

ALMA MATER STUDIORUM - UNIVERSITA' DI BOLOGNA

CAMPUS DI CESENA

SCUOLA DI INGEGNERIA E ARCHITETTURA

---

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA BIOMEDICA

MEMORIA SEMANTICA E LESSICALE:

implementazione di un modello della memoria  
semantica e lessicale basato sull'utilizzo di features

Tesi di laurea in  
Sistemi Neurali LM

Relatore  
Prof. Mauro Ursino

Candidata:  
Fatima Eddaghoughi

III Sessione di Laurea  
Anno Accademico 2012/2013

# INDICE

INTRODUZIONE .....	4
--------------------	---

## Capitolo 1

### 1. LA MEMORIA SEMANTICA E LE TEORIE PER UNA RAPPRESENTAZIONE MODELLISTICA

1.1 Memoria semantica .....	8
1.2 Teorie e modelli per simulare l'attività della memoria semantica e lessicale .....	9

## Capitolo 2

### 2. DESCRIZIONE QUALITATIVA E MATEMATICA DEL MODELLO

2.1 Introduzione al modello .....	11
2.2 Rete semantica	
2.2.1 Descrizione qualitativa .....	17
2.2.2 Descrizione matematica .....	21
2.3 Rete lessicale	
2.3.1 Descrizione qualitativa .....	26
2.3.2 Descrizione matematica .....	26
2.4 Addestramento delle sinapsi	
2.4.1 apprendimento rete semantica .....	33
2.4.2 apprendimento rete lessicale .....	38

## Capitolo 3

### 3. RISULTATI DEL MODELLO

3.1 Simulazione delle attività di riconoscimento oggetto .....	42
3.2 Simulazione delle attività di riconoscimento di parola .....	45
3.3 Riconoscimento oggetto con le features alterate .....	46
3.4 Tempi di risposta del modello .....	50

## Capitolo 4

4. DESCRIZIONE DATABASE .....	53
-------------------------------	----

## Capitolo 5

### 5. MODELLO SEMPLIFICATO

5.1 Descrizione del modello .....	55
5.2 Addestramento delle sinapsi .....	58
5.3 Risultati del modello con una tassonomia .....	62
5.4 Discussione e conclusione .....	72

RINGRAZIAMENTI .....	74
----------------------	----

BIBLIOGRAFIA.....	75
-------------------	----

## **INTRODUZIONE**

L'elaborato utilizza un modello neurale, sviluppato nel lavoro di ricerca del prof. Ursino (2010) e successivamente modificato (2011 e 2012), al fine di analizzare alcuni meccanismi alla base della memoria semantica e lessicale. In particolare, la tesi si riferisce alla versione più recente del modello da cui ne deriva uno più semplificato come possibile modalità con cui l'uomo apprende i concetti, li immagazzina in opportune aree cerebrali e collega tali concetti alla parola corrispondente.

Il principio alla base del suddetto modello è il seguente: gli aspetti lessicali e semantici, propri del linguaggio, sono memorizzati in due aree distinte del cervello e connessi tra loro sulla base dell'esperienza passata, mediante un meccanismo di apprendimento fisiologico. Il modello comprende una rete lessicale che rappresenta forme verbali e una rete semantica composta da diverse aree: ogni area è topologicamente organizzata (somiglianza) e codifica per una caratteristica diversa.

Le caratteristiche (o features), appartenenti ad un dato oggetto, possono essere suddivise in: comuni e distintive, salienti e marginali (non salienti). Le prime sono condivise da più oggetti e insieme consentono la formazione delle

categorie; infatti, quando saranno stimulate, non richiameranno il nome dell'oggetto ma della categoria corrispondente. Le proprietà distintive sono uniche per un oggetto specifico, consentendo il suo riconoscimento. La salienza di una feature dipende in modo particolare, ma non solo, dalla frequenza con cui essa è osservata durante l'addestramento dell'oggetto; si dice "saliente", una feature importante, captata molte volte e ricordata immediatamente, quando si pensa ad esso. Viceversa una non saliente o marginale è poco frequente, non è evocata nella memoria quando si pensa all'oggetto, ma facilita la ricostruzione di oggetti quando fornita come input. I risultati possono fornire indicazioni su quali meccanismi neurali possono essere sfruttati per formare robuste categorie tra oggetti e su quali meccanismi potrebbero essere implementati in sistemi connessionisti artificiali per estrarre concetti e categorie da un flusso continuo di oggetti in ingresso (ciascuno rappresentato come un vettore di features). Inoltre, l'organizzazione topologica di features permette il riconoscimento di oggetti con alcune features modificate.

L'ultima versione del modello vuole estendere quella precedente (Ursino, Cuppini, e Magosso, 2011 e Ursino, Cuppini, e Magosso, 2010) per indagare come feature saliente e feature distintiva possono essere naturalmente

codificate utilizzando un addestramento Hebbiano, e di gestire la somiglianza tra oggetti. Sinapsi nel modello vengono create utilizzando le regole di Hebb con valori diversi per le soglie pre-sinaptiche e post-sinaptiche, producendo modelli di sinapsi asimmetriche all'interno della rete semantica (auto-associazione) e tra semantica e la rete lessicale (etero-associazione), in modo da assegnare un ruolo diverso a ciascuna caratteristica, sulla base della loro importanza. Le successive simulazioni, con la rete addestrata, hanno mostrato come, in linea di massima, evocando alcune proprietà, siano richiamate le altre salienti e la parola associata, situata nell'area lessicale. Fornendo in ingresso la parola, si evocano tutte le proprietà salienti dell'oggetto corrispondente, evitando di richiamare le non salienti. Inoltre la rete conduce automaticamente alla formazione di categorie, sfruttando le proprietà comuni a più concetti.

E' stata utilizzata una semplice tassonomia di oggetti schematici, ovvero vettore di features diverse tra cui features condivise e distintive con diversa frequenza di occorrenza.

La rete addestrata è in grado di risolvere compiti di riconoscimento e denominazione di oggetti semplici, mantenendo una distinzione tra categorie e dei suoi membri, e fornendo un ruolo diverso per features salienti e marginali.

Il ruolo della salienza nel riconoscimento degli oggetti non era stato investigato nei lavori precedenti, nei quali tutte le proprietà, usate per simulare l'apprendimento, avevano un uguale grado di salienza.

Scopo del presente lavoro è quello di analizzare il ruolo della maggiore e minore salienza delle features nel riconoscimento degli oggetti, calcolo dei corrispondenti tempi di risposta e gli eventuali cambiamenti nel comportamento della rete, sulla base delle semplificazioni apportate ai parametri ed alle connessioni sinaptiche del modello. Quest'ultimo è stato testato implementando una tassonomie estratta da un database, con alcune features comuni ed altre distintive ma diversamente salienti.

Il primo capitolo del lavoro presenta alcune teorie riguardo la memoria semantica e lessicale. Il secondo ed il terzo illustrano il modello utilizzato, da un punto di vista qualitativo e matematico, ed i suoi risultati con l'analisi dei diversi tempi di riconoscimento degli oggetti. Nel quarto, viene descritto il database utilizzato; mentre nel quinto capitolo è esposta una breve descrizione del modello semplificato con i corrispondenti risultati, ed in fine vi è la discussione.

## Capitolo 1

### 1. LA MEMORIA SEMANTICA E LE TEORIE PER UNA RAPPRESENTAZIONE MODELLISTICA

#### 1.1 Memoria semantica

Il problema di come il cervello rappresenta il significato di oggetti concreti, collega questo significato ad elementi lessicali, e stabilisce categorie di appartenenza è fondamentale per le attività cognitive dell'uomo.

I ricordi sono il centro della nostra individualità. Per ricordare il cervello ha molti sistemi con differenti caratteristiche, mediati da reti neuronali diverse. Si ritiene che il sistema umano concettuale contenga la conoscenza in grado di supportare tutte le attività cognitive, compresa la percezione, la memoria, il linguaggio ed il pensiero [ *Lawrence W. Barsalou, 2003*].

Studi recenti di brain imaging funzionale suggeriscono che i concetti di un oggetto possono essere rappresentati da reti distribuite in diverse regioni corticali. Inoltre, numerose aree della corteccia prefrontale laterale sinistra possono avere ruoli distinti nel recupero, mantenimento e selezione dell'informazione semantica. Il dominio della memoria semantica è costituito da informazioni memorizzate



riguardanti caratteristiche ed attributi che definiscono quei concetti e processi, i quali consentono di recuperare ed agire su tali informazioni [ *Alex Martin and Linda L Chao, 2001*].

la memoria semantica è un tipo di memoria dichiarativa (esplicita) che contiene tutto ciò che si è immagazzinato ed è esprimibile in una qualche forma di linguaggio; è costituita da informazioni basate sulla conoscenza. Essa contiene il repertorio dei concetti ed il vocabolario che li esprime; è descritta come l'archivio di nozioni e concetti comuni e relativamente costanti tra gli individui. Questi concetti sono immagazzinati in diverse modalità senso-motorie ed in più sistemi cognitivi nel cervello. La nostra capacità di impegnarci nelle interazioni con l'ambiente, dipende dalla capacità di comprendere il significato e l'importanza degli oggetti ed azioni intorno a noi, memorizzati nella memoria semantica [*John Hart, 2007*].

## **1.2 Teorie e modelli per simulare l'attività della memoria semantica e lessicale**

Le teorie di base neurale sulla memoria semantica degli oggetti hanno prodotto modelli sofisticati che hanno incorporato in misura differente i risultati riguardo alle indagini cognitive e neurali. Essi sono raggruppati nelle seguenti categorie:

- 1) modelli cognitivi, nei quali i dati neurali sono utilizzati per rivelare dissociazioni nella memoria semantica a seguito di una lesione cerebrale;
- 2) modelli che incorporano sia l'informazione cognitiva sia l'informazione neuroanatomica;
- 3) modelli che utilizzano dati cognitivi, neuroanatomici e neurofisiologici.

Tra gli obiettivi principali di tali modelli: gli oggetti trattabili come stimoli per la comprensione di un concetto integrato, l'incapacità di nominare e ricordare oggetti, danno comune nei pazienti lesionati. L'importanza della memoria semantica è sottolineata dalla profonda disfunzione cui va incontro un individuo in seguito a lesioni. Questo ha motivato numerosi studi per indagare la sua organizzazione neurale e cognitiva nel cervello, così come le relative funzioni ed operazioni. L'idea centrale è che la conoscenza di un dato oggetto sia organizzata in specifiche caratteristiche sensoriali ed in proprietà motorie, associate con l'uso dello stesso.

In questa prospettiva, i disordini inerenti alla conoscenza di categorie specifiche, si sviluppano quando una lesione distrugge l'informazione su una determinata proprietà o un insieme di features critiche per la definizione della categoria di un oggetto e la distinzione dei suoi membri [*Alex Martin, 2007*].

## **Capitolo 2**

### **2. DESCRIZIONE QUALITATIVA E MATEMATICA DEL MODELLO**

#### **2.1 Introduzione al modello**

Questo lavoro di tesi presenta un modello matematico di una rete neurale che simula l'attività della memoria semantica e lessicale, sviluppato nel lavoro di ricerca condotto dal Prof. Ursino et al. nel 2012, evoluzione dei precedenti: 2009, 2010, 2011.

Molti studi comportamentali e di neuroimaging oggi indicano che la conoscenza concettuale è realizzata attraverso una rappresentazione distribuita, che è flessibile e dipendente dall'esperienza. Inoltre, la maggior parte delle teorie presuppongono che un concetto è rappresentato in memoria come un insieme di features (Collins e Loftus, 1975, Cree, McNorgan, e McRae, 2006, McRae, de Sa, e Seidenberg, 1997 Rosch et al., 1976), che possono essere distribuiti su diverse aree corticali, può essere una modalità specifica, e può essere più o meno rilevante a seconda dell'esperienza precedente e/o del contesto in cui il soggetto opera.

All'interno di questo scenario, diversi problemi emergono. Un primo aspetto è che i concetti devono essere appresi e rappresentati nella memoria a diversi livelli, cioè una

memoria semantica dovrebbe essere in grado di trattare con le categorie e i membri di una categoria con una architettura che non è gerarchica (O'Connor, Cree, e McRae, 2009). Piuttosto un comune presupposto quando si utilizza una rappresentazione distribuita di features è che una categoria può emergere tenendo conto delle differenze tra features distintive e comuni. In particolare, la feature è "condivisa" e consente la formazione di categorie se partecipa alla rappresentazione di diversi oggetti; viceversa una feature è "distintiva" quando è tipica di un particolare concetto, e permette alle persone di discriminare tra i diversi concetti (Cree et al., 2006).

Un altro aspetto significativo nella rappresentazione dell'oggetto è la differenza tra le features "salienti" e "marginali". Come sottolineato da Cree, McNorgan, e McRae (2006), una feature saliente è molto evidente, dovrebbe essere frequentemente riportata tra le features dei task di denominazione, e dovrebbe facilmente saltar fuori quando si pensa a un oggetto. È importante sottolineare che la distintività è differente dalla salienza, infatti una feature potrebbe essere distintiva, ma non saliente. Analogamente, una feature può essere saliente ma non distintiva: in tal caso è importante per la rappresentazione di una categoria a un livello gerarchico più alto. In ciò che segue, si considera che

una feature è "saliente" se è spontaneamente evocata quando si pensa a una rappresentazione di oggetti, altrimenti è considerata una feature "marginale". Un esempio aiuterà a capire meglio i concetti appena espressi: l'animale mucca presenta varie peculiarità tra cui il possedere le corna e sette stomaci. Quando si pensa ad essa, la prima proprietà delle due viene subito in mente ed aiuta a ricostruire il concetto stesso di mucca; per cui il possesso delle corna costituisce una proprietà saliente. Viceversa, il pensiero della mucca non evoca la visione dei suoi sette stomaci anche se contribuiscono a svilupparne il suo concetto; quindi questa è una proprietà non saliente ma distintiva.

La versione recente del lavoro del prof. Ursino è centrata proprio sul concetto, appena nominato, di "salienza": le proprietà degli oggetti sono divise in salienti e non salienti. La salienza è simulata assumendo una diversa frequenza per le proprietà durante l'addestramento. Inoltre, diversamente da quanto sviluppato nel modello del 2011, gli oggetti sono somministrati uno alla volta nella rete e non contemporaneamente, potendo così utilizzare un modello statico. In questo modo si sono svolti i calcoli necessari in un tempo minore.

Un altro aspetto importante, sottolineato da alcuni autori recentemente (Taylor et al. , 2012) è la correlazione tra le

features. Due elementi sono correlati se si verificano spesso insieme o sono percepiti insieme. La correlazione tra features può facilitare e velocizzare il riconoscimento di oggetti .

L'analisi precedente, sottolinea l'importanza delle caratteristiche statistiche delle features nella formazione del significato di concetto; anzi, alcuni autori affermano che "l'elaborazione concettuale dipende dagli effetti combinati delle caratteristiche statistiche delle features rilevanti... tale che nessun ulteriore processo di informazione è necessario" (Taylor et al.,2012).

Naturalmente, i concetti sono anche alla base della parola significato, anzi, nell'uomo la rappresentazione concettuale è sistematicamente collegata con il linguaggio, in modo che un rapporto forte si verifica tra le memorie semantiche e lessicali. Tuttavia, vi è ampio accordo oggi che aspetti semantici e lessicali siano memorizzati in regioni distinte del cervello umano (Anderson, 1983 Collins e Loftus, 1975 Seideberf e McClelland, 1989).

Il modo in cui le features vengono combinate insieme per produrre l'organizzazione concettuale complessa della nostra memoria semantica, e di come questa rappresentazione è legata ad aspetti lessicali, è oggetto di un'intensa attività di ricerca, utilizzando sofisticati paradigmi psicologici (Taylor

et al., 2012, Vigliocco et al., 2004), le tecniche di neuroimaging (Martin, 2007), e modelli matematici connessionisti (Cree et al., 2006 Hinton e Shallice, 1991, O'Connor et al., 2009 Rogers et al, 2004, Vigliocco et al. 2004).

Lo studio di come le regolarità statistiche sono sfruttate nella costruzione di concetti, e il ruolo delle tipologie differenti di features nel riconoscimento dell'oggetto può beneficiare dall'uso di una classe di modelli matematici connessionisti, denominata "reti di attrattori". In queste reti, un concetto è rappresentato come un insieme distribuito di nodi attivi che simula gruppi neurali che codificano per le features, mentre le statistiche regolarità delle features di occorrenza sono codificati nei collegamenti tra i nodi (sinapsi), queste ultime vengono assegnate sulla base di alcune regole di addestramento. Per recuperare un concetto, la rete è attratta verso un set stabile di attività, ripristinando carenti informazioni sulla base dell'esperienza precedente. Utilizzando un modello simile, McRae et al. (1997) studiarono il ruolo della funzione di correlazione nell'organizzazione della memoria semantica e spiegarono diversi effetti di priming semantico; Cree et al. (2006) trovarono che features distintive giocano un ruolo maggiore di quelle condivise nella attivazione di un concetto; O'Connor

et al. (2009) dimostrarono che un singolo strato di nodi di features può rappresentare concetti sia subordinati che sovraordinati, senza la necessità di alcuna organizzazione gerarchica a priori.

In questi modelli, tuttavia, l'addestramento di sinapsi è stato realizzato usando un algoritmo supervisionato usando l'algoritmo recurrent-backpropagationthrough-time. Questa non è una regola di apprendimento neurofisiologico. In un ambiente vero, i concetti sono probabilmente appresi attraverso una presentazione semplice dell'oggetto in modo non supervisionato. Quindi, può essere importante per studiare il ruolo delle reti attrattori nella formazione della memoria semantica, all'interno del paradigma dell'apprendimento Hebbiano. Infatti, il quadro Hebbiano sembra un modo naturale per codificare regolarità statistiche e correlazioni tra features in modo semplice e diretto. Varie reti attrattori con l'apprendimento Hebbiano sono state presentate in passato, queste hanno delle somiglianze e differenze con il modello neurale che, dopo questa introduzione, verrà illustrato ed analizzato sia da un punto di vista qualitativo che matematico.



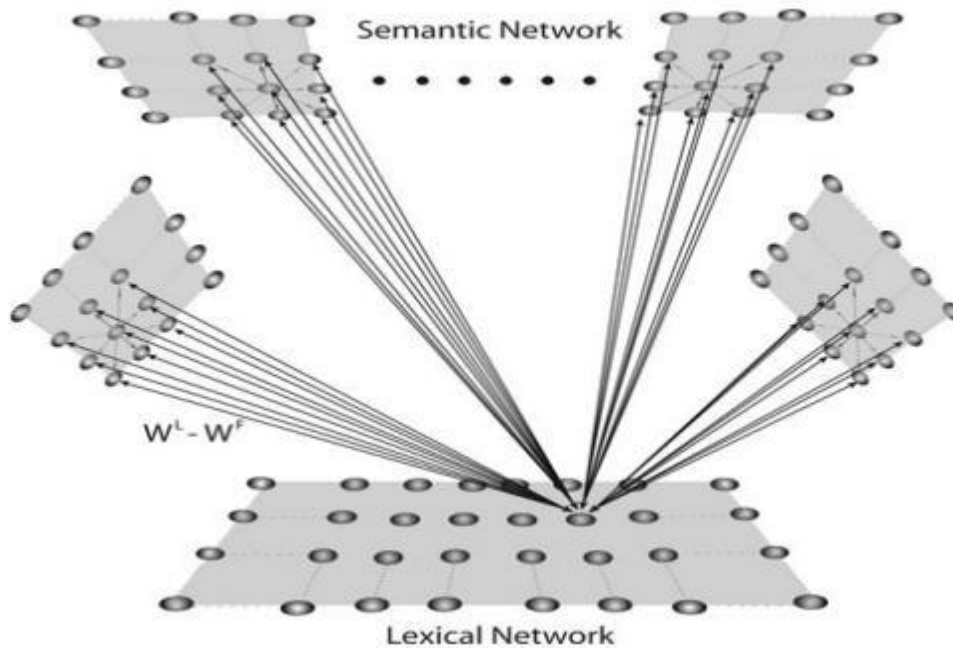
## **2.2 Rete semantica**

### **2.2.1 Descrizione qualitativa**

La struttura di base delle singole reti (dimensioni e connessioni reciproche) è la stessa della versione precedente (Ursino et al., 2010, Ursino et al., 2011). Due principali differenze sono evidenti rispetto alla versione precedente. In primo luogo, per ragioni di semplicità, le parole sono sempre presentate alla rete singolarmente: cioè, non ci si interessa del problema della segmentazione di più oggetti (che richiederebbe sincronizzazione nella banda gamma), e l'obiettivo qui è il problema di realizzare una buona semantica dei singoli oggetti. Per questo motivo, dinamiche neuronali sono descritte con equazioni semplici che non generano oscillazioni banda-gamma. In secondo luogo, abbiamo adottato nuove regole per la formazione delle sinapsi per raggiungere una distinzione tra le features salienti e non-salienti.

Come si evince dalla figure 1, il modello incorpora due reti di neuroni: una rete di features detta "semantica" ed una rete "lessicale". Gli oggetti sono rappresentati come un insieme di features sparse in 9 aree corticali, ognuna costituita da 20x20 neuroni. Quindi è possibile immaginare ogni singola area come una matrice quadrata con 400 neuroni, l'intera rete semantica come una matrice anch'essa quadrata di

dimensioni maggiori, visto che le aree corticali sono disposte a creare una matrice 3x3, cioè con 60x60 neuroni.



**Figura 1:** *Struttura generale del modello che presenta una rete “semantica” ed una “lessicale”. La prima è formata da 9 aree distinte aventi ciascuna 20x20 neuroni. I neuroni di una certa area non possono comunicare con altri della stessa area, ma possono comunicare con neuroni di aree diverse e con quelli della rete lessicale. Quest’ultima consta di uno strato di 20x20 neuroni eccitabili solo da sinapsi della rete semantica.*

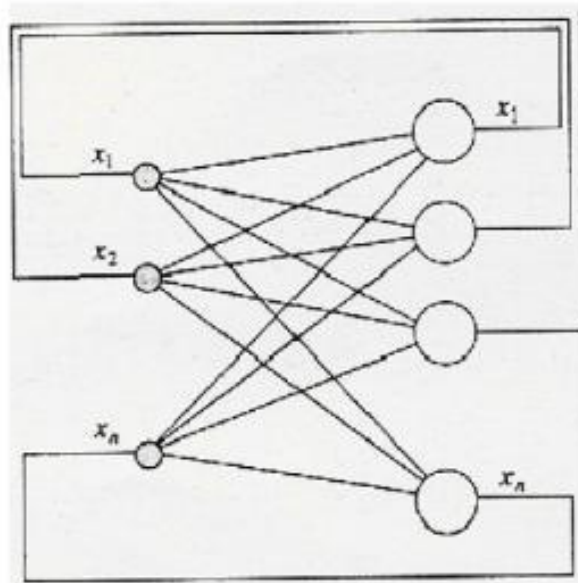
La rete semantica è dedicata alla descrizione di oggetti intesi come un insieme di proprietà senso-motorie. Per semplicità, si utilizzano neuroni non oscillanti e quindi il modello è in grado di gestire un solo oggetto per volta. Ogni feature è rappresentata dall’attività di un singolo neurone, con

coordinate precise all'interno della rete semantica. Si suppone che ciascuna proprietà appartenga ad un'area corticale distinta (sia nella corteccia motoria che nella corteccia sensoriale); all'interno di ogni area, le features sono organizzate secondo un principio di similitudine. Infatti le unità sono collegate le une alle altre, nella stessa area, attraverso sinapsi laterali eccitatorie ed inibitorie, con "disposizione a cappello messicano": cioè i neuroni prossimali si eccitano reciprocamente ed inibiscono quelli più distali. Durante le simulazioni, una proprietà è rappresentata dall'attività di un neurone in una data posizione, attività che può essere evocata da un ingresso esterno localizzato: si assume che esso sia il risultato di una fase di elaborazione a monte che estrae le principali proprietà senso-motorie degli oggetti. Tuttavia, grazie all'organizzazione topologica, l'attività si estende anche alle unità prossimali, sviluppando così "una bolla di eccitazione"; ciò significa che le features simili, situate in posizioni prossimali della rete, diventano moderatamente attive insieme. Una feature può anche essere attivata da un'unità della rete lessicale e può ricevere sinapsi da altre features presenti in diversi settori, realizzando una memoria auto-associativa (figure 2).

Le reti auto-associative, che sfruttano regole Hebbiane, sono costituite da un solo strato di unità; ognuna di esse è connessa

con le altre. Il pattern “Xi” in ingresso, è lo stesso di quello di uscita. La loro funzione principale è quella di apprendere, memorizzare e ricostruire pattern di attivazione. La memorizzazione dei pattern consiste nell’allenare le sinapsi mediante la regola di apprendimento precedentemente nominata: la “regola di Hebb”. Alla fine dell’addestramento, se viene presentata alla rete una versione indebolita di un pattern o una sua parte, la rete è in grado di ricostruire la versione originale del pattern stesso, a seguito di svariati cicli in cui viene propagata l’attivazione [*Dario Floreano, Stefano Nolfi, 2000*]. Inoltre, i pattern vengono memorizzati come punti di equilibrio della rete ed è possibile, come già detto, recuperarli anche se viene dato un ingresso spurio.

Le reti associative (le auto-associative come le etero-associative), diversamente dai computer, non recuperano le informazioni “per indirizzo” ma per somiglianza. Infatti non è necessario precisare un indirizzo di allocazione ma una parte specifica dell’informazione: quest’aspetto le rende molto simili alla memoria degli esseri viventi [*Mauro Ursino, 2011*].



**Figura 2:** rete auto-associativa. Sono presenti due strati di neuroni ma non bisogna lasciarsi ingannare in quanto i pattern in ingresso ed in uscita sono gli stessi; lo strato a monte serve solo per capire che il valore in uscita è retro-azionato su tutta la rete.

### 2.2.2 Descrizione matematica

In questa descrizione gli apici S e L sono usati per indicare una quantità che appartiene alla rete semantica o lessicale, rispettivamente. Una quantità che collega le due reti è rappresentata con due apici, il primo indica la regione destinatario, la seconda regione donatrice. Ogni unità neurale viene indicata con due indici ( $ij$  o  $hk$ ) che rappresentano la sua posizione all'interno della rete. Di conseguenza, una sinapsi tra due unità neurali ha quattro pedici ( $ij, hk$ ) i primi due rappresentano la posizione del neurone postsinaptico, i secondi la posizione del neurone presinaptico.

A livello matematico l'attività del neurone viene descritta da un'equazione differenziale di primo grado nel tempo (detta  $X_{i,j}$ ):

$$\tau^A \frac{d}{dt} x_{ij}^A(t) = -x_{ij}^A(t) + H^A(u_{ij}^A(t)) \quad A = S, L \quad (1)$$

dove l' apice A indica la rete (semantica o lessicale),  $\tau^A$  e la costante di tempo, che determina la velocità di risposta allo stimolo, e  $H^A(u^A(t))$  è una funzione di attivazione sigmoideale che modula gli stimoli in ingresso al neurone ed è descritta dalla seguente equazione:

$$H^A(u^A(t)) = \frac{1}{1 + e^{-(u^A(t) - \varphi^A) p^A}}; \quad (2)$$

dove p è un parametro che imposta la pendenza centrale della sigmoide, e  $\varphi$  imposta la sua posizione centrale. Eq. 2 assume convenzionalmente che l'attività massima è 1 (cioè, tutte le attività neurali sono normalizzate al massimo).

$u_{ij}^A(t)$  nell' eq. 1 è l'ingresso totale raggiunto dal neurone, e ha diverse espressioni nelle reti semantiche e lessicali.

L'ingresso ai neuroni nella rete semantica (apice A = S) viene calcolato come la somma di quattro contributi

$$u_{ij}^S(t) = I_{ij}^S(t) + E_{ij}^{SS}(t) + L_{ij}^{SS}(t) + C_{ij}^{SL}(t) \quad (3)$$

$I_{ij}^S$  rappresenta l'ingresso esterno per l'unità neurale in posizione  $ij$ , proveniente dal motore sensoriale-emotivo catena di lavorazione che estrae le features. può assumere il valore 0 (assenza di feature) o 1 (presenza di feature).  $E_{ij}^{SS}$  rappresenta un termine di accoppiamento eccitatorio proveniente da unità in altre aree della rete semantica (cioè dai neuroni che codificano per una funzione diversa (Eq4),  $L_{ij}^{SS}$  rappresenta l' input da connessioni laterali provenienti da unità nella stessa zona (Eqs 5-9) mentre  $C_{ij}^{SL}$  è un termine di cross-rete, che descrive le connessioni provenienti dalla zona lessicale (Eq. 10) .

Le condizioni di accoppiamento all'interno della rete semantica hanno le seguenti espressioni

$$E_{ij}^{SS} = \sum_h \sum_k W_{ij,hk}^{SS} \cdot x_{hk}^S \quad (4)$$

$$L_{ij}^{SS} = \sum_h \sum_k L_{ij,hk}^{SS,EX} \cdot x_{hk}^S - \sum_h \sum_k L_{ij,hk}^{SS,IN} \cdot x_{hk}^S \quad (5)$$

dove  $ij$  indica la posizione del neurone postsinaptico (target), e  $hk$  la posizione del neurone presinaptico, e le somme si estendono a tutti i neuroni presinaptici nella rete semantica.

Il simbolo  $W_{ij,hk}^{SS}$  rappresenta sinapsi eccitatorie inter-area che realizzano una rete auto-associativa nella semantica. Queste sinapsi sono oggetto di apprendimento, tuttavia, si suppone che siano sempre uguali a zero all'interno della stessa area (ad esempio, solo le features differenti, codificati in aree diverse, possono essere collegate insieme con l'esperienza). I simboli  $L_{ij,hk}^{SS,EX}$  e  $L_{ij,hk}^{SS,IN}$  rappresentano sinapsi laterali eccitatorie ed inibitorie tra neuroni nella stessa area, e sono utilizzati per implementare un principio di similitudine. Di conseguenza, tutti i termini  $L_{ij,hk}^{SS,EX}$  e  $L_{ij,hk}^{SS,IN}$  con i neuroni  $ij$  e  $hk$  appartenenti a diverse aree sono impostati a zero. Le sinapsi laterali realizzano una disposizione di cappello messicano, cioè, una zona eccitatoria tra unità prossimali circondata da un anello di inibizione tra le unità più distali. Quindi:

$$L_{ij,hk}^{SS,EX} = \begin{cases} L_0^{EX} e^{-(d_{ij,hk})^2 / (2\sigma_{ex}^2)} & \text{if } ij \text{ and } hk \text{ are in the same area} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$$L_{ij,hk}^{SS,IN} = \begin{cases} L_0^{IN} e^{-(d_{ij,hk})^2 / (2\sigma_{in}^2)} & \text{if } ij \text{ and } hk \text{ are in the same area} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

dove  $L_0^{EX}$  e  $L_0^{IN}$  sono parametri costanti, che stabiliscono la forza delle sinapsi laterali (eccitatorie ed inibitorie),  $\sigma_{ex}$  e  $\sigma_{in}$  determinano l'estensione di queste sinapsi, e  $d_{ij,hk}$



rappresenta la distanza tra le unità in posizione  $ij$  e  $hk$ . Al fine di evitare effetti di bordo (cioè, che i neuroni vicini al confine non ricevino abbastanza sinapsi dagli altri) le distanze sono state calcolate simulando una struttura sferica per ogni area. Questo è ottenuto utilizzando la seguente espressione per la distanza :

$$d_{ij,hk} = \sqrt{\delta_{i,h}^2 + \delta_{j,k}^2} \quad (8)$$

Dove

$$\delta_{l,m} = \begin{cases} |l-m| & \text{if } |l-m| < N/2 \\ N-|l-m| & \text{otherwise} \end{cases} \quad (9)$$

Infine, il termine cross-rete  $C_{ij}^{SL}$  proveniente dalla zona lessicale viene calcolato come segue

$$C_{ij}^{SL} = \sum_h \sum_k W_{ij,hk}^{SL} \cdot x_{hk}^L \quad (10)$$

Dove  $x_{hk}^L$  rappresenta l'attività del neurone  $hk$  nell'area lessicale ed i simboli  $W_{ij,hk}^{SL}$  sono sinapsi dal lessicale alla rete semantica.

## **2.3 Rete lessicale**

### **2.3.1 Descrizione qualitativa**

Allo scopo di rappresentare gli aspetti lessicali, il modello include un ulteriore strato di neuroni, denominato “rete lessicale”. Ciascuna unità di calcolo codifica una parola ed è associata alla rappresentazione di un singolo oggetto. Anche in questo caso, l’ingresso deve essere considerato come il risultato di una fase di elaborazione a monte, che riconosce le singole word-forms, a partire da fonemi o dall’analisi ortografica. All’interno della rete lessicale, non viene inclusa alcuna sinapsi laterale. Tuttavia, le unità possono essere stimulate dalle sinapsi a lungo raggio provenienti dalla rete semantica. Quest’ultime sono sottoposte ad apprendimento durante una fase di addestramento in cui le parole ed i concetti vengono somministrati insieme.

### **2.3.2 Descrizione matematica**

Nel secondo strato di neuroni (20x20) presenti nel modello detto “rete lessicale”, ogni unità costituisce una parola associata ad un singolo oggetto nella rete semantica.

Poiché gli oggetti semantici costituiscono degli insiemi di features, ciascuna parola è collegata, attraverso le sinapsi, a tutte le features dell’oggetto o della categoria che essa rappresenta.

l'attività del neurone lessicale è la stessa di quella espressa per il neurone semantico.

L'ingresso all'unità neurale lessicale nella posizione  $ij$  ( $u_{ij}^L(t)$ ) (Eq.1) comprende solo due termini in quanto, nel presente modello, non consideriamo i collegamenti laterali tra unità lessicali. Quindi:

$$u_{ij}^L(t) = I_{ij}^L(t) + C_{ij}^{LS}(t) \quad (11)$$

$I_{ij}^L(t)$  è l'ingresso prodotto da una stimolazione linguistica esterna, proveniente sia dall' ascolto di un discorso o dalla lettura di fonemi.  $I_{ij}^L(t)$  assume valore 1 quando la word-form è data alla rete, e zero altrimenti.  $C_{ij}^{LS}(t)$  rappresenta l'intensità dell'ingresso causato da connessioni sinaptiche dalla rete semantica.

Sinapsi dalla semantica alla rete lessicale includono sia un termine eccitatorio che un termine inibitorio ( $W_{ij,hk}^{LS}$  e  $V_{ij,hk}^{LS}$  rispettivamente), che sono addestrati in modi diversi. Ciò comporta una strategia inibitoria-eccitatoria più complessa. In realtà, abbiamo bisogno che una word-form nella zona lessicale sia entusiasta quando sono presenti nello scenario semantico tutte le features salienti, ma viene inibita quando la rete semantica include una feature che non appartiene

all'oggetto (non si accetta un "gatto che abbaia" o una "mucca volante ! "). In altre parole, le features non coinvolte nella rappresentazione dell'oggetto inibiscono la corrispondente unità lessicale. Quindi, si può scrivere

$$C_{ij}^{LS} = \sum_h \sum_k W_{ij,hk}^{LS} \cdot x_{hk}^S - \sum_h \sum_k V_{ij,hk}^{LS} \cdot x_{hk}^S \quad (12)$$

Dove  $x_{hk}^S$  rappresenta l'attività del neurone  $hk$  nella rete semantica (Eq.1),  $W_{ij,hk}^{LS}$  è la forza delle sinapsi eccitatorie e  $V_{ij,hk}^{LS}$  la forza delle sinapsi inibitorie.

## 2.4 Addestramento delle sinapsi

All'inizio dell' addestramento di tutte le sinapsi inter-area eccitatorie nella rete semantica ( $W_{ij,hk}^{SS}$ ) e tutte le sinapsi eccitatorie ed inibitorie tra la semantica e le reti lessicali ( $W_{ij,hk}^{SL}$ ,  $W_{ij,hk}^{LS}$  e  $V_{ij,hk}^{LS}$ ) sono fissate a zero. L'addestramento è quindi diviso in due fasi distinte:

- 1) Nella prima fase di addestramento, i singoli oggetti (descritti dalle loro features in base ad una determinata statistica) sono presentati alla rete uno per uno, e le sinapsi inter-area che collegano le diverse features (es. sinapsi  $W_{ij,kh}^{SS}$  Eq.4) sono apprese. Le features hanno una

diversa frequenza di occorrenza, che stabilisce la loro salienza.

- 2) Nella seconda fase, gli oggetti (descritti dalle features con le stesse statistiche della fase 1) sono presentati alla rete insieme con la corrispondente word-form della rete lessicale, e le sinapsi che collegano le reti semantiche e lessicali ( $W_{ij,kh}^{SL}$ ,  $W_{ij,kh}^{LS}$  e  $V_{ij,kh}^{LS}$ ) sono apprese.

Tutte le sinapsi eccitatorie ( $W_{ij,kh}^{SS}$ ,  $W_{ij,kh}^{SL}$ , e  $W_{ij,kh}^{LS}$ ) sono addestrate con una regola Hebbiana, che modifica il peso sulla base della correlazione tra l'attività presinaptica e postsinaptica. Per tener in considerazione non solo il potenziamento a lungo termine, ma anche la depressione a lungo termine, queste attività sono confrontate con una soglia. In questo modo, un basso livello di attività in un neurone provoca una depressione della forza sinaptica se accompagnata da un alto livello di attività del neurone. Utilizzando il significato dei simboli spiegato sopra, si può scrivere:

$$\Delta W_{ij,kh}^{AB} = \gamma_{ij,hk}^{AB} (x_{ij}^A - \mathcal{G}_{post}^{AB})(x_{hk}^B - \mathcal{G}_{pre}^{AB}) \quad (13)$$

dove gli apici AB possono assumere il significato SS, SL o LS a seconda della particolare sinapsi,  $\Delta W_{ij,kh}^{AB}$  rappresenta la variazione della forza sinaptica, a causa delle attività pre e

post-sinaptiche,  $\mathcal{G}_{post}^{AB}$  e  $\mathcal{G}_{pre}^{AB}$  sono soglie per le attività post e presinaptiche,  $\gamma^{AB}$  denota un fattore di apprendimento,  $x_{ij}^A$  è l'attività dell'unità neurale in posizione  $ij$  dell'area post-sinaptica, e  $x_{hk}^B$  è l'attività dell'unità neurale in posizione  $hk$  nell'area pre-sinaptica.

La regola precedente, tuttavia, ha bisogno di alcuni aggiustamenti per essere davvero fisiologica. In primo luogo, quando entrambe le attività presinaptica e postsinaptica sono basse, non dovrebbe verificarsi nessun cambiamento di peso. Quindi:

$$\text{se } x_{ij}^A < \mathcal{G}_{post}^{AB} \text{ e } x_{hk}^B < \mathcal{G}_{pre}^{AB} \text{ allora } \Delta W_{ij,hk}^{AB} = 0 \quad (14)$$

In secondo luogo, le sinapsi non possono aumentare all'infinito, ma devono raggiungere un livello massimo di saturazione. Questo risultato è ottenuto, nel nostro modello, riducendo progressivamente il tasso di apprendimento quando le sinapsi si avvicinano al loro massimo ( $W_{max}^{AB}$ ). Si ha:

$$\gamma_{ij,hk}^{AB} = \frac{\gamma_0^{AB}}{W_{max}^{AB}} (W_{max}^{AB} - W_{ij,hk}^{AB}) \quad (15)$$

dove  $\gamma_0$  è il tasso di apprendimento quando la sinapsi è zero.

Infine, una sinapsi non può diventare negativa (altrimenti eccitazione verrebbe convertita in inibizione, che non è fisiologicamente accettabile). Quindi, quando calcola il nuovo valore della sinapsi si ha:

$$W_{ij,hk}^{AB} \leftarrow (W_{ij,hk}^{AB} + \Delta W_{ij,hk}^{AB}) U(W_{ij,hk}^{AB} + \Delta W_{ij,hk}^{AB}) \quad (16)$$

dove il simbolo  $\leftarrow$  significa che il valore calcolato a destra è assegnato ai membri di sinistra e  $U(y)$  rappresenta la funzione a gradino (cioè,  $U(y) = 1$  se  $S > 0$ , altrimenti  $U(y) = 0$ ). Le equazioni 13 e 16 sono state applicate in condizioni stazionarie finali (cioè, dopo 200 ms dalla presentazione di ingresso) per evitare eventuali effetti dei cambiamenti transitori di attività.

Infine, un ulteriore aspetto è stato incluso per quanto riguarda le sinapsi eccitatorie dalla semantica alla rete lessicale (cioè,  $W_{ij,kh}^{LS}$  nella (Eq.13). Infatti, si vuole che una word-form nella rete lessicale sia eccitata, se e solo se tutte le features salienti sono attive nella rappresentazione semantica, indipendentemente dal loro numero (cioè, alcune parole possono avere tre features salienti, altre possono avere cinque features salienti e così via. Nel primo caso tre bolle di attivazione nella rete semantica sono sufficienti per eccitare la word-form, nel secondo caso, sono necessari cinque bolle). Per giustificare questa proprietà essenziale, sono state

normalizzate le sinapsi che entrano in una word-form assumendo che la loro somma non può superare un livello massimo. Quindi, la seguente regola viene utilizzata per aggiornare le sinapsi  $W_{ij,kh}^{LS}$

$$\Gamma_{ij}^{LS} = \sum_h \sum_k (W_{ij,hk}^{LS} + \Delta W_{ij,hk}^{LS}) U(W_{ij,hk}^{LS} + \Delta W_{ij,hk}^{LS})$$

$$W_{ij,hk}^{LS} \leftarrow \begin{cases} (W_{ij,hk}^{LS} + \Delta W_{ij,hk}^{LS}) U(W_{ij,hk}^{LS} + \Delta W_{ij,hk}^{LS}) & \text{if } \Gamma_{ij}^{LS} \leq \Gamma_{\max} \\ (W_{ij,hk}^{LS} + \Delta W_{ij,hk}^{LS}) U(W_{ij,hk}^{LS} + \Delta W_{ij,hk}^{LS}) \cdot \Gamma_{\max} / \Gamma_{ij}^{LS} & \text{if } \Gamma_{ij}^{LS} > \Gamma_{\max} \end{cases} \quad (17)$$

dove il simbolo  $\Gamma_{ij}^{LS}$  indica la somma di tutte le sinapsi eccitatorie introdotte nella word-form in posizione ij (quindi, la somma è estesa a tutte le unità neurali hk nella rete semantica). L'eq.17 sostituisce la 15 e la 16.

Le sinapsi inibitorie  $V_{ij,kh}^{LS}$  nell'eq.12 sono addestrate con una regola anti- Hebbiana, cioè, sono indebolite quando entrambe le attività pre e post-sinaptica sono al di sopra della soglia, e sono rafforzate quando le attività sono negativamente correlate. Quindi

$$\Delta V_{ij,hk}^{LS} = -\beta_{ij,hk}^{LS} (x_{ij}^L - \rho_{post}^{LS})(x_{hk} - \rho_{pre}^{LS}) \quad (18)$$

dove i simboli  $\rho$  e  $\beta$  sono usati per la soglia e il tasso di apprendimento, per evitare confusione con quelli utilizzati per le sinapsi eccitatorie.



Naturalmente, queste sinapsi non possono cambiare il loro segno (cioè, una sinapsi non può shiftare da inibitoria ad eccitatoria), e hanno un livello massimo di saturazione. Queste ultime equazioni sono omesse per brevità, essendo analoghe alle eq.15 e 16.

### **2.4.1 Apprendimento rete semantica**

Un punto fondamentale, per realizzare un processo di addestramento corretto, riguarda la scelta di appropriati valori per le soglie pre e post-sinaptiche nella regola di Hebb. Cominciando dalla rete semantica, bisogna dire che questa, affinché sia valida, deve soddisfare i seguenti requisiti:

- 1) Salienza vs non-salienza:
  - a) Le features salienti devono essere evocate da tutte le altre features dell'oggetto (sia salienti che non salienti), di conseguenza, esse dovrebbero ricevere forti sinapsi di input. Al contrario, esse devono inviare forti sinapsi di uscita solo alle altre features salienti.
  - b) le features non salienti non dovrebbero essere evocate dalle altre features dell'oggetto (cioè, che ricevono sinapsi di input deboli), ma favoriscono la ricostruzione dell'oggetto (quindi,

mandano forti sinapsi di uscita verso le features salienti).

Condivise vs distintive:

- a) Le features condivise da diversi oggetti in una categoria dovrebbero ricordare tutte le altre features salienti condivise nella categoria, ma non dovrebbero ricordare le caratteristiche distintive dei singoli membri (ad esempio, la feature "ha una coda" dovrebbe richiamare le features comuni della categoria "animale", come "ha gli occhi", ma non dovrebbe ricordare le features dei singoli membri "gatto", "cane", ecc ..., come "abbaia", "miagola" ecc...).
- b) Una feature distintiva deve ricordare non solo le altre features distintive salienti dello stesso oggetto, ma anche le features comuni salienti (la feature "abbaia" dovrebbe richiamare tutte le proprietà salienti di un cane, tra cui "ha una coda", "ha due occhi" condivise con gli altri animali).

Questo particolare comportamento delle sinapsi semantiche può essere raggiunto assumendo che la soglia per l'attività post-sinaptica sia piuttosto elevata (si suppone,  $g_{post}^{SS} = 0.5$ , cioè, a metà tra la massima eccitazione e la massima

inibizione) mentre la soglia per l'attività pre-sinaptica sia bassa (assunta  $\mathcal{G}_{pre}^{SS} = 0.05$  vicina all' inibizione. Un valore di poco superiore allo zero è stato scelto qui per evitare che una attività neuronale residua provochi un rinforzo di sinapsi indesiderato). Questa scelta comporta le seguenti conseguenze principali :

- i. Se entrambi i neuroni presinaptici e postsinaptici sono attivi (attività vicina a 1), la sinapsi rafforza [si ha  $\Delta W_{ij,hk}^{SS} = \mathcal{V}_{ij,hk}^{SS} \cdot 0.5 \cdot 0.95 = \cdot 0.475$ ].
- ii. Se il neurone postsinaptico viene inibito (attività vicina a 0), mentre il neurone presinaptico è elevato (attività vicino a 1), la sinapsi indebolisce [si ha  $\Delta W_{ij,hk}^{SS} = -\mathcal{V}_{ij,hk}^{SS} \cdot 0.475$ ].

Questa situazione si verifica per la sinapsi che lascia una feature condivisa verso una feature distintiva, quando la feature comune appare in un concetto non contenente quella particolare feature distintiva (ad esempio la feature "ha una coda" verso la feature "miagola", quando uno sta guardando un cane). La stessa situazione si verifica anche per le sinapsi che lasciano una feature frequente verso una feature non-frequente, quando si percepisce il primo e il secondo non viene percepito. Quindi, dopo un addestramento sufficiente, a causa delle statistiche del verificarsi delle features, le features condivise invieranno

sinapsi povere verso le features distintive, e le features salienti invieranno sinapsi povere verso quelle non-salienti.

- iii. Se il neurone post-sinaptico viene eccitato (attività vicino a 1) e il neurone presinaptico viene inibito (attività vicina a zero), le sinapsi mostrano solo un indebolimento moderato [si ha  $\Delta W_{ij,hk}^{SS} = - \gamma_{ij,hk}^{SS} \cdot 0.5 \cdot 0.05 = - \gamma_{ij,hk}^{SS} \cdot 0.025$ ]. Questa è la situazione che si verifica nelle sinapsi che lasciano una feature non-saliente verso una feature saliente, quando la feature non-saliente non è percepita (l'attività pre-sinaptica è spesso vicina allo zero, poiché le features non-salienti sono spesso assenti). Di conseguenza, una feature non-saliente distintiva continua a inviare sinapsi forti verso tutte le features salienti, con scarso indebolimento. La stessa condizione si verifica anche se si considera una sinapsi da qualche feature distintiva verso una saliente condivisa.

Le considerazioni precedenti richiedono due osservazioni sul concetto di salienza utilizzata in questo lavoro. In primo luogo, qui si considera la frequenza di occorrenza come unico aspetto che caratterizza la salienza. Gli autori di questo lavoro sono consapevoli che questa è una limitazione, ma la hanno adottata per semplificare la loro analisi.

In secondo luogo, è opportuno notare che il livello di salienza (cioè, se una feature saliente è o meno basata sulla sua frequenza) è strettamente correlato con il valore usato per la soglia post-sinaptica. Maggiore è tale soglia, maggiore sarà il livello di frequenza richiesto per avere la salienza.

Il valore di saturazione per le sinapsi semantiche ( $W_{\max}^{SS}$ ) è stato assegnato lo stesso come nei precedenti lavori (Ursino et al., 2010, Ursino et al., 2011). Questo valore è stato scelto in modo che, quando le sinapsi sono vicine al massimo, l'attivazione di un' unica "bolla" in un' area semantica può eccitare un neurone bersaglio, in un'altra area semantica, sopra la parte centrale della relazione sigmoideale (cioè al di sopra dell' ascissa centrale  $\vartheta^{SS}$ ). Il tasso di apprendimento,  $\gamma_0^{SS}$ , è stato diminuito rispetto al precedente lavoro: con il nuovo valore, addestrare la semantica di un oggetto avviene più gradualmente, richiedendo circa cento presentazioni di uno stesso oggetto per ottenere sinapsi efficaci. Questo permette una conservazione più accurata delle statistiche delle features di ingresso.

## 2.4.2 Apprendimento rete lessicale

Nell'addestramento della rete lessicale, come nell'addestramento della semantica sopra descritto, sono usati valori dei tempi di apprendimento sufficientemente piccoli per assicurare una convergenza graduale (circa 100 presentazioni di ogni word-form con l'oggetto associato).

Le sinapsi eccitatorie dal lessicale alle unità semantiche ( $W_{ij,hk}^{SL}$ ) sono state addestrate con una bassa soglia pre-sinaptica  $\mathcal{G}_{pre}^{SL} = 0$  e una elevata soglia post-sinaptica  $\mathcal{G}_{post}^{SL} = 0.5$ . La soglia per le sinapsi eccitatorie dalle features semantiche ad una word-form,  $W_{ij,hk}^{LS}$ , è stata data a seguito di un ragionamento simile, ma con un ruolo opposto per le soglie pre e post-sinaptiche ( $\mathcal{G}_{post}^{LS} = 0$ ;  $\mathcal{G}_{pre}^{LS} = 0.5$ ). In altre parole, la soglia per la word-form è sempre a zero, e la soglia per l'unità semantica è sempre 0,5, indipendentemente dal fatto che il neurone sia pre o post-sinaptico. Ciò significa che una word-form deve essere attiva per garantire l'apprendimento. Poi la sinapsi si rafforza quando la feature corrispondente è presente nella rete semantica, e si indebolisce quando la feature è assente. Di conseguenza, solo quelle features che partecipano abbastanza frequentemente alla rappresentazione dell'oggetto (features salienti, in particolare quelle che si verificano più del 50 % delle volte) sono spontaneamente collegate con la

word-form.

L'addestramento delle sinapsi inibitorie,  $V_{ij,hk}^{SL}$ , per le word-forms (apprendimento anti-Hebbiano) richiede una strategia diversa. C'è bisogno che una feature che non partecipa mai (o raramente) alla semantica di tale oggetto (chiamato oggetto 1) ma partecipa spesso alla semantica di altri oggetti (ad esempio oggetto2, oggetto3, ecc ...) inibisce la word-form rispetto all'oggetto1 (per esempio, la feature "fa freddo" dovrebbe inibire la word-form "sole", la feature "abbaia" dovrebbe inibire la word-form "gatto"). Per raggiungere questo obiettivo, è stato deciso di addestrare le sinapsi inibitorie ogni volta che la feature è attiva nella rete semantica (questa è l'unità presinaptica, quindi si assume  $\rho_{pre}^{LS} = 0$  in Eq. 19). Inoltre, la soglia per l'attività post-sinaptica (cioè la word-form) ha ricevuto un valore basso ( $\rho_{post}^{LS} = 0.05$ ). In questo modo, se la feature e la corrispondente word-form sono attive contemporaneamente, la sinapsi inibitoria è drasticamente ridotta ( $\Delta V_{ij,hk}^{LS} = -\beta_{ij,hk}^{SS} \cdot 0.95$  Eq. 18). Ogni volta che una feature è presente senza la sua word-form, la sinapsi inibitoria è aumentata solo leggermente ( $\Delta V_{ij,hk}^{LS} = +\beta_{ij,hk}^{SS} \cdot 0.05$ ). Il risultato finale è che le features che talvolta partecipano all'oggetto (anche non salienti) rimuovono la loro inibizione. Solo quelle features che molto

raramente o mai partecipano all'oggetto, ma spesso partecipano alla semantica di altri oggetti, inviano inibizione all'oggetto 1.

La massima saturazione per la somma delle sinapsi eccitatorie ch

e raggiunge una word-form,  $\Gamma_{\max}$  in Eq. 17, è stata scelta in modo che, quando sono presenti tutte le features salienti, l'attività del neurone word-form,  $x_{ij}^L$ , si trova vicina alla saturazione superiore, ma anche l'assenza di una feature provoca la sua inibizione quasi completa. Questo è possibile grazie alla caratteristica sigmoideale tagliente usata per le unità lessicali. La massima saturazione per le sinapsi inibitorie è stata scelta in modo che anche una singola bolla nella rete semantica, che non partecipa agli oggetti semantici, è in grado di portare l'attività della word-form dalla saturazione superiore all'inibizione.



**Tabella 1***Valori dei parametri numerici utilizzati nelle simulazioni**Rete Semantica*

<b>significato</b>	<b>Simbolo</b>	<b>Valore</b>
Costante di tempo	$\tau$	3ms
Pendenza sigmoide	$p$	40
Posizione sigmoide	$\varphi$	0.55
Forza della eccitazione laterale	$L_0^{EX}$	11
Deviazione standard della eccitazione laterale	$\sigma_{ex}$	0.8
Forza della inibizione laterale	$L_0^{IN}$	3
Deviazione standard della inibizione laterale	$\sigma_{in}$	3.5
Soglia post-sinaptica SS	$\vartheta_{post}$	0.5
Soglia pre-sinaptica SS	$\vartheta_{pre}$	0.05
Tasso di apprendimento SS	$\gamma_0$	0.0025
Massima forza sinaptica SS	$W_{max}$	0.1

*Rete Lessicale*

<b>Significato</b>	<b>Simbolo</b>	<b>Valore</b>
Costante di tempo	$\tau$	1ms
Pendenza sigmoide	$p$	50
Posizione sigmoide	$\varphi$	0.8
Soglia post-sinaptica SL	$\vartheta_{post}$	0.5
Soglia pre-sinaptica SL	$\vartheta_{pre}$	0.0
Tasso di apprendimento SL	$\gamma_0$	0.05
Massima forza sinaptica SL	$W_{max}$	2
Soglia post-sinaptica LS	$\vartheta_{post}$	0.0
Soglia pre-sinaptica LS	$\vartheta_{pre}$	0.5
Tasso di apprendimento LS	$\gamma_0$	0.006
Massima somma delle sinapsi LS	$\Gamma_{max}$	1.0
Soglia post-sinaptica LS	$\rho_{post}$	0.05
Soglia pre-sinaptica LS	$\rho_{pre}$	0.0
Tasso di apprendimento LS	$\beta_0$	0.03
Massima forza sinaptica LS	$V_{max}$	0.06

## Capitolo 3

### 3. RISULTATI DEL MODELLO

Dopo aver eseguito le procedure di apprendimento in cui le features hanno giocato un ruolo diverso (comune vs distintive, saliente vs non salienti). Saranno presentati i risultati della simulazione per quel che riguarda entrambi i compiti di riconoscimento oggetto e compiti di riconoscimento parola. Infine, sarà analizzato il riconoscimento di oggetti con alcune features modificate, quindi sottolineare il ruolo del principio di somiglianza nel modello.

#### 3.1 Simulazione delle attività di riconoscimento oggetto

I valori delle sinapsi ottenuti dopo l'addestramento incorporano le differenze tra features condivise e distintive, e le differenze tra features salienti e non-salienti.

Quindi, il modello addestrato è stato controllato simulando le attività di riconoscimento oggetto. Durante queste prove, è stata fornita alla rete la descrizione incompleta di un oggetto (ad esempio, alcune features sono eccitate dall'ingresso esterno, mentre le altre non sono stimolate) ed è stato verificato che queste altre features sono spontaneamente

evocate nella rete semantica grazie alle sinapsi auto-associative, e se la word-form corrispondente viene evocata nella rete lessicale. Diciamo che un oggetto viene riconosciuto correttamente quando tutte le sue features salienti sono eccitate (anche quelle che non erano originariamente stimulate) e di conseguenza la word-form riceve una sufficiente eccitazione per passare dallo stato "off" allo stato "on". È da notare che l'eccitazione di una feature consiste sempre nell'eccitazione di una intera "bolla" di neuroni centrati nella posizione della feature interessata.

I risultati delle attività di riconoscimento oggetto sono riassunti nella Tabella 3, in cui la prima colonna rappresenta le features utilizzate come ingressi, la seconda tutte le features evocate nella condizione di stato stazionario finale, e la terza l'eventuale word-form evocata nella rete lessicale.

I risultati mostrano che:

- a) se solo alcune features comuni sono utilizzate come ingresso, la rete riconosce correttamente la categoria corrispondente e non evoca le features distintive dei singoli oggetti;
- b) se le features distintive sono stimulate, la rete riconosce l'oggetto corrispondente;
- c) riconoscimenti ai punti a e b si conseguono solo se la rete riceve informazioni sufficienti per recuperare tutte

le features salienti attraverso le sue sinapsi auto-associative. Se viene utilizzato un numero insufficiente di ingressi, la rete non recupera tutte le features salienti e non riconosce l'oggetto corrispondente. Il numero di features necessarie per il riconoscimento dipende dal particolare oggetto (o categoria) e dalla soglia. Con i valori attuali della soglia due features sono sufficienti per riconoscere la maggior parte degli oggetti;

- d) Una feature che non è saliente non è evocata dalle altre features. L'unico modo per eccitare una tale feature è quello di nutrirla da ingresso esterno;
- e) features che sono non-salienti possono ancora svolgere un ruolo nel riconoscimento di oggetti (purché evocate dall'ingresso esterno), in quanto contribuiscono all'eccitazione delle features salienti e, in questo modo, aiutano il riconoscimento di oggetti;
- f) il numero di passi (da cui l'intervallo di tempo complessivo) necessari per evocare una word-form dipende dagli input utilizzati. Generalmente è richiesto un tempo più lungo quando le features non-salienti sono usate come input, dal momento che queste devono evocare le features salienti prima che appaia la word-form corrispondente. Inoltre, le fasi temporali

diminuiscono con il numero di features utilizzate come input.

### 3.2 Simulazione delle attività di riconoscimento di parola

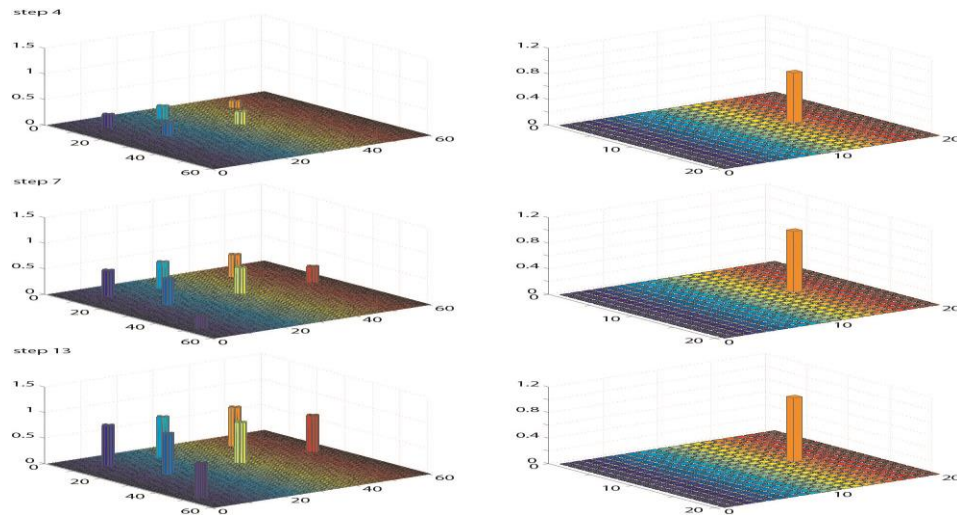
In queste simulazioni si è generata una sola word-form nella rete lessicale, e osservato che le features sono state evocate nella rete semantica. Si può dire che una parola viene riconosciuta correttamente se è in grado di evocare tutte le features salienti corrispondenti. I risultati, riassunti nella tabella 2, confermano quest'ultimo punto: ogni word-form evoca tutte le features salienti, inoltre, forme verbali che rappresentano categorie non evocano le features distintive dei suoi singoli membri.

Word-form	Features
Super-Category	A
Category 1	A, B, C
Category 2	A, D
Object 1	A, B, C, F1, G1, H1
Object 2	A, B, C, F2, G2
Object 3	A, B, C, F3, G3, H3, I3
Object 4	A, D, F4, G4, H4
Object 5	A, D, F5, G5

**Tabella 2:** Risultati di diversi esperimenti di riconoscimento di parola, in cui la word-form viene data come input nella rete lessicale, e le features vengono evocate dalle dinamiche della rete semantica. È fondamentale notare che solo quelle salienti (con percentuale di almeno il 60%) sono evocate.

Un esempio delle simulazioni è riportato nella figura 3, mostrando la risposta della rete alla word-form che rappresenta object3 (Fig. 3). Questo evoca le sue sette features salienti, e la feature marginale L3 non viene evocata.

**Figura 3**



### **3.3 Riconoscimento oggetto con le features alterate**

Una proprietà importante del modello è che le features sono topologicamente organizzate, cioè, features analoghe occupano posizioni prossimali nella rete semantica e tendono ad essere attive insieme nella stessa "bolla di attivazione".

La conseguenza è che un oggetto potrebbe essere rilevato anche se alcune delle sue proprietà sono moderatamente

alterate rispetto al valore "standard" utilizzato durante l'addestramento.

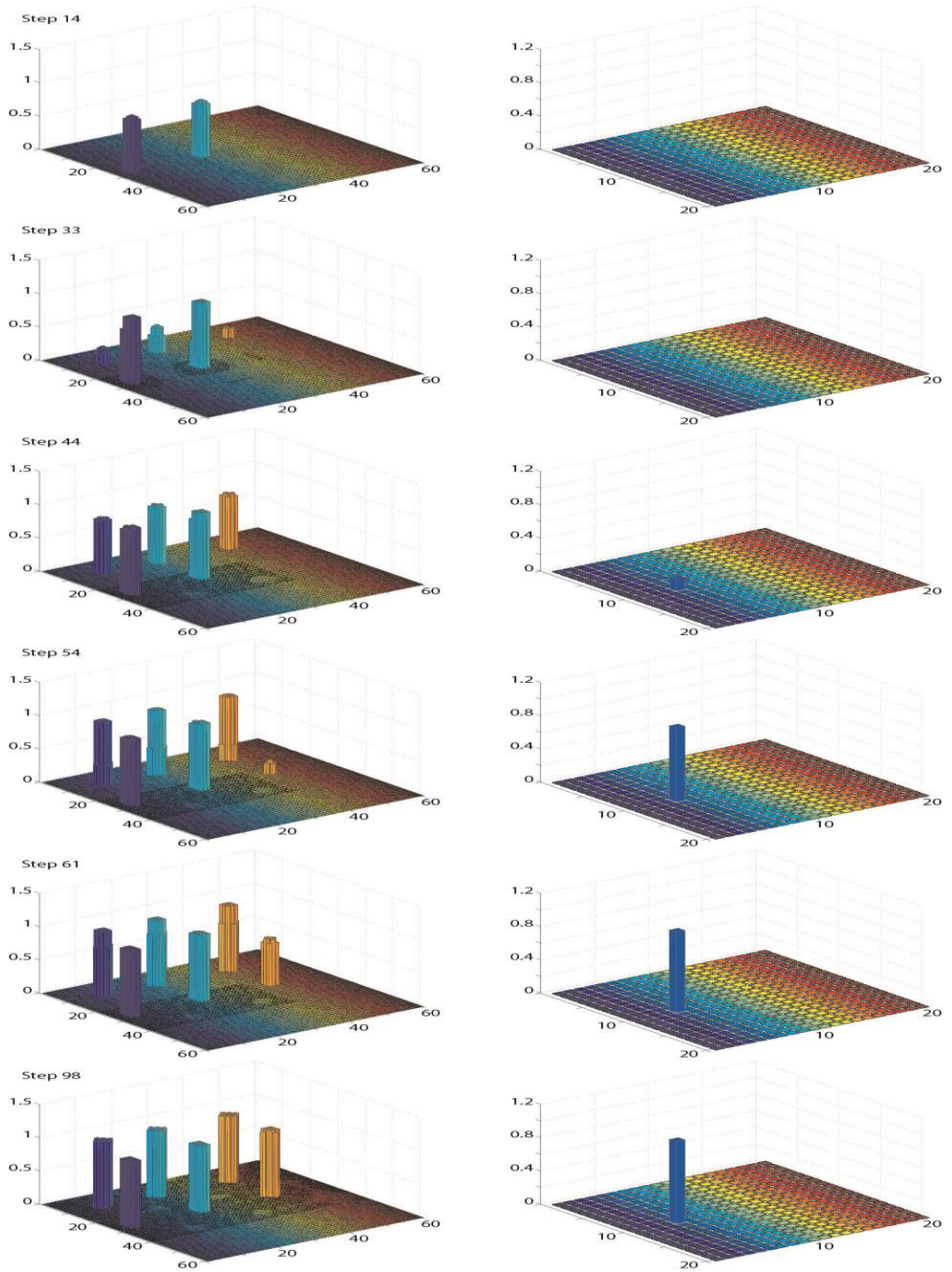
Per testare la capacità del modello di riconoscere un oggetto anche in presenza di un'informazione danneggiata (recuperando tutte le features salienti nella rete semantica ed evocando la word-form corrispondente nella rete lessicale) abbiamo fornito due features come input per la rete semantica, ma con la loro posizione spostata da una o due posizioni rispetto ai valori standard utilizzati durante l'addestramento. Nel seguito considereremo un esempio, in cui Object1 dovrebbe essere riconosciuto dando come input due features F1 e G1, ma con F1 spostata di due posizioni e G1 spostata di una posizione. Risultati simili possono essere ottenuti usando altri oggetti con un paio di features spostate.

Utilizzando gli stessi valori dei parametri indicati nella tabella1, il modello non riesce a riconoscere l'oggetto partendo dalle due features danneggiate. La ragione è che le sinapsi eccitatorie laterali nella rete semantica hanno un'estensione lieve ( $\sigma_{ex} = 0.8$ ), da cui la bolla di attivazione comprende solo le features immediatamente vicine alla "feature centrale". Questa bolla è troppo piccola per spiegare un cambiamento nella posizione della feature.

Al contrario, la rete riesce a riconoscere questo "oggetto danneggiato" se usiamo un'estensione più grande per le sinapsi laterali (in particolare, è utilizzato  $\sigma_{ex} = 1.1$  invece di 0.8). Questo cambiamento può replicare l'effetto di meccanismi attenti, che inducono un'attivazione più ampia nella rete semantica per risolvere questo compito di riconoscimento più difficile. I risultati di questa simulazione sono riassunti in fig. 4. È da notare che, in conseguenza del maggior valore utilizzato per  $\sigma_{ex}$ , le bolle di attivazione sono ora più ampie e comprendono features più distali. Questo provoca l'eccitazione per diffondere dalla feature "corrotta" nella posizione sbagliata alla feature "standard" attiva durante l'addestramento. Questo, tramite sinapsi auto-associative, provoca il conseguente ripristino di tutte le features salienti in altri settori e della word-form. È anche evidente che il compito richiede ora molto più tempo (66 passi di simulazione) rispetto allo stesso compito eseguito con features intatte (46 gradini).



Figura 4



### 3.4 Tempi di risposta del modello

Nel seguito viene testata la velocità del modello nel riconoscimento della parola fornite determinate features, tenendo conto delle differenze tra le features comuni vs distintive e salienti vs non-salienti.

Un esempio di simulazione è il riconoscimento dell'oggetto6: in questo caso solo tre features sono state fornite come input, e sono tutte e tre con salienza più o meno bassa. Quindi è da notare che il tempo necessario per evocare una word-form dipende dagli input utilizzati. Infatti, qui è richiesto un tempo più lungo perchè sono usate features non-salienti come input, dal momento che queste devono evocare le features salienti, attraverso le sinapsi auto-associative, prima che appaia la word-form corrispondente.

Implementazione in Matlab	features	percentage	Stimulated features	Recognition	Time
I(a(1),a(2))=0.;	<b>a</b>	<b>90%</b>	<b>h3, i3, l3</b>	<b>ok</b>	<b>12ms</b>
I(b(1),b(2))=0.;	<b>b</b>	<b>90%</b>			
I(c(1),c(2))=0;	<b>c</b>	<b>70%</b>			
I(f3(1),f3(2))=0.;	<b>f3</b>	<b>100%</b>			
I(g3(1),g3(2))=0.;	<b>g3</b>	<b>80%</b>			
I(h3(1),h3(2))=0.8;	<b>h3</b>	<b>70%</b>			
I(i3(1),i3(2))=0.8;	<b>i3</b>	<b>60%</b>			
I(l3(1),l3(2))=0.8;	<b>l3</b>	<b>40%</b>			

**Tabella 3:** Nella tabella si hanno le features dell'oggetto 6 con le percentuali, ed il tempo di riconoscimento della word-form.

Questa volta, invece, sono state fornite delle features salienti, per cui la word-form corrispondente emerge in un tempo minore

Implementazione in Matlab	features	percentage	Stimulated features	Recognition	Time
I(a(1), a(2))=0.8;	<b>a</b>	<b>90%</b>	<b>a, b, f3, g3</b>	<b>ok</b>	<b>8.4 ms</b>
I(b(1), b(2))=0.8;	<b>b</b>	<b>90%</b>			
I(c(1), c(2))=0;	<b>C</b>	<b>70%</b>			
I(f3(1), f3(2))=0.8;	<b>f3</b>	<b>100%</b>			
I(g3(1), g3(2))=0.8;	<b>g3</b>	<b>80%</b>			
I(h3(1), h3(2))=0.;	<b>h3</b>	<b>70%</b>			
I(i3(1), i3(2))=0.;	<b>i3</b>	<b>60%</b>			
I(l3(1), l3(2))=0.;	<b>l3</b>	<b>40%</b>			

**Tabella 4:** Nella tabella si hanno le features dell'oggetto 6, di cui fornite quelle con percentuale di occorrenza più alta, che sono quelle salienti.

Anche in questa simulazione per l'oggetto 7, ho eseguito i due casi: il primo con delle features meno salienti come input, ed è ciò che vediamo nella seguente tabella:

Implementazione in Matlab	features	percentage	Stimulated features	Recognition	Time
I(a(1), a(2))=0.;	<b>a</b>	<b>90%</b>	<b>i4, l4, m4</b>	<b>ok</b>	<b>14.4 ms</b>
I(d(1), d(2))=0.;	<b>d</b>	<b>90%</b>			
I(e(1), e(2))=0.;	<b>e</b>	<b>50%</b>			
I(f4(1), f4(2))=0.;	<b>f4</b>	<b>90%</b>			
I(g4(1), g4(2))=0.;	<b>g4</b>	<b>90%</b>			
I(h4(1), h4(2))=0.;	<b>h4</b>	<b>60%</b>			
I(i4(1), i4(2))=0.8;	<b>i4</b>	<b>40%</b>			
I(l4(1), l4(2))=0.8;	<b>l4</b>	<b>50%</b>			
I(m4(1), m4(2))=0.8;	<b>m4</b>	<b>40%</b>			

**Tabella 5:** Nella tabella si hanno le features dell'oggetto 7 con le percentuali, ed il tempo di riconoscimento della word-form.

Mentre nel secondo caso sono fornite features con alta salienza. Infatti il tempo di emergenza word-form è minore.

Inoltre, le fasi temporali dipendono anche dal numero delle features fornite, in quanto più features sono date come ingresso minore sarà il tempo impiegato nel riconoscimento della word-form corrispondente nella rete lessicale.

Implementazione in Matlab	features	percentage	Stimulated features	Recognition	Time
I(a(1), a(2))=0.8;	<b>a</b>	<b>90%</b>	<b>a, d, f4, g4, h4</b>	<b>ok</b>	<b>10 ms</b>
I(d(1), d(2))=0.8;	<b>d</b>	<b>90%</b>			
I(e(1), e(2))=0.;	<b>e</b>	<b>50%</b>			
I(f4(1), f4(2))=0.8;	<b>f4</b>	<b>90%</b>			
I(g4(1), g4(2))=0.8;	<b>g4</b>	<b>90%</b>			
I(h4(1), h4(2))=0.8;	<b>h4</b>	<b>60%</b>			
I(i4(1), i4(2))=0.;	<b>i4</b>	<b>40%</b>			
I(l4(1), l4(2))=0.;	<b>l4</b>	<b>50%</b>			
I(m4(1), m4(2))=0.;	<b>m4</b>	<b>40%</b>			

**Tabella 6:** Nella tabella si hanno le features dell'oggetto 7, di cui fornite quelle con percentuale di occorrenza più alta, che sono quelle salienti.

## Capitolo 4

### 4. DESCRIZIONE DATABASE

Il database utilizzato in questo lavoro di tesi, ricevuto dal Dipartimento di Neuropsicologia dell'ospedale San Raffaele a Milano, è costituito dai dati relativi ad oggetti reali. Per produrre tale database, un gruppo di soggetti sani si è espresso enunciando gli attributi caratteristici di alcuni oggetti, appartenenti a categorie differenti.

Di conseguenza, è una grande raccolta di informazioni varie, ed è una tabella con le righe contenenti i diversi concetti ovvero gli oggetti (gatto, cane...), che sono stati presentati come input queste persone per avere in output gli aspetti e le idee stimulate nel loro cervello; in altre parole, data la parola (cane), dovevano dire ciò che pensavano riguardo il cane (abbaia, animale...ecc) e queste sono le famose features.

Nel database i concetti sono stati classificati in categorie (veicoli, animali, verdure...) e in viventi / non viventi, mentre le features sono state classificate nella colonna "Etichette" in forma, colore, gusto, tatto, olfatto, funzionale, enciclopedico e tassonomico. Mentre una elaborazione statistica ha permesso di classificare le features assegnando delle percentuali a ciascuna feature nelle seguenti grandezze:

**Distintività:** rappresenta una feature tipica proprio ad un concetto.

**Dominanza:** è la frequenza di una feature in un concetto normalizzata al numero di persone e può essere utilizzata come output del modello.

**Rilevanza semantica:** è l'importanza di una feature in un concetto, ed è una combinazione non lineare della dominanza e la distintività.

**Frequenza totale:** è la frequenza con la quale una feature si presenta negli oggetti. Nel database, quando una feature è distintiva presenta un valore di frequenza totale molto basso.

**Accessibilità:** è la rappresentazione delle features in ordine di importanza semantica in un concetto, potrebbe essere utilizzata come output del modello.

Per aver un'idea migliore del database sopra descritto, presento una sua immagine:

**Figura 5**

FEATUES	CONCETTO	L/NL	categoria	ETICHETTE	sens/non sens	DOMINANZA NORM	FREQ tot	DISTINTIVITA' Wn	RILEVANZA SEMANTICA	frequenza per categoria	distintività_garrard	
atterra	aereo	NL	veicoli	enciclopedica	ns		5	2	41	26.79	2	0.40
decolla	aereo	NL	veicoli	enciclopedica	ns		4	2	41	21.43	2	0.40
è allungato	aereo	NL	veicoli	forma	s		6	13	6	15.94	1	0.20
è di metallo	aereo	NL	veicoli	forma	s		5	18	5	10.94	4	0.80
è grande	aereo	NL	veicoli	forma	s		7	16	5	16.50	2	0.40
è pilotato	aereo	NL	veicoli	funzionale	ns		8	2	41	42.86	2	0.40
è rumoroso	aereo	NL	veicoli	suono	s		3	5	16	12.11	3	0.60
è un mezzo di trasporto	aereo	NL	veicoli	tassonomico	ns		20	6	14	75.45	5	1.00
è usato per viaggiare	aereo	NL	veicoli	funzionale	ns		3	2	41	16.07	2	0.40
è veloce	aereo	NL	veicoli	enciclopedica	ns		6	6	14	22.64	1	0.20
ha due ali	aereo	NL	veicoli	forma	s		4	2	41	21.43	1	0.20
ha gli assistenti di bordo	aereo	NL	veicoli	enciclopedica	ns		3	1	82	19.07	1	0.20
ha i finestrini	aereo	NL	veicoli	forma	s		5	2	41	26.79	2	0.40
ha il carrello	aereo	NL	veicoli	forma	s		3	1	82	19.07	1	0.20
ha il motore a reazione	aereo	NL	veicoli	enciclopedica	s		3	1	82	19.07	1	0.20
ha il motoreli	aereo	NL	veicoli	forma	s		12	4	21	52.29	4	0.80
ha i sedileli	aereo	NL	veicoli	forma	s		3	2	41	16.07	2	0.40
ha l'elica	aereo	NL	veicoli	forma	s		3	2	41	16.07	2	0.40
ha la cabina	aereo	NL	veicoli	forma	s		9	2	41	48.22	2	0.40
ha la cabina di pilotaggio	aereo	NL	veicoli	forma	s		9	2	41	48.22	2	0.40
ha la coda	aereo	NL	veicoli	forma	s		3	19	4	6.33	1	0.20
ha le ali	aereo	NL	veicoli	forma	s		20	7	12	71.00	1	0.20
ha le ruote	aereo	NL	veicoli	forma	s		3	4	21	13.07	4	0.80
si trova in aereoporto	aereo	NL	veicoli	enciclopedica	ns		8	1	82	50.86	1	0.20
trasporta	aereo	NL	veicoli	funzionale	ns		10	8	10	33.58	5	1.00

Ovviamente, in questo lavoro di tesi, ho utilizzato solo alcuni concetti con alcune features per formare la tassonomia che sarà l'input del modello semplificato.

## Capitolo 5

### 5. MODELLO SEMPLIFICATO

#### 5.1 Descrizione del modello

Il modello che sto per illustrare in questo capitolo è semplicemente quello descritto precedentemente, con delle semplificazioni ad esso apportate. Il modello incorpora sempre due reti di neuroni, la rete semantica e quella lessicale. Gli oggetti sono rappresentati come un insieme di features sparse non più in 9 aree corticali, ognuna costituita da 20x20 neuroni; ma semplicemente in un vettore di lunghezza  $N$ , dove ogni feature riceve delle sinapsi da tutte le altre formando una matrice delle sinapsi  $N \times N$ ; di conseguenza sono state eliminate le sinapsi laterali; quindi non ho più una matrice di matrici, ma lavoro con una semplice matrice, così ogni unità neurale viene indicata con uno solo indice ( $i$ ) invece di due ( $i, j$ ); mentre una sinapsi tra due unità neurali avrà solo due pedici  $ij$ , il primo rappresenta la posizione del neurone postsinaptico, il secondo la posizione del neurone presinaptico.

L'attività del neurone è descritta dalla stessa equazione differenziale di primo grado nel tempo, con un indice solo ( $X_i$ ):

$$\tau^A \frac{d}{dt} x_i^A(t) = -x_i^A(t) + H^A(u_i^A(t)) \quad A = S, L$$

dove l' apice  $A$  indica la rete (semantica o lessicale),  $\tau^A$  è la costante di tempo, che determina la velocità di risposta allo stimolo, e  $H^A(u^A(t))$  è una funzione di attivazione

sigmoideale che modula gli stimoli in ingresso al neurone ed è descritta dalla seguente equazione:

$$H^A(u^A(t)) = \frac{1}{1 + e^{-(u^A(t) - \varphi^A) p^A}};$$

dove  $p$  è un parametro che imposta la pendenza centrale della sigmoide, e  $\varphi$  imposta la sua posizione centrale. Questa equazione assume convenzionalmente che l'attività massima sia 1.

$u_i^A(t)$  è l'ingresso totale raggiunto dal neurone, e ha diverse espressioni nelle reti semantiche e lessicali.

L'ingresso ai neuroni nella rete semantica (apice  $A = S$ ) viene calcolato come la somma di tre contributi perché manca il quarto contributo che riguardava le sinapsi laterali nel modello precedente.

$$u_i^S(t) = I_i^S(t) + E_i^{SS}(t) + C_i^{SL}(t)$$

$I_i^S$  rappresenta l'ingresso esterno per l'unità neurale in posizione  $i$ , può assumere il valore 0 o 1.  $E_i^{SS}$  rappresenta un termine di accoppiamento eccitatorio proveniente da unità in altre aree della rete semantica; mentre  $C_i^{SL}$  è un termine di cross-rete, ed è calcolato come segue:

$$C_i^{SL} = \sum_h W_{i,h}^{SL} \cdot x_h^L$$



Dove  $x_h^L$  rappresenta l'attività del neurone  $h$  nell'area lessicale ed i simboli  $W_{i,h}^{SL}$  sono sinapsi dal lessicale alla rete semantica.

Le condizioni di accoppiamento all'interno della rete semantica hanno la seguente espressione:

$$E_i^{SS} = \sum_h W_{i,h}^{SS} \cdot x_h^S$$

dove  $i$  indica la posizione del neurone postsinaptico, e  $h$  la posizione del neurone presinaptico, e le somme si estendono a tutti i neuroni presinaptici nella rete semantica.

Il simbolo  $W_{i,h}^{SS}$  rappresenta sinapsi eccitatorie nella semantica, e sono oggetto di apprendimento.

L'ingresso, invece, all'unità neurale lessicale  $u_i^L(t)$  comprende solo due termini in quanto, nel presente modello, non consideriamo collegamenti laterali tra unità lessicali. Quindi:

$$u_i^L(t) = I_i^L(t) + C_i^{LS}(t) \quad (11)$$

$I_i^L(t)$  è l'ingresso prodotto da una stimolazione linguistica esterna.  $I_i^L(t)$  assume valore 1 quando la word-form è data alla rete, e zero altrimenti.  $C_i^{LS}(t)$  rappresenta l'intensità

dell'ingresso causato da connessioni sinaptiche dalla rete semantica.

Sinapsi dalla semantica alla rete lessicale includono sia un termine eccitatorio che un termine inibitorio ( $W_{ij,hk}^{LS}$  e  $V_{ij,hk}^{LS}$ ) che sono addestrati in modi diversi. Ciò comporta una strategia inibitoria-eccitatoria più complessa. In realtà, abbiamo bisogno che una word-form nella zona lessicale sia entusiasta quando sono presenti tutte le features salienti, ma viene inibita quando si presenta una feature che non appartiene all'oggetto. Quindi:

$$C_i^{LS} = \sum_h W_{i,h}^{LS} \cdot x_h^S - \sum_h V_{i,h}^{LS} \cdot x_h^S$$

Dove  $x_h^S$  rappresenta l'attività del neurone h nella rete semantica,  $W_{i,h}^{LS}$  è la forza delle sinapsi eccitatorie e  $V_{i,h}^{LS}$  la forza delle sinapsi inibitorie.

## 5.2 Addestramento delle sinapsi

L'addestramento, la memorizzazione e la ricostruzione di pattern di attivazione, quindi l'allenamento delle sinapsi avviene sempre utilizzando la regola di Hebb; in modo che, alla fine dell'addestramento, se fornisco una versione danneggiata, la rete sia in grado di ricostruire la versione

originaria del pattern stesso che sarà memorizzato nella rete come punto di equilibrio.

All'inizio dell'addestramento, come sappiamo, tutte le sinapsi eccitatorie nella rete semantica ( $W_{i,h}^{SS}$ ) e tutte le sinapsi eccitatorie ed inibitorie tra la semantica e la rete lessicale

( $W_{i,h}^{SL}$ ,  $W_{i,h}^{LS}$  e  $V_{i,h}^{LS}$ ) sono fissate a zero.

La regola di addestramento di Hebb è la seguente:

$$\Delta W_{i,h}^{AB} = \gamma_{i,h}^{AB} (x_i^A - \mathcal{G}_{post}^{AB})(x_h^B - \mathcal{G}_{pre}^{AB})$$

dove gli apici AB possono assumere il significato SS, SL o LS a seconda della particolare sinapsi,  $\Delta W_{i,h}^{AB}$  rappresenta la variazione della forza sinaptica, a causa delle attività pre e post-sinaptiche,  $\mathcal{G}_{post}^{AB}$  e  $\mathcal{G}_{pre}^{AB}$  sono soglie per le attività post e presinaptiche,  $\gamma^{AB}$  denota un fattore di apprendimento,  $x_i^A$  è l'attività dell'unità neurale in posizione i dell'area post-sinaptica, e  $x_h^B$  è l'attività dell'unità neurale in posizione h nell'area pre-sinaptica.

Affinchè sia fisiologica la regola di Hebb, ha bisogno di alcuni aggiustamenti:

- quando entrambe le attività presinaptica e postsinaptica sono basse, non dovrebbe verificarsi nessun cambiamento di peso. Quindi:

$$\text{se } x_i^A < \mathcal{G}_{post}^{AB} \text{ and } x_h^B < \mathcal{G}_{pre}^{AB} \text{ allora } \Delta W_{i,h}^{AB} = 0$$

- le sinapsi non possono aumentare all'infinito, ma devono raggiungere un livello massimo di saturazione. Questo risultato è ottenuto riducendo progressivamente il tasso di apprendimento quando le sinapsi si avvicinano al loro massimo ( $W_{\max}^{AB}$ ). Si ha:

$$\gamma_{i,h}^{AB} = \frac{\gamma_0^{AB}}{W_{\max}^{AB}} (W_{\max}^{AB} - W_{i,h}^{AB})$$

dove  $\gamma_0$  è il tasso di apprendimento quando la sinapsi è zero.

Infine, una sinapsi non può diventare negativa. Quindi, quando calcola il nuovo valore della sinapsi si ha:

$$W_{i,h}^{AB} \leftarrow (W_{i,h}^{AB} + \Delta W_{i,h}^{AB}) U(W_{i,h}^{AB} + \Delta W_{i,h}^{AB})$$

dove il simbolo  $\leftarrow$  significa che il valore calcolato a destra è assegnato ai membri di sinistra e  $U(y)$  rappresenta la funzione a gradino ( $U(y) = 1$  se  $S > 0$ , altrimenti  $U(y) = 0$ ).

Infine, un ulteriore aspetto è stato incluso per quanto riguarda le sinapsi eccitatorie dalla semantica alla rete lessicale ( $W_{i,h}^{LS}$ ).

Infatti, si vuole che una word-form nella rete lessicale sia eccitata, se e solo se tutte le features salienti sono attive nella

rappresentazione semantica, indipendentemente dal loro numero. Per giustificare questa proprietà essenziale, sono state normalizzate le sinapsi che entrano in una word-form assumendo che la loro somma non possa superare un livello massimo. Quindi, la seguente regola viene utilizzata per aggiornare le sinapsi  $W_{i,h}^{LS}$

$$\Gamma_i^{LS} = \sum_h (W_{i,h}^{LS} + \Delta W_{i,h}^{LS}) U(W_{i,h}^{LS} + \Delta W_{i,h}^{LS})$$

$$W_{i,h}^{LS} \leftarrow \begin{cases} (W_{i,h}^{LS} + \Delta W_{i,h}^{LS}) U(W_{i,h}^{LS} + \Delta W_{i,h}^{LS}) & \text{if } \Gamma_i^{LS} \leq \Gamma_{\max} \\ (W_{i,h}^{LS} + \Delta W_{i,h}^{LS}) U(W_{i,h}^{LS} + \Delta W_{i,h}^{LS}) \cdot \Gamma_{\max} / \Gamma_i^{LS} & \text{if } \Gamma_i^{LS} > \Gamma_{\max} \end{cases}$$

dove il simbolo  $\Gamma_i^{LS}$  indica la somma di tutte le sinapsi eccitatorie introdotte nella word-form in posizione i; quindi, la somma è estesa a tutte le unità neurali h nella rete semantica.

Le sinapsi inibitorie  $V_{ij,kh}^{LS}$  sono addestrate con una regola anti-Hebbiana, cioè, sono indebolite quando entrambe le attività pre e post-sinaptica sono al di sopra della soglia, e sono rafforzate quando le attività sono negativamente correlate. Quindi

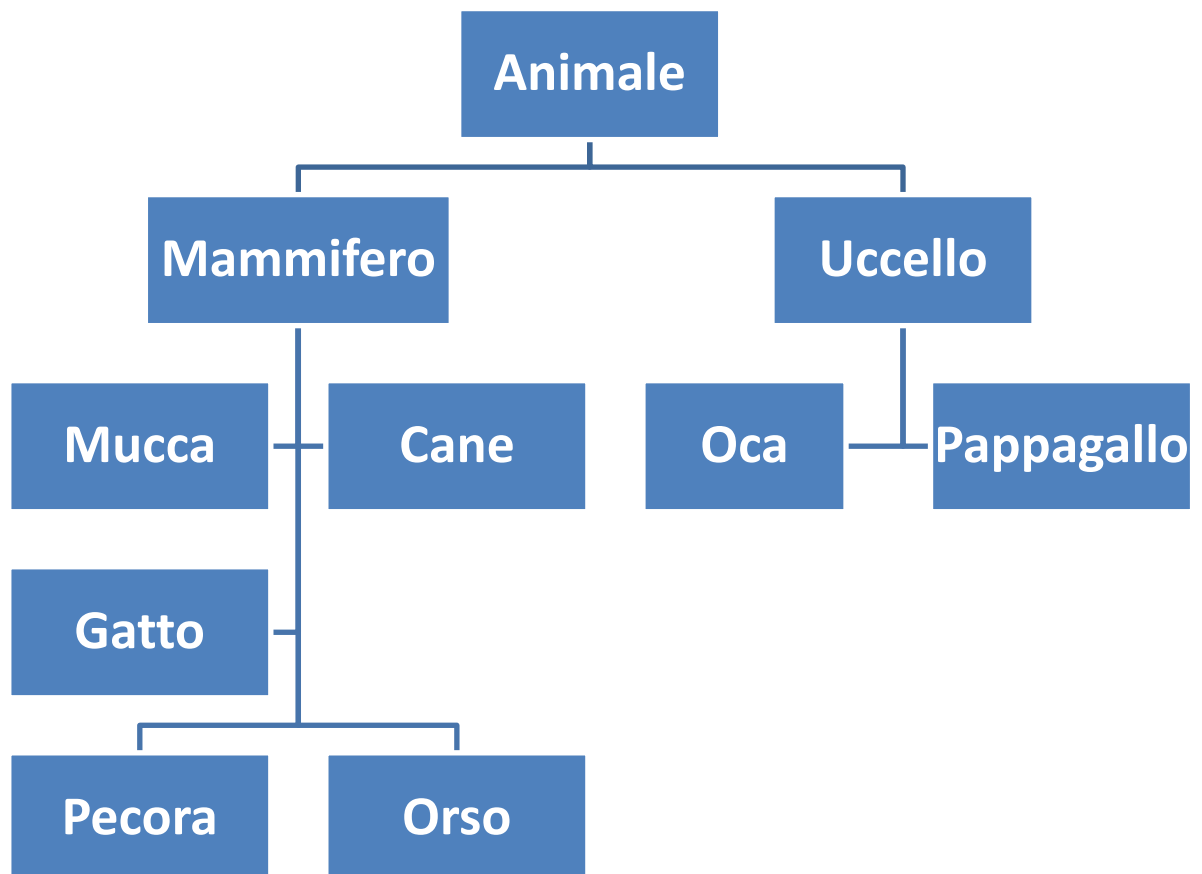
$$\Delta V_{ij,hk}^{LS} = -\beta_{ij,hk}^{LS} (x_{ij}^L - \rho_{post}^{LS}) (x_{hk} - \rho_{pre}^{LS})$$

dove i simboli  $\rho$  e  $\beta$  sono usati per la soglia e il tasso di apprendimento. Naturalmente, queste sinapsi non possono cambiare il loro segno, e hanno un livello massimo di

saturazione. Queste ultime equazioni sono omesse per brevità, essendo analoghe alle precedenti.

### 5.3 Risultati del modello con una tassonomia

La tassonomia utilizzata nella rete semantica, è costituita da una super-categoria da cui si dipartono due sottocategorie costituite dai singoli membri; ognuno di questi membri è costituito da un determinato numero di features.



**Figura 8:** nella figura è illustrata la tassonomia utilizzata: è costituita da una supercategoria che si suddivide in due sottocategorie, che a loro volta si suddividono in singoli membri, ognuno dei quali è descritto da un determinato numero di features.

**Tabella 7**

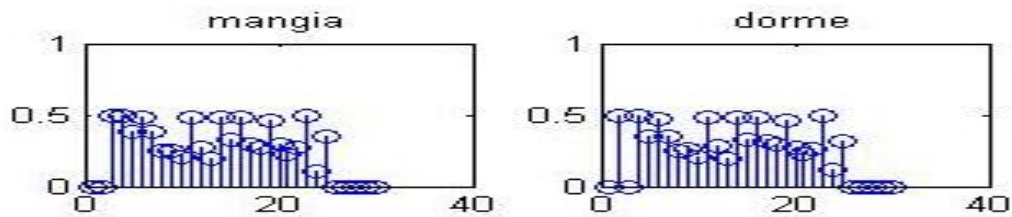
*Nella tabella sono illustrate le features con le corrispondenti posizioni nella rete semantica e le percentuali di occorrenza.*

<b>Oggetto</b>	<b>Feature</b>	<b>Posizione</b>	<b>Percentuale</b>
Super categoria	mangia Dorme	2	80%
		3	80%
Da terra 4zampe	Ha il pelo Ha quattro zampe	4	80%
		23	70%
Uccelli	Ha le piume Ha due zampe vola	14	80%
		16	90%
		11	80%
Cane	Abbaia Scodinzola domestico	25	80%
		5	90%
		6	80%
Gatto	miagola domestico Fa le fusa	22	70%
		6	80%
		7	90%
Orso	Va in letargo pericoloso polare	8	70%
		9	70%
		24	70%
Pecora	Mangia erba bela Si tosa	19	70%
		18	70%
		17	70%
Mucca	Ha le corna muggisce Mangia erba	21	60%
		20	80%
		19	70%
Pappagallo	Ripete i suoni Ha il becco grosso	10	50%
		13	60%
Oca	Ha il becco arancione starnazza	12	75%
		15	80%

In seguito, osserviamo simulazioni tra unità semantiche:

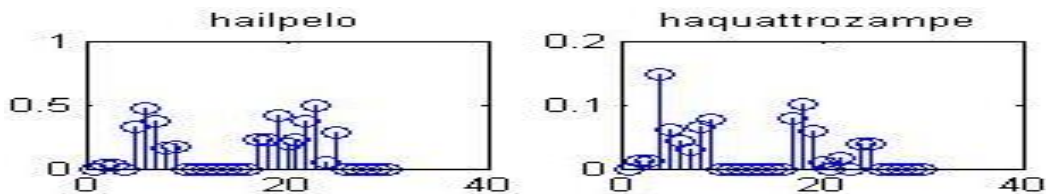
Mangia 2	Va in letargo 8	Ha le piume 14	Muggisce 20
Dorme 3	Pericoloso 9	Starnazza 15	Ha le corna 21
Ha il pelo 4	Ripete i suoni 10	Ha due zampe 16	Miagola 22
Scodinzola 5	Vola 11	Si tosa 17	Ha quattro zampe 23
Domestico 6	Ha il becco arancione 12	Bela 18	Polare=24
Fa le fusa 7	Ha il becco grosso 13	Mangia erba 19	Abbaia=25

### Super-categoria “animale”



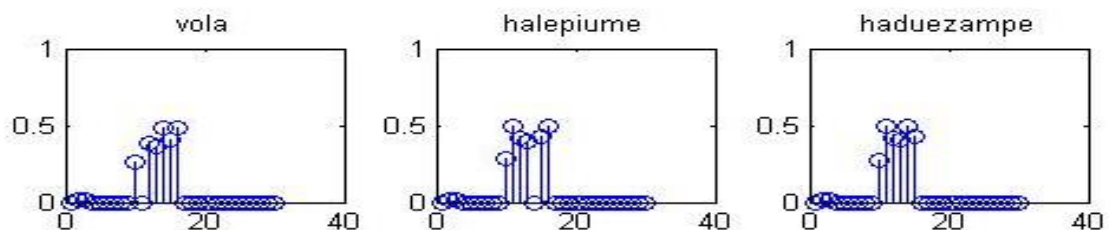
le features condivise “mangia” e “dorme” sono chiamate da tutte le features della super-categoria “animale”.

### Sottocategoria “mammifero”



“ha il pelo” e “ha quattro zampe” ricevono sinapsi dalle features della sottocategoria “mammifero”.

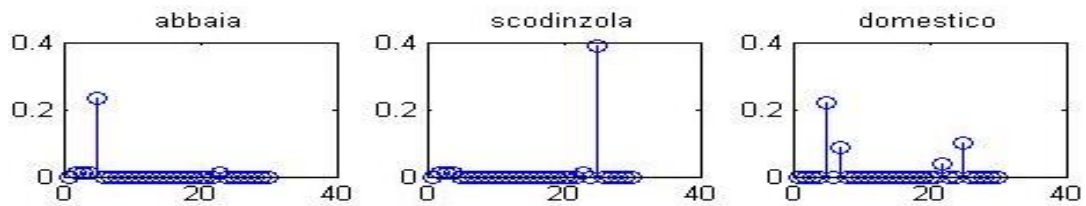
### Sottocategoria “uccello”



Le tre features comuni agli uccelli “ha le piume”, “ha due zampe” e “vola” ricevono sinapsi dalle altre features della sottocategoria uccello.

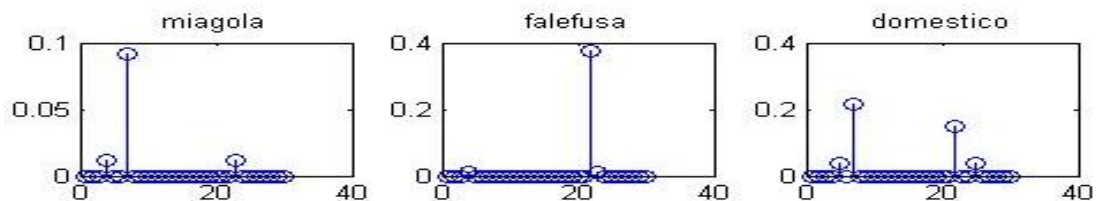


## Cane



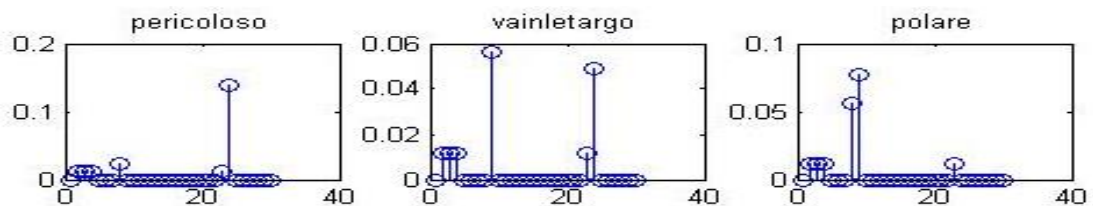
Le due features distintive “scodinzola” e “abbaia” si chiamano a vicenda, “domestico” è chiamato dal cane e il gatto.

## Gatto



Le features distintive “miagola” e “fa le fusa” si stimolano a vicenda. Inoltre, la feature “miagola” è stimolata anche da qualche feature comune saliente come “mangia” e “ha quattro zampe”.

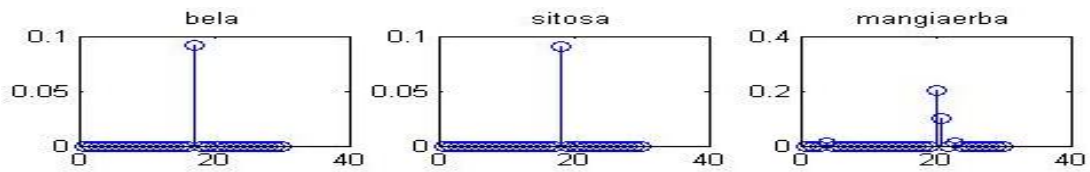
## Orso



In questa immagine, si hanno delle sinapsi tra le features distintive “va in letargo”, “pericoloso” e “polare”. Esse chiamano anche qualche altra feature comune perché hanno salienza alta.

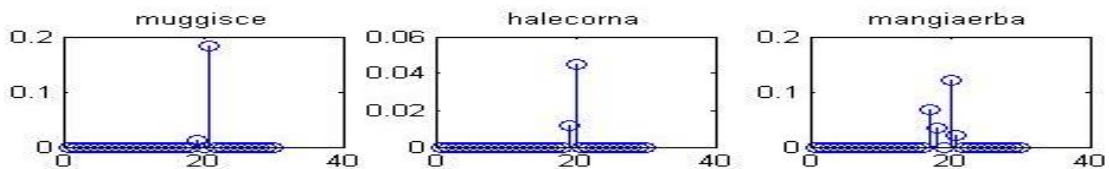
Mangia 2	Va in letargo 8	Ha le piume 14	Muggisce 20
Dorme 3	Pericoloso 9	Starnazza 15	Ha le corna 21
Ha il pelo 4	Ripete i suoni 10	Ha due zampe 16	Miagola 22
Scodinzola 5	Vola 11	Si tosa 17	Ha quattro zampe 23
Domestico 6	Ha il becco arancione 12	Bela 18	Polare=24
Fa le fusa 7	Ha il becco grosso 13	Mangia erba 19	Abbaia=25

## Pecora



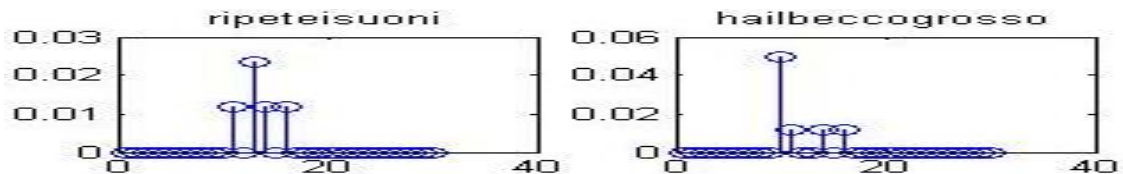
Anche qui, si creano sinapsi tra le features distintive “si tosa” e “bela”, la feature “mangia erba” è richiamata dalle features riguardanti la mucca e la pecora.

## Mucca



Si sono create delle sinapsi tra le features distintive “muggisce” e “ha le corna”. Queste chiamano leggermente anche la features comune saliente “mangiaerba”.

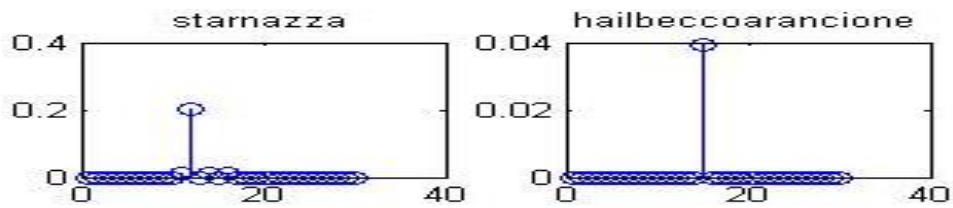
## Pappagallo



La feature distintiva “ripete i suoni” è stimolata dalla feature distintiva “ha il becco grosso” e dalle features comuni salienti “vola” e “ha le piume” e “ha due zampe”. Le stesse features comuni salienti stimolano la feature “ha il becco grosso”.

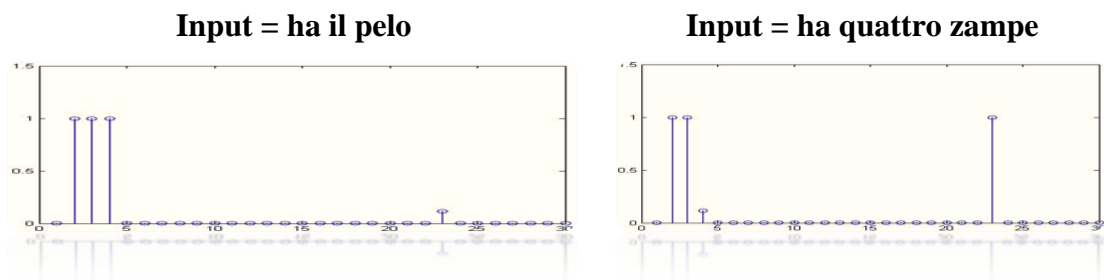
Mangia 2	Va in letargo 8	Ha le piume 14	Muggisce 20
Dorme 3	Pericoloso 9	Starnazza 15	Ha le corna 21
Ha il pelo 4	Ripete i suoni 10	Ha due zampe 16	Miagola 22
Scodinzola 5	Vola 11	Si tosa 17	Ha quattro zampe 23
Domestico 6	Ha il becco arancione 12	Bela 18	Polare=24
Fa le fusa 7	Ha il becco grosso 13	Mangia erba 19	Abbaia=25

## Oca

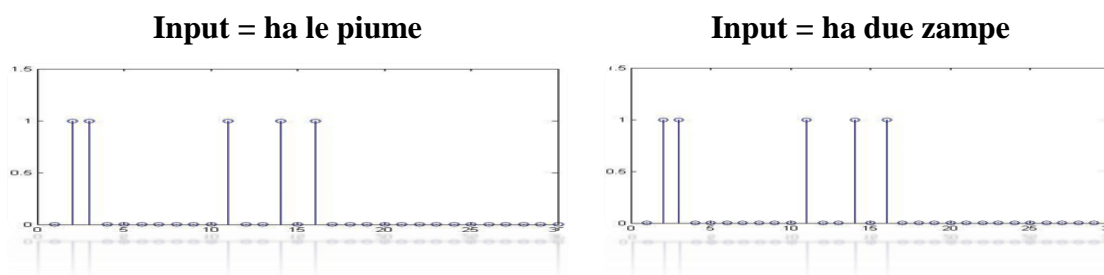


In quest' ultima immagine vediamo che le due distintive “starnazza” e “ha il becco arancione” si chiamano.

Ora, considerando le features di un determinato oggetto, provo a fornire in ingresso qualcuna per osservare quali attività neurali sono evocate nella rete semantica:



Quando fornisco la feature “ha il pelo” si creano sinapsi tra questa e le features “mangia”, “dorme” e “ha quattro zampe”; lo stesso succede per l’immagine a destra: le features condivise sono chiamate dalle altre condivise salienti della stessa categoria di appartenenza.

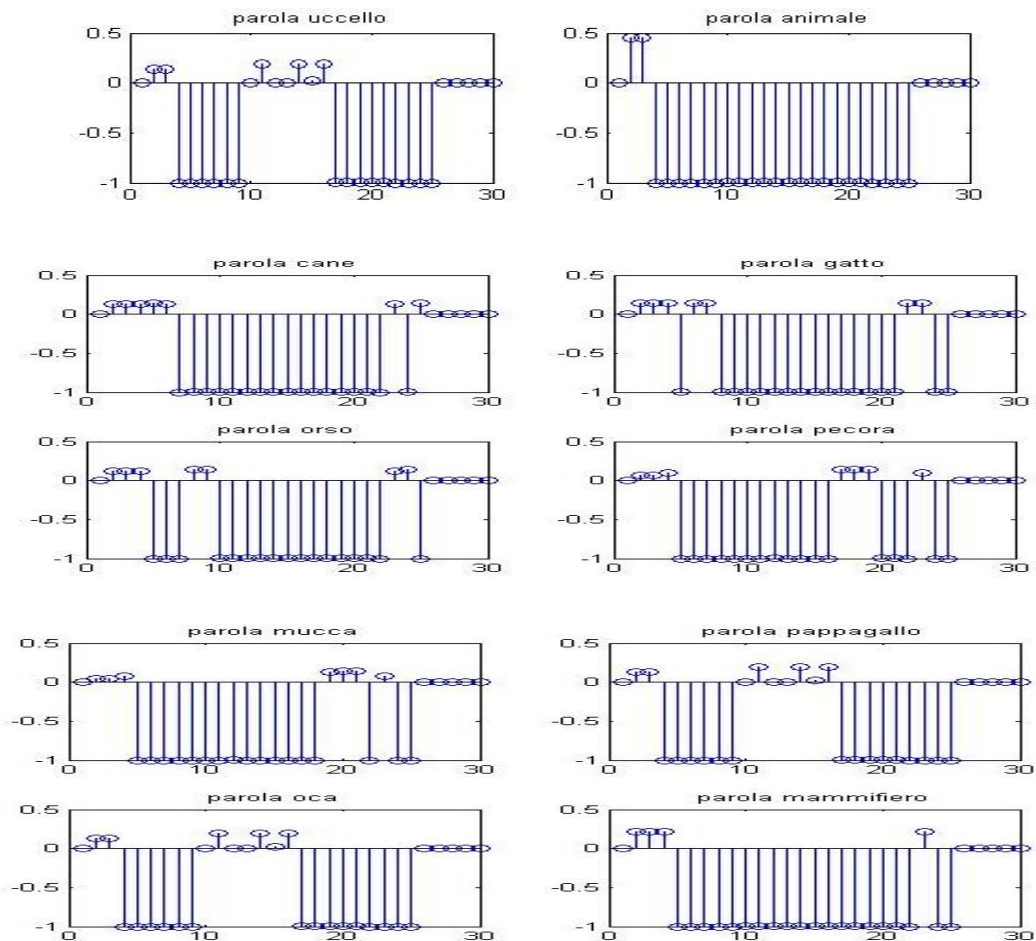


Dando in input la feature “ha le piume” sono evocate le altre condivise salienti della sottocategoria uccello: “mangia”, “dorme”, “vola”, “ha le piume” e “ha due zampe”.

Viste le simulazioni tra le unità della rete semantica, ora vediamo come si comportano le sinapsi tra le due reti semantica e lessicale. In primo luogo, osserviamo le sinapsi semantico-lessicali “WLC”, in altre parole quali features della rete semantica stimolano quali parole della rete lessicale:

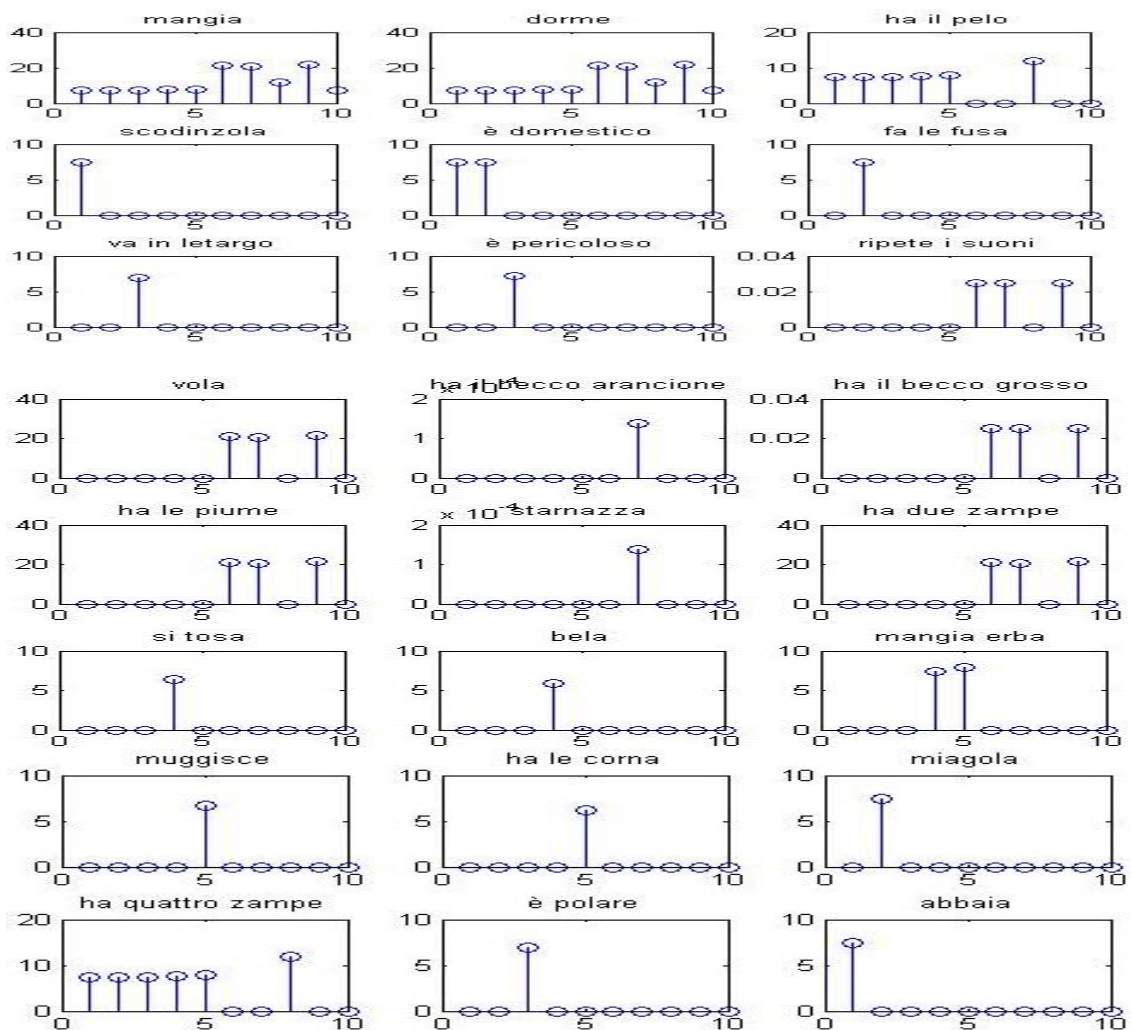
Cane	Gatto	Orso	Pecora	Mucca	Pappagallo	Oca	Mammifero	Uccello	Animale
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>

Mangia 2	Va in letargo 8	Ha le piume 14	Muggisce 20
Dorme 3	Pericoloso 9	Starnazza 15	Ha le corna 21
Ha il pelo 4	Ripete i suoni 10	Ha due zampe 16	Miagola 22
Scodinzola 5	Vola 11	Si tosa 17	Ha quattro zampe 23
Domestico 6	Ha il becco arancione 12	Bela 18	Polare=24
Fa le fusa 7	Ha il becco grosso 13	Mangia erba 19	Abbaia=25



In queste immagini, osserviamo che ogni word-form è stimolata dalle features corrispondenti nella rete semantica: la supercategoria è stimolata dalle features condivise “mangia” e “dorme”, le sottocategorie sono stimolate dalle features comuni agli oggetti che le compongono ed i singoli membri sono stimolati dalle loro features comuni e distintive. Ma la sottocategoria “uccello” e le word-forms “pappagallo” e “oca” sono stimolate dalle stesse features e non dovrebbe essere così.

In fine, osserviamo le sinapsi lessico-semantiche “WCL”, cioè quali word-forms nella reta lessicale stimolano quali features nella rete semantica:

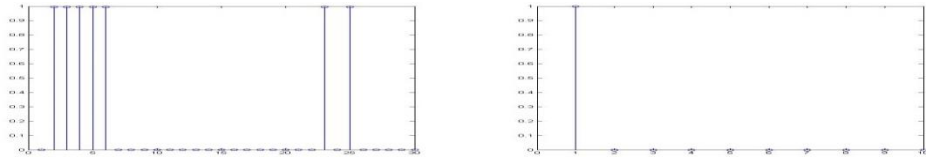


In questi risultati, avendo le word-forms sulle ascisse, notiamo che esse stimolano le corrispondenti features nella rete semantica. In particolare, le features comuni sono stimulate da più oggetti, mentre le features distintive sono stimulate dal singolo membro corrispondente. L'unica situazione indesiderata sta nella sottocategoria "uccello", in quanto le features "ripete i suoni" e "ha il becco grosso" dovrebbero essere stimulate solo dalla word-form "pappagallo"; invece, quello che succede è che sono stimulate anche dalla word-form "uccello". Questo non va bene, perché per il modello queste due word-forms sono equivalenti ed il motivo potrebbero essere le basse percentuali assegnate alle features. Lo stesso problema si presentava con la features "ha il becco arancione" che dopo un aumento della percentuale (da 50% a 75%), la feature è stimolata solo dalla word-form corrispondente "oca".

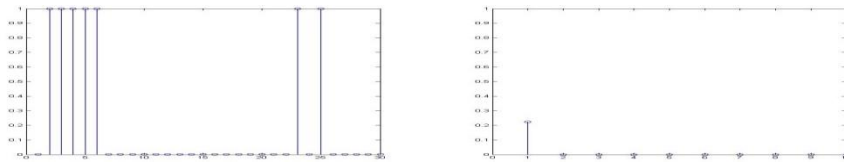
Osserviamo le ultime due simulazioni sulla parte lessicale:

- Nella prima fornisco in ingresso la word-form "cane", e come vediamo nell'immagine seguente, sono chiamate tutte le features corrispondenti.

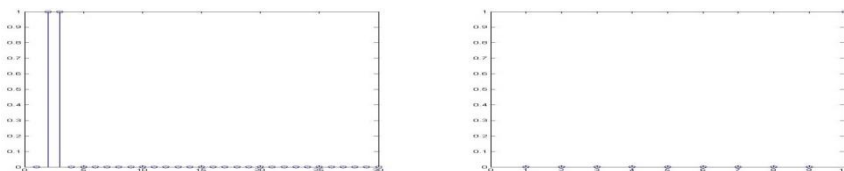
Cane	Gatto	Orso	Pecora	Mucca	Pappagallo	Oca	Mammifero	Uccello	Animale
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>



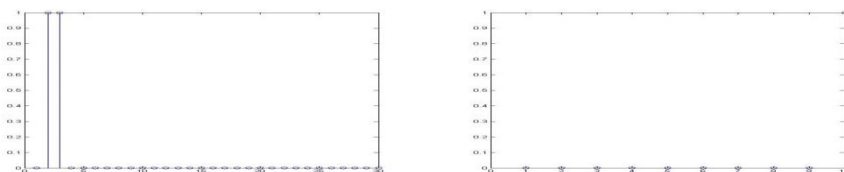
Adesso fornisco, invece, alcune features: “scodinzola”, “ha il pelo” e “domestico”. Quello che succede è che queste features chiamano tutte le altre che riguardano il cane, ed in fine compare proprio la word-form “cane”.



- La seconda simulazione riguarda la super-categoria “animale”. fornita in ingresso la word-form “animale”, vengono chiamate le features condivise salienti alla super-categoria: “mangia” e “dorme”.



Se viene data soltanto la feature “dorme”, questa chiama la seconda feature condivisa saliente “mangia” ed in fine è evocata la word-form “animale”.



## 5.4 Discussione e conclusione

Il presente lavoro ha lo scopo di presentare un modello di rete neurale, che possa acquisire automaticamente il significato di oggetti concreti e categorie, che sfrutti la statistica co-occorrenza delle features in modo da collegarle a forme verbali. Questo permette di analizzare il ruolo dell'apprendimento Hebbiano, nelle reti associative, con particolare attenzione alle soglie pre- e post-sinaptica, e rivelare il ruolo delle features condivise vs distintive e salienti vs marginali, per sfruttare al meglio la somiglianza tra gli oggetti.

La differenza tra le features comuni e distintive era già presente nei precedenti lavori del Prof. Ursino (Ursino et al, 2009, Ursino et al, 2010). Mentre nella versione trattata in questa tesi e di conseguenza nel modello semplificato è stata introdotta una novità rilevante: nella formazione della rete abbiamo ritenuto che le features possano verificarsi con una frequenza differente, alcune si presentano alla rete più spesso di altre. Così è stato valutato il ruolo della salienza delle features, nel riconoscimento degli oggetti, e nel tempo necessario per tale riconoscimento.

L'algoritmo del modello è stato testato con una tassonomia costituita da sette oggetti, aventi numero di proprietà



differente, e ha prodotto nella maggior parte dei casi i risultati attesi.

In particolare, sono stati rispettati i due aspetti fondamentali di una feature: per quanto concerne la capacità di evocare altre proprietà, le distintive hanno richiamato le altre salienti, distintive e comuni, evitando le non salienti; le comuni hanno evocato le loro simili ma non le distintive, permettendo così la rappresentazione di categorie.

Per ottenere i risultati desiderati sono stati eseguiti diversi tentativi, questi hanno suggerito qualche volta di modificare le percentuali di occorrenza di alcune features, nella rete semantica; mentre in quella lessicale c'è stata qualche modifica della soglia post-sinaptica e del numero di iterazioni. Inoltre, è stato modificato il numero delle simulazioni per poter correggere alcune situazioni indesiderate.

In futuro potrebbe rivelarsi interessante valutare l'attendibilità di un eventuale algoritmo che preveda il setting automatico del valore di soglia, in base alla salienza delle features presentate.

In conclusione si può affermare che il modello manifesti un comportamento adeguato alle ipotesi iniziali e sia capace di rappresentare gli oggetti, come anche le relative categorie, in modo consono a quanto previsto per una memoria semantica.

## **RINGRAZIAMENTI**

Desidero innanzitutto ringraziare il prof. Mauro Ursino, per aver dedicato parte del suo tempo prezioso a migliorare la mia tesi.

Si è sempre dimostrato disponibile al dialogo ed al confronto, qualità che ho molto apprezzato; non si è mai sottratto ai suoi doveri di relatore e come tale, ha sempre fugato i miei dubbi inerenti all'argomento trattato. Spero davvero che il mio elaborato potrà essergli utile per i suoi futuri lavori di ricerca.

Un ringraziamento speciale va sicuramente alla mia famiglia, in particolare ai miei cari genitori, mio marito e mio figlio "Rayan", che mi hanno sostenuta ed incoraggiata lungo tutto il percorso universitario. Senza il loro appoggio e le loro dolci parole di conforto, non ce l'avrei mai fatta: sono e resteranno sempre il mio più grande punto di riferimento.

Per questo, con immenso amore e profonda ammirazione, dico loro: grazie di tutto!

Intendo inoltre ringraziare tutti gli amici ed i compagni di corso che mi sono stati accanto e non hanno esitato a tendermi la mano nei momenti di bisogno: Giuseppe, Conny, Simon, Mario, Marco, Simona, Annabel, Najwa.

## BIBLIOGRAFIA

Ursino Mauro, Cristiano Cuppini and Elisa Magosso, Department of Electronics, Computer Science and Systems, University of Bologna, Bologna, Italy (2011) *A computational model of the lexical-semantic system based on a grounded cognition approach.*

Ursino Mauro, Cristiano Cuppini, Elisa Magosso, University of Bologna, Italy (2012) *The formation of categories and the representation of feature saliency: analysis with a computational model trained with an Hebbian paradigm.*

Barsalou Lawrence W., W. Kyle Simmons, Aron K. Barbey and Christine D. Wilson (2003). *Grounding conceptual knowledge in modality-specific systems.* Lawrence W. Barsalou, Department of Psychology, Emory University, Atlanta, Georgia (2008) *Grounded Cognition.*

Barsalou Lawrence W., Department of Psychology, Emory University, Atlanta (1999), *Perceptual symbol systems.*

Barsalou Lawrence W., Department of Psychology, Emory University, Atlanta, Georgia (2003), *Grounded Cognition.*

Barsalou Lawrence W., Department of Psychology, Emory University, Atlanta, USA (2009), *Simulation, situated conceptualization, and prediction.*

Barsalou Lawrence W. , Department of Psychology, Emory University, Atlanta, Georgia (2008) *Grounding Symbolic Operations in the Brain's Modal Systems.*

Alex Martin and Linda L. Chao (2001) *Semantic memory and the brain: structure and processes.*

Hart John Jr., Raksha Anand, Sandra Zoccoli, Mandy Maguire, Jacque Gamino, Gail Tillman, Richard King, Michael A. Kraut (2007) *Neural substrates of semantic memory.*

Martin Alex, Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, Maryland (2007) *The Representation of Object Concepts in the Brain.*

Collins, A. M., & Loftus, E. F. (1975). *A spreading-activation theory of semantic processing*. *Psychological review*, 82(6), 407.

O'Connor, C. M., Cree, G. S., & McRae, K. (2009). *Conceptual hierarchies in a flat attractor network: Dynamics of learning and computations*. *Cognitive Science*, 33(4), 665-708.

Cree, G. S., McNorgan, C., & McRae, K. (2006). *Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 643.

Taylor Kirsten I., Angeliki Salamoura, Billi Randall, Helen Moss, and Lorraine K. Tyler, University of Cambridge (2006) *Clarifying the Nature of the Distinctiveness by Domain Interaction in Conceptual Structure*.

Vigliocco, G., Vinson, D.P., Lewis, W., & Garrett, M.F. (2004). *Representing the meanings of object and action words: The featural and unitary semantic space hypothesis*. *Cognitive Psychology*, 48, 422-488

Taylor, K.I., Devereux, B.J., Acres, K., Randall, B., & Tyler, L.K. (2012). *Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects*. *Cognition*, 122, 363-374

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). *On the nature and scope of featural representations of word meaning*. *J. Exp. Psychol. Gen.*, 126, 99–130

Anderson, J.R. (1983). *A spreading activation theory of memory*. *J Verb Learn Verb Behav*, 22, 261-295