisogni e sensazioni dell'agente.

L'evoluzione continua di questi sistemi consente di approfondire gli studi su cervello e corpo umano. In [21], Kurzweil afferma: *reverse engineering the brain (is the way) to understand precisely how it works and using that knowledge to create even more intelligent machines.*

Per questo motivo Hebb, avendo conciliato fisiologia e psicologia, è considerato uno dei padri fondatori dell'apprendimento nelle reti neurali. La sua teoria fonda le basi per studi più evoluti come il riconoscimento di pattern nella neocorteccia e la formulazione di idee e pensieri complessi. Esteso dal value system, l'apprendimento può essere influenzato dalle motivazioni e dalle necessità interne dell'individuo. L'agente sembra così scoprire istanze psichiche simili a quelle Freudiane.

La frontiera della robotica sta oggi riscoprendo e approfondendo nozioni note in psicologia. Nonostante il caso studio qui descritto non si avvicini affatto alla realizzazione di una forma d'intelligenza, esso tenta di delineare una sequenza di piccoli passi per mirare a un obiettivo così lontano.

Bibliografia

- [1] E. M. Azoff. *Neural Network Time Series Forecasting of Financial Markets*. John Wiley & Sons, Inc., July 1994.
- [2] V. Braitenberg. *Vehicles: Experiments in Synthetic Psychology*. MIT Press, 1986.
- [3] C. L. Breazeal. *Designing Sociable Robots*. MIT Press, 2004.
- [4] C. L. Breazeal. Cynthia Breazeal: The rise of personal robots (visionato il 20/09/2014), 2010.
- [5] R. Brooks. A robust layered control system for a mobile robot. *Robotics and Automation, IEEE Journal of*, 2(1):14–23, 1986.
- [6] R. Brooks. Intelligence without reason. *The artificial life route to artificial intelligence: Building embodied, situated agents*, (1293):25–81, 1995.
- [7] F. Cecconi and D. Parisi. Neural networks with motivational units. In *From animals to animats: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pages 346–355. Aug. 1993.
- [8] W. J. Clancey. The frame of reference problem in cognitive modeling, June 1989.
- [9] G. M. Edelman, G. N. Reeke, W. E. Gall, G. Tononi, D. Williams, and O. Sporns. Synthetic neural modeling applied to a real-world artifact. *Proceedings of the National Academy of Sciences of the United States of America*, 89(15):7267–71, Aug. 1992.

-
- [10] J. Fleischer and G. M. Edelman. Brain-based devices. *IEEE Robot. Automat. Mag.*, 16(3):33–41, 2009.
- [11] K. J. Friston, G. Tononi, G. N. Reeke, O. Sporns, and G. M. Edelman. Value-dependent selection in the brain: simulation in a synthetic neural model. *Neuroscience*, 59(2):229–43, Mar. 1994.
- [12] H. Gardner. *Frames of Mind: The Theory of Multiple Intelligences*. Basic books, 1985.
- [13] M. S. Gazzaniga. *Nature’s Mind: The Biological Roots of Thinking, Emotions, Sexuality, Language, and Intelligence*. BasicBooks, 1992.
- [14] W. Gerstner and W. M. Kistler. *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press, 2002.
- [15] E. Gibson. Exploratory Behavior In The Development Of Perceiving, Acting, And The Acquiring Of Knowledge. *Annual Review of Psychology*, 39(1):1–42, 1988.
- [16] D. M. Gordon. *Ant Encounters: Interaction Networks and Colony Behavior*. Princeton University Press, 2010.
- [17] D. O. Hebb. *The organization of behavior: a neuropsychological theory*. Wiley, 1949.
- [18] A. Jain, J. Mao, and K. Mohiuddin. Artificial neural networks: A tutorial. *Computer*, 1996.
- [19] J. Jones and D. Roth. *Robot Programming: A Practical Guide to Behavior-Based Robotics*. McGraw-Hill Education, 2003.
- [20] D. G. Kneller, F. E. Cohen, and R. Langridge. Improvements in protein secondary structure prediction by an enhanced neural network. *Journal of molecular biology*, 214(1):171–82, July 1990.
- [21] R. Kurzweil. *How to Create a Mind: The Secret of Human Thought Revealed*, volume 2012. Penguin, 2012.

BIBLIOGRAFIA

- [22] S. Legg and M. Hutter. A Collection of Definitions of Intelligence. In *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*, volume 157 of *Frontiers in Artificial Intelligence and Applications*, pages 17–24. IOS Press, June 2007.
- [23] R. Llinas. The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science*, 242(4886):1654–1664, Dec. 1988.
- [24] J. L. McClelland. How far can you go with Hebbian learning and when does it lead you astray. In *In Y. Munakata & M.H. Johnson (Eds.), Attention and Performance XXI: Processes of change in brain and cognitive development*, pages 33–39. University Press, 2005.
- [25] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133, Dec. 1943.
- [26] K. D. K. Miller, D. J. C. Mackay, and D. MacKay. The role of constraints in Hebbian learning. *Neural Computation*, 126:100–126, 1994.
- [27] E. Oja. Simplified neuron model as a principal component analyzer. *Journal of mathematical biology*, 15(3):267–273, Nov. 1982.
- [28] R. C. O’Reilly. *The Leabra model of neural interactions and learning in the neocortex*. PhD thesis, Carnegie Mellon University, 1996.
- [29] I. Pavlov. *Conditioned reflex: An investigation of the physiological activity of the cerebral cortex*. Dover Publications, 1960.
- [30] R. Pfeifer and J. Bongard. *How the body shapes the way we think*. MIT press, 2007.
- [31] R. Pfeifer and R. Heifer. Emotions in robot design. In *Proceedings of 1993 2nd IEEE International Workshop on Robot and Human Communication*, pages 408–413. IEEE, 1993.

-
- [32] R. Pfeifer and D. Kumar. Curriculum descant: teaching “New AI”. *intelligence*, 11(2):17–19, June 2000.
- [33] R. Pfeifer and C. Scheier. *Understanding Intelligence*. MIT Press, 2001.
- [34] R. Pfeifer and P. F. Verschure. Distributed adaptive control: a paradigm for designing autonomous agents. In F. J. Varela and P. Bourguine, editors, *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*, pages 21 – 30. MIT Press, 1992.
- [35] R. Pfeifer and P. F. Verschure. Designing efficiently navigating non-goal-directed robots. In *From Animals to Animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pages 31–38. MIT Press, 1993.
- [36] S. Russel and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2009.
- [37] R. Salomon. Improving the DAC architecture by using proprioceptive sensors. In *Proceedings of the fifth international conference on simulation of adaptive behavior on From animals to animats*, pages 331–339. Sept. 1998.
- [38] M. Suwara. Is Cognitive Science a science at all ? In *Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on*, pages 733–738. IEEE, 2012.
- [39] M. Tan, J. Pu, and B. Zheng. Reduction of false-positive recalls using a computerized mammographic image feature analysis scheme. *Physics in medicine and biology*, 59(15):4357–4373, July 2014.
- [40] E. Thelen, L. B. Smith, A. Karmiloff-Smith, and M. H. Johnson. A dynamic systems approach to the development of cognition and action. *Nature*, 372(6501):53, 1994.
- [41] E. L. Thorndike. *Animal Intelligence: Experimental Studies*. Macmillan, 1911.

BIBLIOGRAFIA

- [42] M. Versace. Robots Are Everywhere On Youtube (visionato il 10/01/2013), 2013.
- [43] M. Versace and B. Chandler. The brain of a new machine. *IEEE spectrum*, (December), 2010.
- [44] P. F. Verschure. Distributed adaptive control: explorations in robotics and the biology of learning. *Informatik/Informatique*, 1:25 – 29, 1998.
- [45] P. F. Verschure. Distributed Adaptive Control: A theory of the Mind, Brain, Body Nexus. *Biologically Inspired Cognitive Architectures*, 1:55–72, July 2012.
- [46] P. F. Verschure and A. Coolen. Adaptive fields: distributed representations of classically conditioned associations. *Network: Computation in Neural Systems*, 2(2):189–206, May 1991.
- [47] P. F. Verschure, B. Kröse, and R. Pfeifer. Distributed Adaptive Control: The self-organization of structured behavior. *Robotics and Autonomous Systems*, 9(3):181 – 196, 1992.
- [48] P. F. Verschure and R. Pfeifer. Categorization, Representations, and the Dynamics of System-Environment Interaction: a case study in autonomous systems. pages 210–217, Aug. 1993.
- [49] P. F. Verschure and T. Voegtlin. A bottom up approach towards the acquisition and expression of sequential representations applied to a behaving real-world device: Distributed Adaptive Control III. *Neural Networks*, 11(7):1531–1549, 1998.
- [50] P. F. Verschure, J. Wray, O. Sporns, G. Tononi, and G. M. Edelman. Multilevel analysis of classical conditioning in a behaving real world artifact. *Robotics and Autonomous Systems*, 16(2-4):247–265, Dec. 1995.
- [51] M. Wertheimer. *Gestalt Theory*. Hayes Barton Press, 1938.
- [52] A. Wissner-Gross and C. Freer. Causal Entropic Forces. *Physical Review Letters*, 110(16):168702, Apr. 2013.

- [53] K. Zahedi, N. Ay, and R. Der. Higher coordination with less control—a result of information maximization in the sensorimotor loop. *Adaptive Behavior*, 49(0):1–22, 2010.
- [54] B. Ziebart, A. Maas, J. Bagnell, and A. Dey. Maximum Entropy Inverse Reinforcement Learning. *AAAI*, pages 1433–1438, 2008.
- [55] J. J. Zurada, A. Malinowski, and P. Przestrzelski. Modified Hebbian learning rule for single layer learning. *Circuits and Systems, 1993., ISCAS'93, 1993 IEEE International Symposium on*, pages 2407–2410, 1993.