
ALMA MATER STUDIORUM – UNIVERSITA' DI BOLOGNA
SEDE DI CESENA
SECONDA FACOLTA' DI INGEGNERIA CON SEDE A CESENA
CORSO DI LAUREA MAGISTRALE IN INGEGNERIA BIOMEDICA

Titolo della tesi

**MEMORIA SEMANTICA E LESSICALE: ANALISI
ATTRAVERSO UNA RETE NEURALE.**

Tesi in

Sistemi Neurali LM

Relatore
Prof. Mauro Ursino

Presentata da
Simona Fiori

Sessione II
Anno Accademico 2011/2012

INDICE

Introduzione

1. La memoria semantica

- 1.1 Memoria semantica e cervello: struttura e processi
- 1.2 Supporti neurali della memoria semantica
- 1.3 Grounded ed Embodied Cognition

2. Descrizione matematica del modello

- 2.1 Introduzione al modello
- 2.2 Rete semantica
 - 2.2.1 Descrizione qualitativa
 - 2.2.2 Descrizione matematica
 - 2.2.3 Addestramento delle sinapsi
- 2.3 Rete lessicale
 - 2.3.1 Descrizione qualitativa
 - 2.3.2 Descrizione matematica
 - 2.3.3 Addestramento delle sinapsi

3. Risultati del modello con 2 oggetti

- 3.1 Introduzione
- 3.2 Struttura del modello
- 3.3 Modifiche ai parametri del modello
 - 3.3.1 I modifica
 - 3.3.2 II modifica

4. Risultati del modello con 5 oggetti

4.1 Introduzione

4.2 Struttura del modello

4.3 Risultati

Conclusioni

Ringraziamenti

Bibliografia

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 e riproduce una 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. L'aspetto semantico di un oggetto è individuato da un set di features; queste ultime sono topologicamente organizzate per sviluppare il principio di somiglianza.

Le caratteristiche, appartenenti ad un dato oggetto, possono essere suddivise in: comuni e distintive, salienti e 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 caratteristica 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" è poco frequente, non viene ricordata ma può aiutare nel ricostruire il concetto dell'oggetto stesso. 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 stato di analizzare il ruolo della maggiore o minore salienza delle caratteristiche nel riconoscimento degli oggetti, e gli eventuali cambiamenti nel comportamento della rete, sulla base delle modifiche apportate ai parametri del modello. Quest'ultimo è stato testato usando prima due oggetti, con alcune proprietà comuni ed altre distintive ma diversamente salienti; successivamente è stata implementata una più complessa tassonomia formata da cinque oggetti.

L'addestramento della memoria semantica, mediante potenziamenti e de-potenziamenti, ha permesso di sviluppare un pattern di sinapsi asimmetriche, 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.

Il primo capitolo del lavoro presenta alcune teorie riguardo alla memoria semantica e lessicale, evidenziando l'idea di "Grounded ed "Embodied Cognition". Il successivo espone il modello utilizzato, da un punto di vista qualitativo e matematico. Negli ultimi due, rispettivamente il terzo ed il quarto, sono analizzati i risultati ottenuti dalle simulazioni, svolte sulla tassonomia con due e cinque oggetti.

CAPITOLO 1

LA MEMORIA SEMANTICA

1.1 Memoria semantica e cervello: struttura e processi

Questo lavoro di tesi presenta un modello matematico di una rete neurale che simula l'attività della memoria semantica e lessicale.

Prima di affrontare la sua spiegazione in termini qualitativi e matematici, è opportuno chiarire alcuni concetti necessari alla comprensione di quanto seguirà. 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*].

Un concetto ampiamente accettato è che la memoria semantica, un tipo di memoria dichiarativa (si dice memoria dichiarativa o esplicita tutto ciò che si è immagazzinato ed esprimibile in una qualche forma di linguaggio), sia 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 (ad esempio, la memoria di cosa sia un bicchiere). Questo concetti sono

immagazzinati in diverse modalità senso-motorie ed in più sistemi cognitivi nel cervello (ad esempio, come il bicchiere si tiene in mano, la sua forma, la sua funzione). 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].

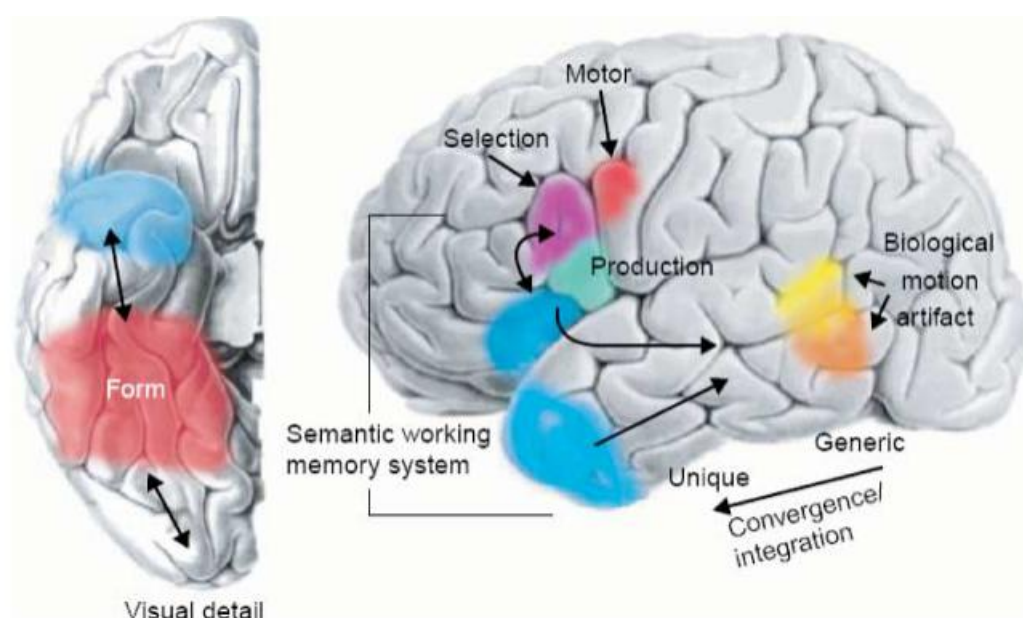


Illustrazione 1: rappresentazione schematica della superficie ventrale (sinistra) e laterale (destra) del cervello. Le aree colorate designano le posizioni approssimative delle regioni in cui possono essere memorizzate le informazioni semantiche sulla forma, movimento e relativi schemi motori dell'oggetto. L'informazione semantica può essere integrata nei lobi temporali, con specificità che aumenta procedendo dall'asse anteriore verso il posteriore. Le regioni specifiche della corteccia parietale inferiore sinistra ed i lobi temporali possono essere impegnati nelle operazioni di recupero, selezione ed accesso alle informazioni semantiche. [Martin, A., & Chao, L.L., 2001].

1.2 Supporti neurali della memoria semantica

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.

Nel complesso, essi sfruttano approcci differenti per illustrare la rappresentazione ed organizzazione della memoria semantica. 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 (ad esempio forma, colore, movimento) 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*].

1.3 Grounded ed Embodied cognition

Le teorie standard sulla conoscenza suppongono che essa sia localizzata in un sistema di memoria semantica separato dagli altri sistemi modali del cervello adibiti alla percezione (ad esempio vista, udito), azione (ad esempio movimento, propriocezione) ed introspezione (ad esempio stati mentali, affetto). Inoltre propongono che le rappresentazioni nei sistemi modali siano trasdotte in simboli amodali che esprimono la conoscenza sull'esperienza nella memoria semantica. La "grounded cognition- la conoscenza basata sulla realtà" - respinge invece le concezioni tradizionali per cui la cognizione sia il calcolo sui simboli amodali in un sistema modulare, indipendente dai sistemi modali del cervello. La suddetta concezione propone l'idea che le simulazioni modali, gli stati del corpo e l'azione, sotto intendano alla conoscenza. Esistono diverse varianti di tale teoria; in generale esse rifiutano la visione classica per la quale i simboli amodali rappresentano la conoscenza nella memoria semantica. È inoltre improbabile che il cervello contenga simboli amodali; se così fosse, essi lavorerebbero assieme alle rappresentazioni modali per creare la conoscenza stessa [Lawrence W. Barsalou, 2008].

La maggioranza dei resoconti riguardanti la "Grounded cognition" si focalizza sui ruoli della simulazione nell'ambito della cognizione [Barsalou 1999, Decety & Grezes 2006]. La simulazione consiste nella rievocazione di stati percettivi, motori ed introspettivi, acquisiti durante l'esperienza con il mondo, il corpo e la mente. A seguito di un episodio (ad esempio riposarsi su una sedia), il cervello cattura gli stati attraverso le modalità e le integra con una rappresentazione multimodale posta in memoria (ad esempio l'azione del sedersi, le introspezioni di confort e relax).

Successivamente, dopo che è stata memorizzata la conoscenza per generare una categoria (ad esempio la sedia), le rappresentazioni

multimodali, catturate durante le esperienze, vengono riattivate con le loro istanze per simulare; in questo modo il cervello ha espresso la percezione, azione, introspezione ad essa associate. Secondo il precedente resoconto, la variegata collezione di meccanismi di simulazione, che mostrano un sistema di rappresentazione comune, supporta lo spettro delle attività cognitive e suggerisce che la simulazione fornisce una forma di computazione nel cervello. L'immaginario mentale costituisce l'esempio migliore di tali meccanismi di simulazione

[*Kosslyn, 1980*]; questo tipicamente risulta dai tentativi volontari di costruire rappresentazioni conscie della memoria di lavoro.

Spesso sembra che si attivino automaticamente ed inconsciamente altre forme di simulazione al di fuori della memoria di lavoro.

Ulteriori ricerche sulla "Grounded cognition" si focalizzano sull'interazione sociale e sull'ambiente [*Barsalou 2003, Barsalou et al. 2007, Rizzolatti & Craighero 2004, Robbins & Aydede 2007*].

Da questo punto di vista il sistema cognitivo si è evoluto a sostenere l'azione in situazioni differenti, inclusa l'interazione sociale. Quest'ultimi lavori sottolineano le interazioni tra la percezione, l'azione, il corpo, l'ambiente ed altri agenti, durante il raggiungimento degli obiettivi.

È importante notare come l'espressione "Embodied cognition-la conoscenza basata sul corpo"-, sia spesso usata per riferirsi alla precedente collezione di articoli scientifici. Tuttavia essa produce l'assunzione errata per cui tutti i ricercatori considerano gli stati del corpo necessari per la cognizione, focalizzando la loro attenzione nelle ricerche solo su di essi. L'etichetta "Embodied cognition" è stata utilizzata negli ultimi 10-15 anni per riferirsi a teorie differenti che coinvolgono ambiti disciplinari disparati: dall'intelligenza artificiale alla robotica, dalle neuroscienze cognitive alla psicologia cognitiva, dalla filosofia alla linguistica e all'antropologia

cognitiva. Ciò che accomuna “Grounded ed Embodied cognition” è il fatto di voler sottolineare l’importanza del ruolo ricoperto dal corpo nella cognizione; l’idea di base è che i nostri processi cognitivi siano vincolati non solo al nostro cervello ma anche al nostro corpo, in particolare al nostro sistema sensoriale e motorio. A seconda dei diversi ambiti ed autori, questa posizione può assumere livelli più o meno radicali [*Goldman e de Vignemont, 2009*]. Alcuni autori enfatizzano l’importanza del corpo e la centralità nei riguardi dell’azione guidata dagli obiettivi [*Glenberg, 1997*], altri invece tendono a preferire l’espressione “Grounded cognition”, evidenziando come la cognizione non sia situata solo nel corpo [*Barsalou, 2008*]. Nonostante i punti in comune, le due prospettive intendono diversamente il modo di rappresentare gli oggetti. In entrambi i casi si ritiene che, quando ad esempio pensiamo ad un telefono, o anche quando sentiamo e produciamo la parola “telefono”, riattiviamo le esperienze che abbiamo svolto con quell’entità o quell’oggetto. Attraverso queste esperienze multimodali, riattiveremo (e il nostro cervello risponderà di conseguenza) l’esperienza acustica di sentire lo squillo del telefono, quella motoria di sollevare la cornetta, quella visiva di distinguerne il colore e così via. Per la posizione “Embodied”, nel pensare al telefono, attiviamo immediatamente le parti di esso legate alle azioni che più frequentemente svolgiamo: ad esempio, la cornetta, che di solito solleviamo. Secondo la versione “Grounded” abbiamo invece una rappresentazione del telefono più globale, non direttamente legata all’azione. Solo in seguito, in funzione dell’azione specifica che dobbiamo svolgere, attiveremo anche le sue caratteristiche più rilevanti. Questi due modi in cui si rappresentano gli oggetti, non sono necessariamente in contrasto ed è probabile che siano entrambi corretti.

Chiaramente la conoscenza spesso è indipendente dal corpo; per cui molti ricercatori si indirizzano verso altre forme di “Grounded cognition”, la quale riflette l’assunto per cui la conoscenza è basata sul corpo in modi diversi, comprese le simulazioni e gli stati dello stesso, testimoniato anche dai risultati di esperimenti condotti negli ultimi dieci anni. I ricercatori che un tempo negavano il fatto che le modalità avessero a che fare con la cognizione, ora riconoscono la loro potenziale rilevanza. È ormai conclamata l’evidenza empirica per la quale le modalità hanno a che fare con la cognizione.

Nonostante ciò, molti dei ricercatori in psicologia cognitiva e scienze cognitive, non sono ancora pronti per abbandonare completamente le teorie tradizionali. Costituirà un aspetto di fondamentale importanza per le ricerca futura valutare la visione mista della situazione.

A questo punto, dopo questa introduzione, nel capitolo successivo analizzerò dettagliatamente il modello neurale sfruttato in questo lavoro di tesi, sia da un punto di vista qualitativo che soprattutto matematico.

CAPITOLO 2

DESCRIZIONE MATEMATICA DEL MODELLO

2.1 Introduzione al modello

Una questione importante nei modelli di memoria semantica è la formazione di categorie ed il diverso ruolo svolto dalle proprietà distintive vs condivise, salienti vs marginali/non salienti.

Il modello usato in questo elaborato è stato sviluppato nel lavoro di ricerca condotto dal Prof. Ursino et al nel 2012, evoluzione dei precedenti: 2009, 2010, 2011. Lo scopo è stato quello di indagare i meccanismi che portano alla formazione di categorie e capire in che modo le caratteristiche salienti possano essere apprese attraverso l'esperienza passata [*Mauro Ursino, 2012*].

Il modello si basa sulla rappresentazione di oggetti intesi come una collezione di features sparse nelle varie aree corticali.

Per l'addestramento delle sinapsi, cioè i collegamenti tra i neuroni, si sfrutta una regola di apprendimento senza supervisore nota come "Regola di Hebb"; è previsto l'uso del potenziamento e della depressione delle sinapsi, di un valore soglia per l'attività pre-sinaptica ed uno per la post-sinaptica, al fine di trovare le norme compatibili con una memorizzazione auto-organizzante robusta delle categorie. Il modello è stato addestrato usando come ingressi degli oggetti semplici, descritti attraverso vettori di features, aventi alcune caratteristiche comuni (che appartengono a più di un concetto e rappresentano una categoria) ed alcune distintive (che appartengono ad un unico concetto e consentono la rappresentazione dei membri) con una differente "saliienza".

Proprio il lavoro del prof. Ursino del 2012 è centrato sul concetto, appena nominato, di "saliienza": le proprietà degli oggetti sono divise in salienti e non salienti. La saliienza è simulata assumendo una diversa frequenza per le proprietà durante l'addestramento.

Le caratteristiche salienti sono molto evidenti e svolgono un ruolo importante nella rappresentazione di un concetto

[*Kirsten I. Taylor, 2006*].

Esse vengono subito in mente quando si pensa ad un oggetto, aiutando la sua ricostruzione in memoria. Invece le proprietà non salienti possono cooperare nella ricostruzione ma non vengono alla mente nel momento in cui si pensa ad un oggetto. 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. 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. Nei paragrafi successivi discuterò del modello attraverso una descrizione matematica e qualitativa delle parti che lo compongono. Prima di proseguire è opportuno fare un'ulteriore precisazione: i capitoli dedicati ai risultati saranno due poiché lo stesso modello matematico è stato usato sia per la semplice tassonomia di 2 oggetti, analizzando in modo particolare il ruolo delle soglie nell'addestramento semantico, sia per il caso più complesso costituito da 5 elementi. Nel primo caso, ciascuno dei due oggetti è caratterizzato da 7 features, di cui 3 comuni e 4 distintive; nel secondo caso, ognuno dei cinque oggetti possiede un numero di caratteristiche diverso l'uno dall'altro. Gli oggetti costituiscono una

tassonomia più complessa, comprendente due categorie, incluse in una superiore chiamata: “super-categoria”.

2.2 Rete semantica

2.2.1 Descrizione qualitativa

Come si evince dall'illustrazione 2, il modello incorpora due reti di neuroni: una rete di caratteristiche 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.

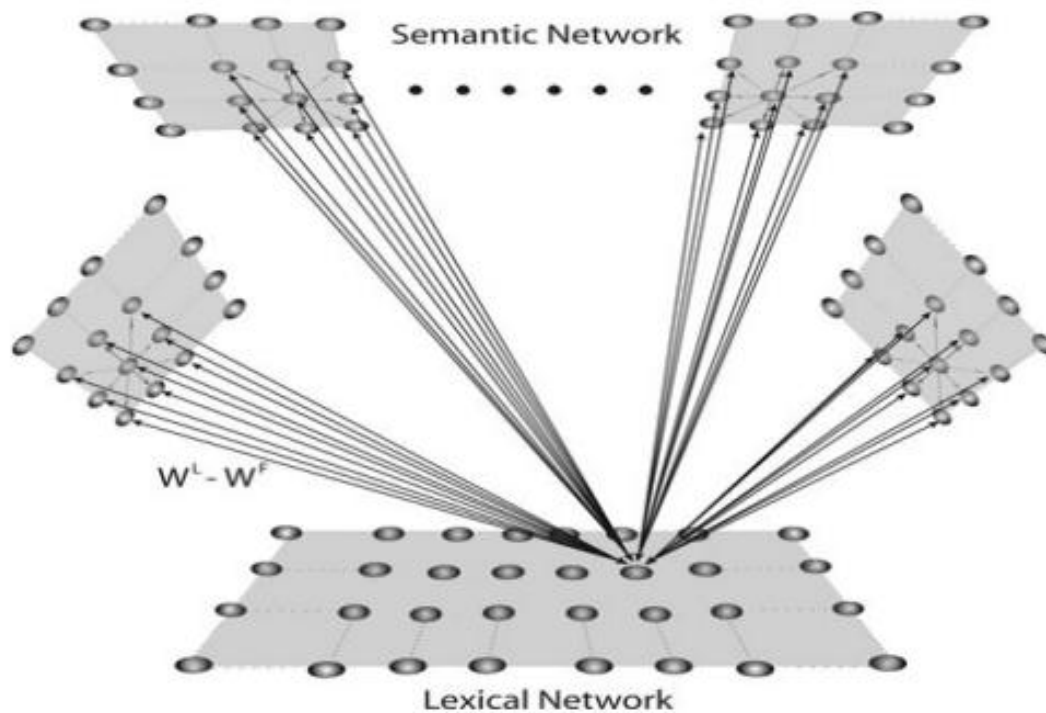


Illustrazione 2: Struttura generale del modello che presenta una rete “semantica” ed una “lessicale”. La prima è formata da 9 aree distinte aventi ciascuna 20x20 neuroni. I neuroni di una certa area non possono comunicare con altri della stessa area, ma possono comunicare con neuroni di aree diverse 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 caratteristica 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 sia 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 caratteristiche simili, situate in posizioni prossimali della rete, diventano moderatamente attive insieme. Una caratteristica può anche essere attivata da un'unità della rete lessicale (sarà mostrato nel dettaglio nei successivi capitoli) e può ricevere sinapsi da altre features presenti in diversi settori, realizzando una memoria auto-associativa (illustrazione 3).

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 delle 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*].

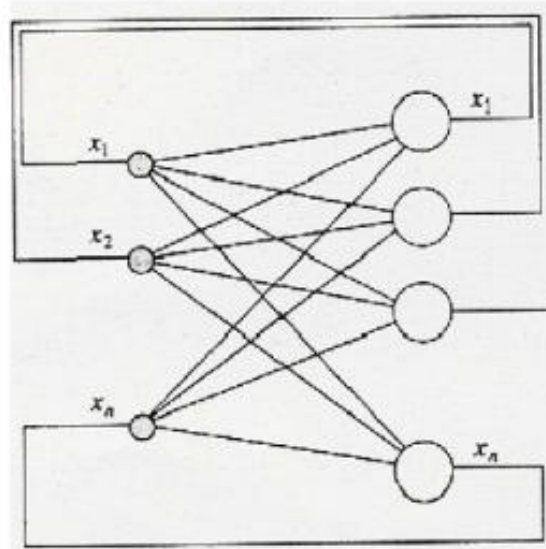


Illustrazione 3: rete auto-associativa. Sono presenti due strati di neuroni ma non bisogna lasciarsi ingannare in quanto i pattern in ingresso ed in uscita sono gli stessi; lo strato a monte serve solo per capire che il valore in uscita è retro-azionato su tutta la rete.

2.2.2 Descrizione matematica

A livello matematico l'attività del neurone viene descritta da un'equazione differenziale di primo grado nel tempo (risolta tramite il metodo di Eulero):

$$\tau \frac{dY}{dt} = -Y + H(\Psi) \quad \text{dove,}$$

$H(\psi)$: è una funzione di attivazione sigmoideale che modula gli stimoli in ingresso al neurone e vale:

$$H(\Psi) = \frac{1}{1 + e^{\frac{-\Psi}{T}}}$$

Gli stimoli in ingresso sono a loro volta rappresentati da ψ :

$$\Psi = I + Se - Si - \Phi x \quad \text{dove,}$$

Φx : è un termine di offset per la sigmoide;

T: è un parametro legato alla pendenza della sigmoide;

Si: rappresenta la somma delle attività provenienti dalle sinapsi laterali inibitorie e vale:

$$Si_{i,j} = \sum_{i,j} (Lin_{i,j} * x_k)$$

Se: è la somma delle attività provenienti dalle sinapsi laterali eccitatorie, comprese quelle provenienti dall'area lessicale e vale:

$$Se_{i,j} = \sum_{i,j} (Lex_{i,j} * x_k) + \sum_{i,j} (W_{i,j} * x_k) + SCL_{i,j}$$

SLC: è la somma delle attività sinaptiche verso il neurone in questione provenienti dalla rete semantica e vale:

$$SCL_{i,j} = \sum_{h,k} (WCL_{i,j} * Xl_k)$$

I: costituisce l'ingresso esterno cioè il richiamo delle proprietà.

Esso è assunto come il risultato di un processo proveniente dal canale senso-motorio. Nell'ambito delle simulazioni, tale stimolo

viene settato dall'operatore e consente di decidere quali proprietà evocare dall'esterno, per simulare un processo di riconoscimento dell'oggetto. L'andamento dell'equazione differenziale di primo grado rappresenta l'attività del neurone, la quale giunge a regime qualora gli stimoli forniti in ingresso, dopo essere stati modulati dalla sigmoide, raggiungano un valore d'attivazione sufficientemente elevato.

2.2.3 Addestramento delle sinapsi

Nel seguito indicherò con il pedice ij la posizione di un neurone nella rete semantica; in particolare il simbolo “ X_{ij} ” sarà usato per esprimere l'attività di un neurone in posizione ij ; tale attività può variare tra 0, inibizione totale ed 1, attivazione massima. Il simbolo “ $W_{ij,hk}$ ” rappresenterà il peso di una sinapsi che collega un neurone

pre-sinaptico in posizione hk ad uno post-sinaptico in posizione ij .

L'addestramento mediante la regola di Hebb, presuppone che il cambiamento nella sinapsi sia dipendente dalla correlazione tra le attività pre e post-sinaptiche. Tuttavia, per poter tenere in considerazione non solo il potenziamento ma anche la depressione a lungo termine, molti lavori di ricerca [*Goldman Mark S., 2001*] confermano la necessità di paragonare tali attività con una soglia. In questo modo, se il livello di attività di uno dei due neuroni fosse basso, produrrebbe indebolimento; viceversa, se entrambi fossero alti, si avrebbe rinforzo.

La regola è la seguente:

$$\Delta W_{ij,hk} = \gamma_{ij,hk} (X_{ij} - \mathcal{I}_{post})(X_{hk} - \mathcal{I}_{pre}) \quad \text{dove,}$$

$\Delta W_{ij,hk}$: rappresenta la variazione della forza delle sinapsi, a causa delle attività pre e post-sinaptiche;

θ_{post} e θ_{pre} : sono le soglie per le attività post e pre-sinaptiche;

$\gamma_{ij,hk}$: è un fattore di apprendimento.

Tuttavia la precedente regola necessita di alcuni aggiustamenti per essere del tutto fisiologica. In primo luogo: nel caso entrambe le attività pre e post-sinaptiche avessero un valore basso, non dovrebbe verificarsi alcun cambiamento di peso. Quindi:

$$\text{se } X_{ij} < \mathcal{I}_{post} \text{ e } X_{hk} < \mathcal{I}_{pre} \rightarrow \Delta W_{ij,hk} = 0$$

In secondo luogo: le sinapsi non possono aumentare all'infinito ma devono raggiungere un livello massimo di saturazione. Ciò è stato ottenuto, nel mio caso, riducendo il tasso di apprendimento " γ ", nel momento in cui le sinapsi si avvicinavano al loro massimo chiamato: " W_{max} ". Per cui si ha:

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

γ_0 : è il tasso di apprendimento quando le sinapsi sono a zero.

Analizziamo ora i 2 aspetti fondamentali di una feature e cioè le capacità di evocare altre caratteristiche e di essere evocata. La prima dipende dalle sinapsi uscenti da un neurone. Nel dettaglio sappiamo che una caratteristica distintiva è importante nella ricostruzione di un oggetto e quindi sarà in grado di evocare le altre proprietà salienti, distintive e comuni, ma non quelle non salienti; infatti quest'ultime aiutano a ricostruire l'oggetto ma non sono evocate spontaneamente. Le features comuni evocano le altre

comuni ricostituendo la categoria ma non le distintive. Le non salienti possono evocare le salienti ma con fatica poichè risultano scarsamente importanti. Nel momento in cui una proprietà riesce ad evocarne un'altra, si può dire che la sinapsi verso quest'ultima è forte, altrimenti è debole o nulla. Un secondo aspetto, correlato con il precedente, è la possibilità che una caratteristica venga evocata dalle altre: questo è un problema legato alle sinapsi che entrano in una caratteristica. Tutte le proprietà salienti devono essere rapidamente richiamate dalle altre salienti, poiché vengono subito alla mente nel pensare ad un oggetto; diversamente le non salienti non devono essere evocate spontaneamente, siano esse distintive o meno. Le proprietà comuni sono evocate dalle altre comuni e da tutte le distintive. Questi concetti si possono riassumere con la seguente espressione: ciò che è distintivo richiama ciò che è comune e ciò che è saliente non richiama ciò che non lo è.

Le sinapsi della rete semantica sono dunque asimmetriche e per realizzare tale aspetto è stato necessario usare due valori di soglia diversi: uno per l'attività pre- ed un altro per l'attività post-sinaptica.

Sono state effettuate le simulazioni scegliendo differenti valori di soglia, per valutare la "saliienza" delle caratteristiche, sulla base delle loro percentuali di presenza durante l'addestramento. Come ci si aspettava, i valori di soglia superiori a 0.5, hanno comportato un indebolimento maggiore delle sinapsi rispetto al loro rinforzo; viceversa, i valori inferiori a 0.5, hanno causato un ridotto indebolimento se paragonato al loro rinforzo.

I risultati ottenuti dalle simulazioni svolte sulle due tassonomie considerate, saranno discussi in modo approfondito nei successivi capitoli.

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”.

Ciascun’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 formazione in cui le parole ed i concetti vengono somministrati insieme: di questo tratterò nei paragrafi successivi.

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. In riferimento alle disposizioni considerate in questo elaborato avremo:

- nel caso della tassonomia di due oggetti, una parola:
per l’oggetto 1;
per l’oggetto 2;
per la categoria che li rappresenta;
- nel caso della tassonomia di cinque oggetti, una parola:
per la “supercategoria” o “superconcetto”;
per la categoria 1 o I concetto;
per la categoria 2 o II concetto;

per l'oggetto 1;
 per l'oggetto 2;
 per l'oggetto 3;
 per l'oggetto 4;
 per l'oggetto 5.

Poiché gli oggetti semantici costituiscono degli insiemi di features, ciascuna parola è collegata, attraverso le sinapsi, a tutte le caratteristiche dell'oggetto o della categoria che essa rappresenta. Anche l'attività dei neuroni lessicali è esprimibile mediante la seguente formula:

$$\tau \frac{dY_l}{dt} = -Y_l + H(\Psi L) \quad \text{dove,}$$

$H(\psi_l)$: è una funzione di attivazione sigmoideale che modula gli stimoli in ingresso al neurone ed ha tale espressione:

$$H(\Psi l) = \frac{1}{\left(1 + e^{-\frac{\Psi l}{T}}\right)}$$

Gli stimoli in ingresso sono rappresentati da ψ_l :

$$\Psi l = I_l + SCL - \Phi x_l$$

2.3.3 Addestramento delle sinapsi

Si ipotizza che l'uomo apprenda dapprima la rappresentazione multi modale di un oggetto, in un secondo momento la colleghi alla parola che lo denota e soprattutto che tali processi non siano contemporanei.

Per questo, le parole sono associate alla rappresentazione semantica dell'oggetto mediante un secondo addestramento. Il suddetto allenamento produce due tipi di sinapsi:

- WCL, le sinapsi dalle parole (rete lessicale) alle features (rete semantica);
- WLC, le sinapsi dalle caratteristiche verso le parole.

La rete è caratterizzata da due aspetti fondamentali. Innanzitutto, come sarà evidenziato nei risultati, ciascuna parola verrà evocata dalla corrispondente rappresentazione dell'oggetto, solo se saranno attive

simultaneamente le sue principali caratteristiche. In secondo luogo, la parola non potrà essere richiamata nel caso siano attive features appartenenti ad un altro oggetto. Per risolvere questi eventuali inconvenienti, sono state adottate delle misure preventive.

Prima dell'addestramento, le sinapsi WLC sono state settate ad un livello negativo, cioè tutte le unità semantiche inviano sinapsi inibitorie alle unità lessicali. Esse sono state addestrate mediante la seguente regola di Hebb:

$$\Delta WLC = \gamma * YL * (YC - \Theta LC) \quad \text{dove,}$$

YL: costituisce il neurone lessicale post-sinaptico;

YC: rappresenta il neurone semantico pre-sinaptico.

Inoltre, WLC è composta da una parte inibitoria ed una eccitatoria:

$$WLCex_{i,j} = WLCex_{i,j} + \gamma * Xl_{i,j,k} * (X_k - 0,5)$$

$$WLCinib_{i,j} = WLCinib_{i,j} - \gamma * Xl_{i,j,k} * X_k$$

Quando YC è inattivo, non si registra alcun cambiamento delle sinapsi; mentre se YL risulta attivo, si produce rinforzo quando anche YC è attivo ed indebolimento se YC è inattivo.

C'è un'ulteriore limitazione: durante l'allenamento la somma delle sinapsi eccitatorie non può superare il valore 1; le sinapsi possono subire variazioni nel loro valore, fermo restando che la somma delle sinapsi entranti ad un'unità lessicale non valga più dell'unità, indipendentemente dal numero di features.

Come già espresso in questo stesso paragrafo, la parola viene evocata nel caso siano attive un numero di caratteristiche, la cui somma delle attività sinaptiche sia pari ad 1, cioè quando sono attive contemporaneamente le sue caratteristiche principali. Questo aspetto viene realizzato usando una sigmoide piuttosto ripida con una soglia di attivazione pari ad 1.

Le sinapsi nella direzione opposta WCL non hanno bisogno di condizioni particolari: esse devono essere eccitatorie verso le parole dell'oggetto considerato e nulle verso le altre unità.

Vengono addestrate attraverso paradigma Hebbiano:

$$\Delta WCL = \gamma * YL * (YC - \Theta_{WCL})$$

La formula risulta praticamente identica alla precedente; sono solo invertite le posizioni pre e post-sinaptiche.

Nei prossimi capitoli saranno mostrati i risultati relativi alle modalità con cui il modello risponde agli stimoli esterni.

CAPITOLO 3

RISULTATI DEL MODELLO CON 2 OGGETTI

3.1 Introduzione

È ormai accertato che se una caratteristica è considerata “saliente” viene percepita spesso durante il processo di memorizzazione di un oggetto.

Nel modello presentato dal prof. Ursino nel lavoro di ricerca del 2011, una feature veniva considerata saliente nel caso apparisse sempre durante l’allenamento; invece veniva data come non saliente nel caso la si registrasse nel 50% delle volte. Tale divisione netta, 100% saliente e 50% non saliente, è risultata piuttosto semplificativa. Al fine di capire come si modificava il comportamento della rete, sono state introdotte alcune modifiche alle percentuali di presenza degli attributi durante l’allenamento. Nello specifico, per ottenere risultati sempre migliori e più vicini alle mie aspettative, ho apportato delle modifiche in corso d’opera a due valori importanti: la soglia post-sinaptica della rete semantica ed il numero delle prove da eseguire per un efficiente addestramento del modello. Nei successivi paragrafi questi ultimi aspetti saranno trattati nel dettaglio.

3.2 Struttura del modello

L’illustrazione 4 mostra lo schema dei 2 oggetti usati per allenare la rete. Ad un livello superiore si pone la categoria con le 3 caratteristiche comuni ad entrambi gli oggetti, ognuno caratterizzato da 4 features distintive: il totale è di 7 proprietà per ciascuno. Accanto ad ogni attributo è riportata la percentuale di presenza durante il training delle sinapsi.

Una delle proprietà, comuni ad ambedue gli oggetti, è presente nel 50% dei casi per cui è da considerarsi non saliente. L'oggetto 1 ha due caratteristiche distintive con il 100% della presenza, una al 70% e l'altra al 50%; l'oggetto 2 ha solo una feature, presente il 100% delle volte mentre le altre sono caratterizzate dai seguenti valori di presenza: 90%, 80%, 60%.

Sulla base delle modifiche apportate ai parametri sopra nominati quali: valore di soglia post-sinaptica della rete semantica e numero di prove da eseguire durante l'addestramento, le caratteristiche con percentuali diverse dal 100% saranno considerate più o meno salienti, contribuendo in misura maggiore o minore alla ricostruzione dell'oggetto.

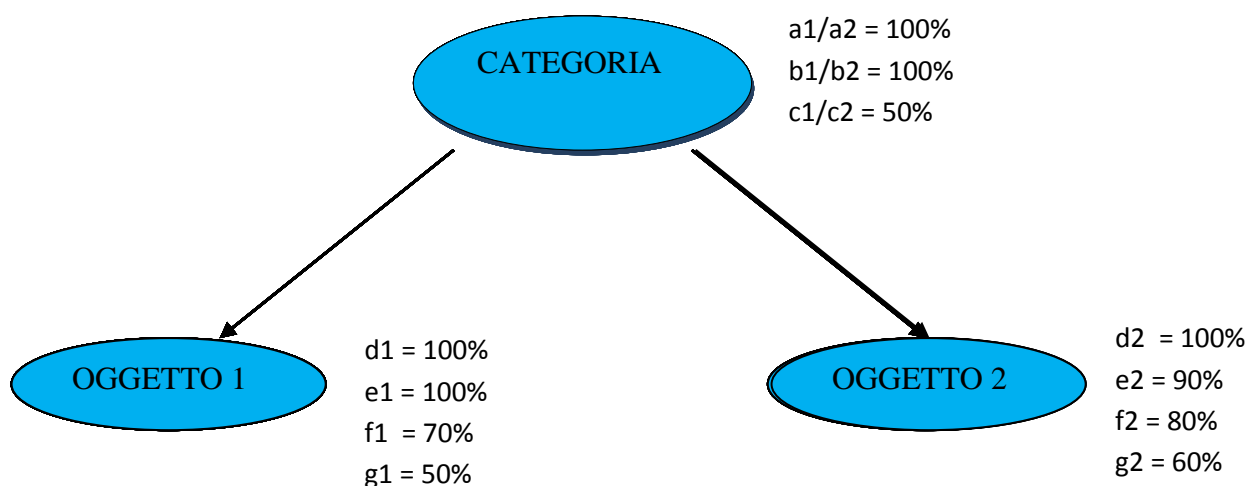


Illustrazione 4: schematizzazione degli oggetti 1 a sinistra e 2 a destra. Ad un livello superiore è presente la categoria con le sue proprietà comuni: una delle tre compare il 50% delle volte. Sotto i due oggetti con quattro caratteristiche distintive ognuno e relative percentuali di presenza durante

l'allenamento: tali valori distinguono le caratteristiche in salienti e non salienti.

Prima di entrare nel dettaglio delle modifiche apportate ai parametri sopra citati e commentare i risultati ottenuti, è necessario fare delle precisazioni.

I neuroni della memoria semantica e lessicale sono rappresentati come punti della rete, caratterizzati da due coordinate, espresse nella seguente tabella.

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7
OGGETTO 1	[10,10]	[10,30]	[10,50]	[30,10]	[50,10]	[30,25]	[50,25]
OGGETTO 2	[10,10]	[10,30]	[10,50]	[30,50]	[50,50]	[30,35]	[50,35]

Tabella 1: posizioni delle sette proprietà appartenenti ai due oggetti utilizzati nel modello. Le prime tre sono comuni, con le stesse coordinate, le altre quattro sono distintive e ricoprono posizioni diverse.

3.3 Modifiche ai parametri del modello

3.3.1 I modifica

Come già espresso sopra, la rete è stata allenata con la regola di Hebb.

Per apprendere correttamente la rappresentazione degli oggetti, è stato necessario somministrarli più volte.

Lo scopo è stato quello di valutare la salienza delle features, sulla base della loro differente percentuale di presenza durante l'addestramento, al variare dei parametri sopra esplicitati.

Ho effettuato 100 prove per il training della memoria semantica e 20 per l'allenamento della memoria lessicale, allo scopo di simulare la fase iniziale dell'addestramento stesso; la rete si trova ancora nella fase di memorizzazione e non si comporterà nel modo più appropriato e desiderato. Inoltre ho assegnato valori ogni volta diversi alla soglia

post-sinaptica e 0.1 alla soglia pre-sinaptica per quanto riguarda la memoria semantica, mantenendo fisso a 0.65 il valore della soglia della memoria lessicale.

In particolare ho eseguito le simulazioni con i seguenti valori: 0.3, 0.4, 0.5, 0.6, 0.65 per la soglia post-sinaptica della rete semantica ottenendo i risultati attesi. Un valore di soglia molto basso come 0.3 o 0.4 ha causato un indebolimento delle sinapsi ridotto rispetto al loro rinforzo. I valori di soglia alti come 0.6 o 0.65 hanno prodotto invece un indebolimento maggiore se confrontato con il relativo rinforzo.

Riporterò solo i risultati ottenuti per l'oggetto 1 con il valore 0.5 (l'oggetto 2 produce gli stessi risultati).

SIMULATION	STIMULATED FEATURES	OUTPUT FEATURES	OBJECT RECOGNITION
1	a1/b1	a1, b1	NO RECOGNITION
2	a1, b1		<u>NO RECOGNITION</u>
3	c1		NO RECOGNITION
4	a1, c1	b1	NO RECOGNITION
5	c1, d1	a1, b1	NO RECOGNITION
6	c1, e1	a1, b1	NO RECOGNITION
7	d1		NO RECOGNITION
8	e1		NO RECOGNITION
9	a1, d1/e1	b1	NO RECOGNITION
10	a1, f1/g1	b1	NO RECOGNITION
11	c1, f1	a1, b1	NO RECOGNITION
12	c1, g1	a1, b1	NO RECOGNITION
13	d1, e1 d1, f1 d1, g1	a1, b1, f1 a1, b1, e1 a1, b1, e1, f1	MEMBER
14	d1, e1, f1 d1, e1, g1	a1, b1 a1, b1, f1	MEMBER
15	f1/g1		NO RECOGNITION
16	f1, g1	a1, b1, d1, e1	MEMBER
17	a1, b1, c1		NO RECOGNITION
18	a1, b1, c1, d1, e1, f1, g1		MEMBER
19	MEMBER		
20	CATEGORY		

Tabella 2: risultati delle simulazioni, ottenuti dall'addestramento del modello (rete semantica e rete lessicale) con il valore di soglia post-sinaptico pari a 0,5.

Le prime 4 simulazioni sono relative alle caratteristiche comuni: le proprietà a1, b1 si richiamano fra loro ma non evocano mai c1, comune non saliente al 50%. Quest'ultima da sola non è in grado di evocare alcunché mentre assieme ad un'altra comune riesce a richiamare l'altra della stessa tipologia. Stesso discorso vale per le caratteristiche distintive f1, g1: singolarmente non evocano alcuna proprietà, assieme ad altre recuperano la parola relativa all'oggetto in esame. Quest'ultima situazione si presenta anche nel caso in cui la rete riceva in ingresso tutte e 7 le features dell'oggetto. Come si può notare dalla tabella, il difetto maggiore, evidenziato con il colore giallo, consiste nell'incapacità di richiamare la categoria corrispondente alle due features comuni: a1, b1, nel momento in cui esse sono somministrate al modello. Un ulteriore aspetto negativo riguarda l'impossibilità di evocare alcuna proprietà dell'oggetto, dando in ingresso la parola relativa ad esso o alla categoria cui esso appartiene.

In figura 5 sono riportate le sinapsi WLC, cioè quelle che partono dalla rete semantica (le caratteristiche) e vanno verso la rete lessicale (le parole).

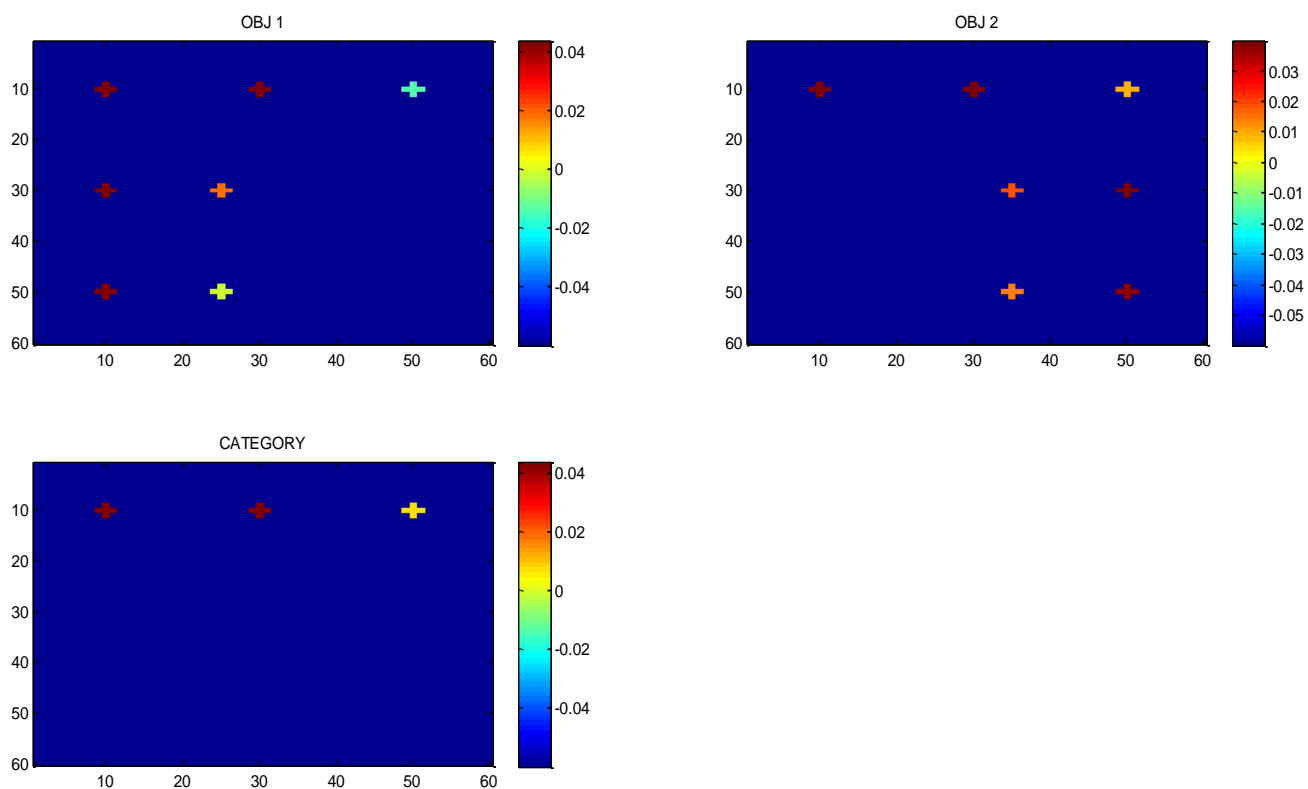


Illustrazione 5: sinapsi dirette verso le parole WLC. In alto a sinistra le sinapsi verso l'oggetto 1; in alto a destra le sinapsi dirette all'oggetto 2 ed in basso a sinistra quelle verso la categoria.

Le sinapsi provenienti dalle 4 caratteristiche salienti, tutte al 100%, hanno un valore alto di circa 0.04; anche le proprietà f1, g1 contribuiscono al raggiungimento dell'unità anche se con valori inferiori mentre c1, non saliente al 50%, ha sinapsi di valore nullo. Diversa è la situazione per l'oggetto 2: le 4 proprietà salienti, di cui una al 90%, hanno valori di circa 0.03. In questo caso le proprietà c2, f2, g2 contribuiscono al raggiungimento dell'unità con valori all'incirca paritari.

È quindi possibile concludere che per ambedue gli oggetti sono necessarie tutte le caratteristiche attive per evocare la parola corrispondente. In ultimo la categoria, nel riquadro in basso a destra, necessita di tutte e 3 le sinapsi per raggiungere il valore 1. Le sinapsi inverse e cioè WCL, da una parola (rete lessicale) verso una caratteristica (rete semantica), sono mostrate in figura 6.

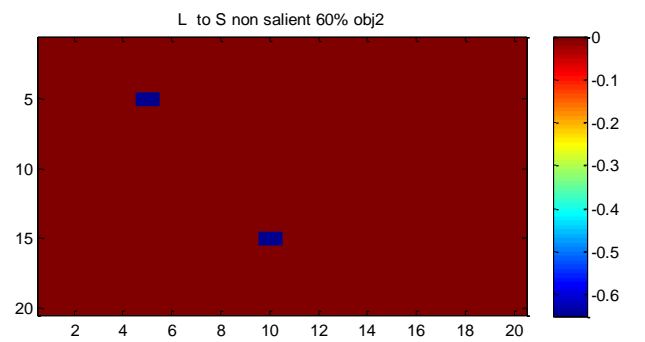
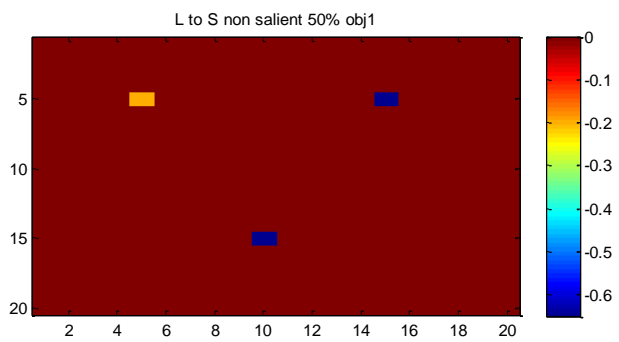
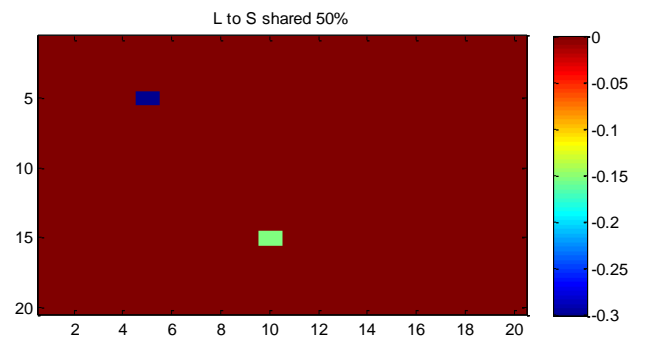
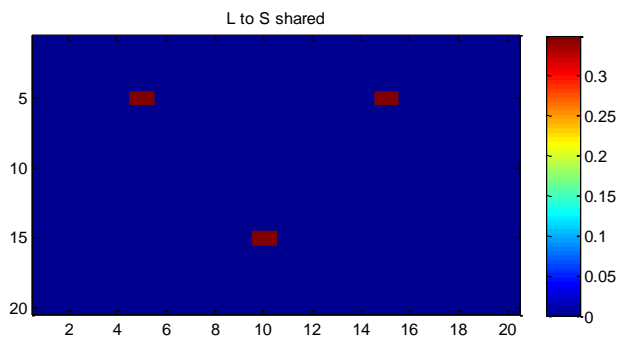


Illustrazione 6: sinapsi entranti alle caratteristiche WCL. In alto a sinistra sono riportate le sinapsi dirette ad una feature comune al 100%; affianco le sinapsi verso c1/c2 comune al 50%; in basso le sinapsi verso le proprietà non salienti al 50% e 60%.

Verso una proprietà comune al 100% giungono sinapsi di valore elevato, in ugual misura dalle 3 parole che designano i 2 oggetti e la categoria cui essi appartengono. La proprietà comune non saliente al 50%, riceve solo sinapsi negative, in accordo con il fatto che una non saliente non viene evocata spontaneamente perché ritenuta poco importante.

Discorso identico per le altre due features rappresentate: g1, la non saliente al 50% dell'oggetto 1 e g2, la non saliente al 60% propria dell'oggetto 2.

3.3.2 II modifica

Il successivo tentativo è stato quello di incrementare il numero di prove per l'addestramento della rete lessicale, portandole da 20 a 60. In questo modo il comportamento inadeguato evidenziato in precedenza, giustificato dalla fase immatura in cui si trovava la rete, in questa circostanza è assente. Come mostrato nella tabella 3, le due sole caratteristiche comuni: a e b, riescono ad evocare la categoria corrispondente. Inoltre ho eseguito le simulazioni sul modello mantenendo costante il valore della soglia della memoria lessicale a 0.5, mentre ho variato di volta in volta la soglia post-sinaptica della rete semantica. I valori analizzati sono stati: 0.5, 0.55, 0.6, 0.65, 0.7. Per completezza ho eseguito anche dei tentativi con valori inferiori a 0.5, di cui non mostrerò i risultati poiché ritenuti non validi: infatti, accadeva che un oggetto era capace di evocare anche le proprietà dell'altro. Come si era previsto l'aumento del valore di soglia post-sinaptica della memoria semantica ha prodotto un indebolimento crescente delle sinapsi corrispondenti. Per brevità di descrizione riporterò e commenterò solo i risultati ottenuti con gli

estremi dell'intervallo di valori considerati: 0.5 e 0.7. Consideriamo prima il caso con valore di soglia 0.5.

La tabella 3 mostra i risultati con il suddetto valore inerenti all'oggetto1.

SIMULATION	STIMULATED FEATURES	OUTPUT FEATURES	OBJECT RECOGNITION
1	a1/b1	a1, b1	CATEGORY
2	a1, b1		<u>CATEGORY</u>
3	c1		NO RECOGNITION
4	a1, c1	b1	CATEGORY
5	c1, d1	a1, b1, e1	MEMBER
6	c1, e1	a1, b1, d1	MEMBER
7	d1		NO RECOGNITION
8	e1		NO RECOGNITION
9	a1, d1/e1	b1	NO RECOGNITION
10	a1, f1/g1	b1	NO RECOGNITION
11	c1, f1	a1, b1	NO RECOGNITION
12	c1, g1	a1, b1	NO RECOGNITION
13	d1, e1 d1, f1 d1, g1	a1, b1, f1 a1, b1, e1 a1, b1, e1, f1	MEMBER
14	d1, e1, f1 d1, e1, g1	a1, b1 a1, b1, f1	MEMBER
15	f1/g1		NO RECOGNITION
16	f1, g1	a1, b1, d1, e1	MEMBER
17	a1, b1, c1		CATEGORY
18	a1, b1, c1, d1, e1, f1, g1		MEMBER
19	MEMBER	a1, b1, d1, e1, f1	
20	CATEGORY	a1, b1	

Tabella 3: risultati delle simulazioni, ottenuti dall'addestramento del modello (rete semantica e rete lessicale) con il valore di soglia post-sinaptico pari a 0.5

È subito evidente che, come espresso precedentemente, questa volta le 2 caratteristiche comuni al 100% sono in grado di richiamare la categoria. Quest'aspetto è presente anche nel caso dell'oggetto 2, i cui risultati sono riportati in tabella 4.

SIMULATION	STIMULATED FEATURES	OUTPUT FEATURES	OBJECT RECOGNITION
1	a2/b2	a2, b2	CATEGORY
2	a2, b2		<u>CATEGORY</u>
3	c2		NO RECOGNITION
4	a2, c2	b2	CATEGORY
5	c2, d2	a2, b2	NO RECOGNITION
6	c2, e2	a2, b2	NO RECOGNITION
7	d2		NO RECOGNITION
8	e2		NO RECOGNITION
9	a2, d2/e2	b2	NO RECOGNITION
10	a2, f2/g2	b2	NO RECOGNITION
11	c2, f2	a2, b2	NO RECOGNITION
12	c2, g2	a2, b2	NO RECOGNITION
13	d2, e2 d2, f2 d2, g2	a2, b2, f2, g2 a2, b2, e2, g2 a2, b2, e2, f2	MEMBER
14	d2, e2, f2 d2, e2, g2	a2, b2, g2 a2, b2, f2	MEMBER
15	f2/g2		NO RECOGNITION
16	f2, g2	a2, b2, d2e2	MEMBER
17	a2, b2, c2		CATEGORY
18	a2, b2, c2, d2, e2, f2, g2		MEMBER
19	MEMBER	a2, b2, d2, e2, f2, g2	
20	CATEGORY	a2, b2	MEMBER

Tabella 4: risultati delle simulazioni, ottenuti dall'addestramento del modello (rete semantica e rete lessicale) con il valore di soglia post-sinaptico pari a 0.5 per l'oggetto 2.

Allo scopo di favorire un confronto rapido ed efficiente tra le due situazioni, i casi di simulazioni per ambo gli oggetti sono identici. In generale le proprietà comuni non salienti al 50%: c1/c2 e le proprietà distintive non salienti rispettivamente al 50% e 60%: g1/g2, aiutano la ricostruzione dell'oggetto ma non vengono evocate spontaneamente insieme ad altre proprietà. Esaminiamo i casi di simulazioni 5 e 6: nel primo, sono date in ingresso alla rete le caratteristiche: "c", comune non saliente al 50% e "d", distintiva saliente al 100%. L'oggetto 1, oltre ad "a" e "b", riesce ad evocare anche "e", saliente al 100%, richiamando così la parola corrispondente; l'oggetto 2 invece evoca solo le proprietà comuni non riuscendo così a richiamare la parola ad esso relativa. Ugualmente nel caso 6, l'oggetto 1, avendo in ingresso "c" ed "e", distintiva saliente al 100%, è in grado di evocare "a", "b", "d" con la relativa parola; diversamente l'altro, richiamando solo le due comuni "a" e "b", non riesce a ricordare la word-form corrispondente.

Le situazioni 13 e 14 dimostrano come l'oggetto 2 abbia bisogno anche della proprietà g2, non saliente al 60%, per evocare la parola corrispondente. Lo stesso si riflette nel caso 19: somministrando la word-form riguardante l'oggetto 2, si richiamano tutte le sue caratteristiche ad eccezione di c2; mentre per l'oggetto 1 sono evocate tutte tranne le due features non salienti al 50%: c1 e g1. In figura 7 sono riportate le sinapsi, WLC, entranti alle parole. L'oggetto 1 riceve sinapsi forti dalle 4 caratteristiche salienti al 100% e sinapsi 4 volte più deboli dalle altre 3: c1, f1, g1. Anche l'oggetto 1 riceve sinapsi forti di valore pari circa a 0.04 dalle 4 salienti, sinapsi deboli da c2 e g2 ma di valore superiore da f2 rispetto ad f1. Infatti queste sono caratterizzate da percentuali di presenza diverse: f2 è all'80% mentre f1 al 70% presente. In ultimo, il riquadro in basso a sinistra riguarda le sinapsi ricevute

dalla categoria: quelle provenienti dalle caratteristiche comuni a, b sono molto forti, invece la non saliente al 50% c invia sinapsi deboli.

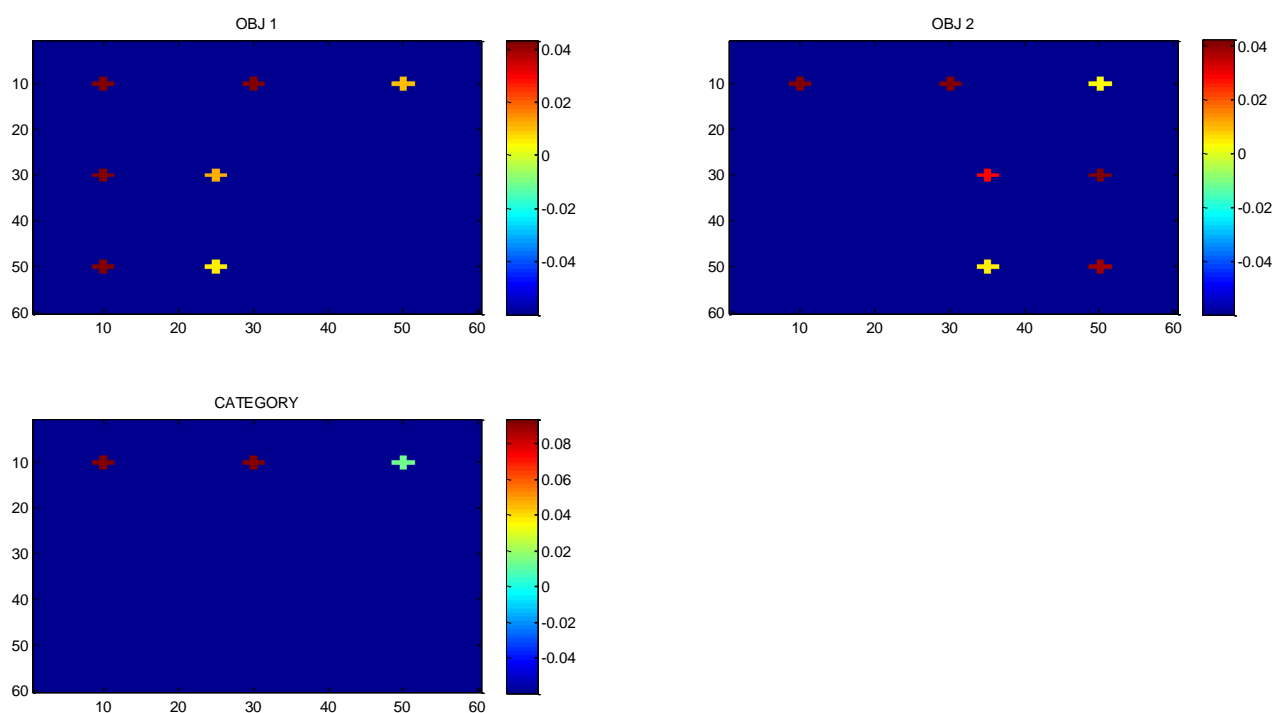


Illustrazione 7: sinapsi entranti alle parole WLC. In alto a sinistra le sinapsi verso l'oggetto 1; in alto a destra le sinapsi dirette all'oggetto 2 ed in basso a sinistra quelle verso la categoria.

La figura 8 successiva mostra le sinapsi, WCL, entranti alle features da parte delle parole. In essa sono proposte diverse caratteristiche sia dell'oggetto 1 che dell'oggetto 2. Una proprietà condivisa riceve sinapsi forti da tutte e 3 le parole. La caratteristica c, non saliente al 50%, in linea con la sua definizione, riceve sinapsi nulle. La feature saliente al 100% dell'oggetto 1 riceve solo dalla parola che lo designa così come la corrispondente saliente al 90% dell'altro oggetto riceve sinapsi positive solo dalla word-form ad esso relativa. Lo stesso comportamento hanno anche le caratteristiche non salienti, una presente l'80% e l'altra il 70% delle volte, che ricevono sinapsi forti solo dalla parola corrispondente all'oggetto considerato e sinapsi nulle dalle altre 2. Riducendosi la percentuale di presenza durante l'allenamento, come nel caso degli ultimi riquadri a destra, si nota come diminuisca il valore delle sinapsi che le caratteristiche non salienti al 50% e 60% ricevono dalle parole corrispondenti.

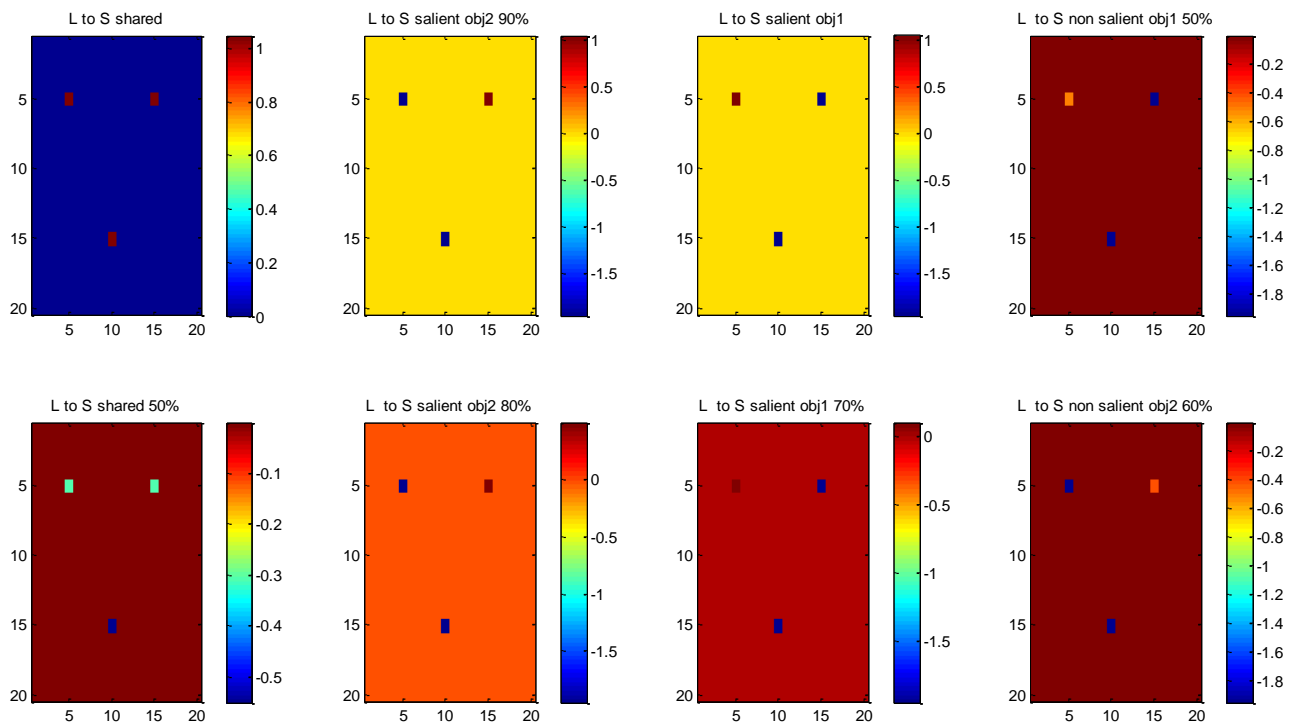


Illustrazione 8: sinapsi entranti alle caratteristiche WCL con valore di soglia pari a 0.5.

Ora passiamo a considerare il caso di addestramento con valore di soglia pari 0.7. L'aspetto di maggiore importanza da sottolineare è il seguente: le proprietà non salienti perdono la capacità di aiutare la ricostruzione dell'oggetto e non vengono più richiamate né somministrando una proprietà né una parola.

Le tabelle successive mostrano i risultati ottenuti dalle simulazioni con i 2 oggetti. Questa volta, le 2 caratteristiche comuni sono sufficienti per richiamare la categoria corrispondente. I casi evidenziati con il colore giallo favoriscono il confronto di situazioni che differiscono tra i due oggetti del modello.

Le simulazioni 11 e 13 sottolineano come dando in ingresso una caratteristica saliente accoppiata ad una non saliente, indipendentemente dalle loro percentuali di presenza, quest'ultima non è capace di fornire aiuto nel recupero della parola corrispondente. Il caso 19 dimostra invece come la parola relativa all'oggetto 2, oltre ad a, b, d, e, richiami anche f2, saliente all'80%.

SIMULATION	STIMULATED FEATURES	OUTPUT FEATURES	OBJECT RECOGNITION
1	a1/b1	a1/b1	NO RECOGNITION
2	a1, b1		CATEGORY
3	c1		NO RECOGNITION
4	a1, c1	b1	CATEGORY
5	c1, d1	a1, b1	NO RECOGNITION
6	c1, e1	a1, b1	NO RECOGNITION
7	d1		NO RECOGNITION
8	e1		NO RECOGNITION
9	a1, d1/e1	b1	NO RECOGNITION
10	a1, f1/g1	b1	NO RECOGNITION
11	c1, f1		NO RECOGNITION
12	c1, g1		NO RECOGNITION
13	d1, e1 d1, f1 d1, g1	a1, b1 a1, b1, d1 a1, b1, e1	MEMBER MEMBER NO RECOGNITION
14	d1, e1, f1 d1, e1, g1	a1, b1 a1, b1	MEMBER MEMBER
15	f1/g1		NO RECOGNITION
16	f1, g1		NO RECOGNITION
17	a1, b1, c1		CATEGORY
18	a1, b1, c1, d1, e1, f1, g1		MEMBER
19	MEMBER	a1, b1, d1, e1	
20	CATEGORY	a1, b1	

Tabella 5: risultati delle simulazioni, ottenuti dall'addestramento del modello (rete semantica e rete lessicale) con il valore di soglia post-sinaptico pari a 0.7 per l'oggetto 1.

SIMULATION	STIMULATED FEATURES	OUTPUT FEATURES	OBJECT RECOGNITION
1	a2/b2	a2/b2	NO RECOGNITION
2	a2, b2		CATEGORY
3	c2		NO RECOGNITION
4	a2, c2	b2	CATEGORY
5	c2, d2	a2, b2	NO RECOGNITION
6	c2, e2	a2, b2	NO RECOGNITION
7	d2		NO RECOGNITION
8	e2		NO RECOGNITION
9	a2, d2/e2	b2	NO RECOGNITION
10	a2, f2/g2	b2	NO RECOGNITION
11	c2, f2	a2, b2	NO RECOGNITION
12	c2, g2		NO RECOGNITION
13	d2, e2 d2, f2 d2, g2	a2, b2 a2, b2	MEMBER NO RECOGNITION NO RECOGNITION
14	d2, e2, f2 d2, e2, g2	a2, b2 a2, b2	MEMBER MEMBER
15	f2/g2		NO RECOGNITION
16	f2, g2		NO RECOGNITION
17	a2, b2, c2		CATEGORY
18	a2, b2, c2, d2, e2, f2, g2		MEMBER
19	MEMBER	a2, b2, d2, e2, f2	
20	CATEGORY	a2, b2	

Tabella 6: risultati delle simulazioni, ottenuti dall'addestramento del modello (rete semantica e rete lessicale) con il valore di soglia post-sinaptico pari a 0.7 per l'oggetto 2.

Non riporterò i grafici inerenti alle sinapsi WCL, dirette alle caratteristiche e WLC, entranti alle parole; infatti sono identici al caso precedente con il valore di soglia pari a 0.5.

Le differenze sostanziali si registrano relativamente alle sinapsi entranti agli oggetti: all'aumentare del valore di soglia, da 0.5 a 0.7, esse vanno gradualmente riducendosi di valore.

A tal proposito, nelle figure successive sono riportate le sinapsi in ingresso all'oggetto 1 con i suddetti valori di soglia, rispettivamente: 0.5 e 0.7.

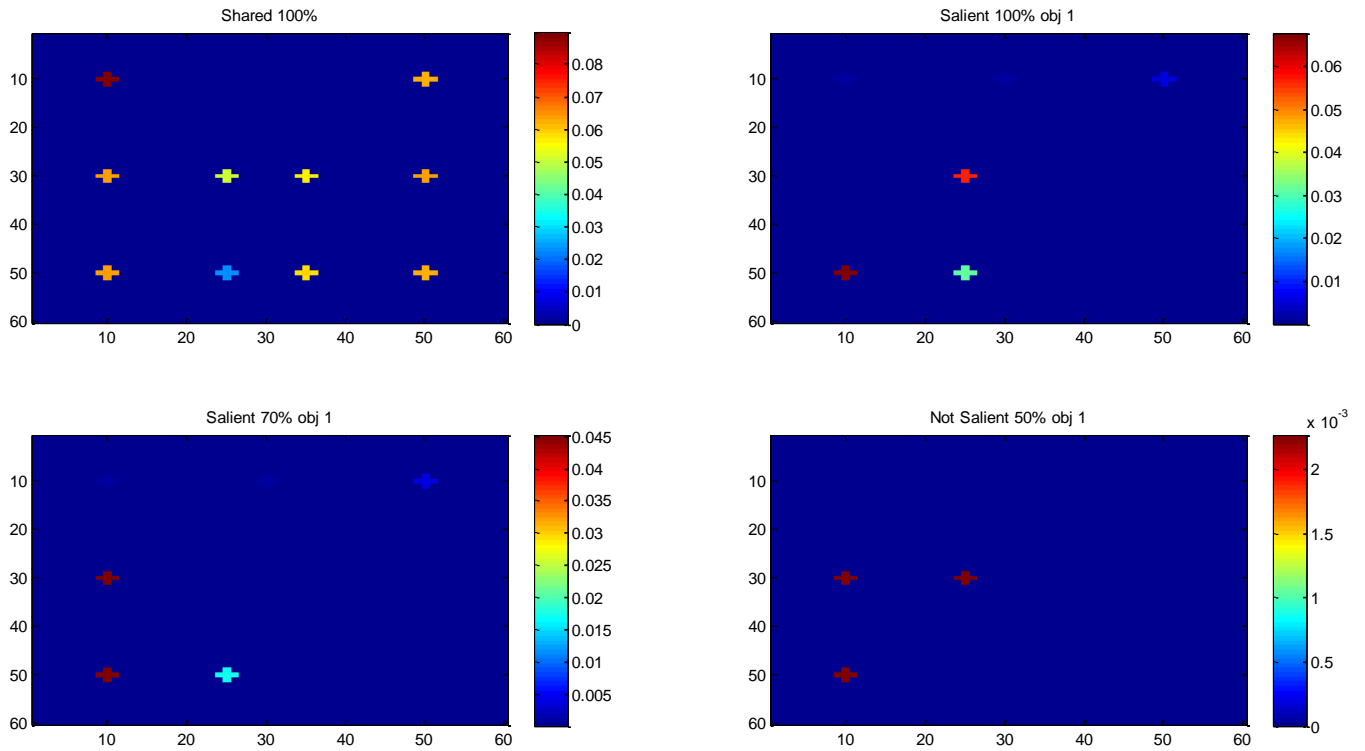


Illustrazione 9: sinapsi in ingresso alle caratteristiche dell'oggetto 1 con valore di soglia pari a 0.5. In alto è destra le sinapsi entranti verso la comune al 100%; di fianco quelle dirette alla saliente 100%; sotto a sinistra le sinapsi verso la proprietà saliente al 70%; in basso a destra quelle direzionate alla non saliente al 50%.

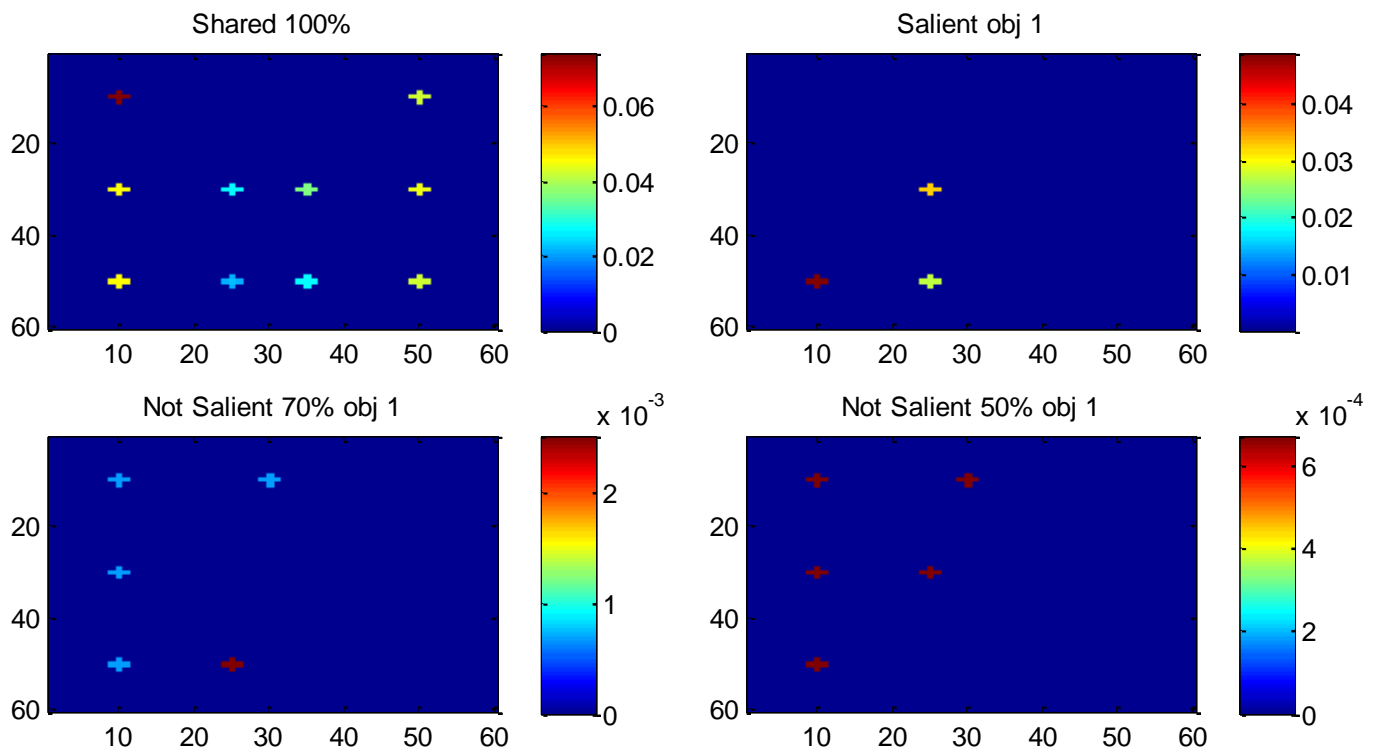


Illustrazione 10: sinapsi in ingresso alle caratteristiche dell'oggetto 1 con valore di soglia pari a 0.7. In alto è destra le sinapsi entranti verso la comune al 100%; di fianco quelle dirette alla saliente al 90%; sotto a sinistra le sinapsi verso la proprietà saliente all' 80%; in basso a destra quelle direzionate alla non saliente al 60%.

Per brevità di descrizione, commenterò solo un paio delle 4 features rappresentate: i casi più significativi riguardano le sinapsi entranti alle proprietà poco salienti. Nel caso di soglia: 0.5, la proprietà con percentuale di presenza del 70%, nel riquadro in basso a sinistra, è da considerarsi saliente: infatti essa riceve sinapsi seppur debolissime, quasi nulle, dalle 3 features comuni e sinapsi di valore superiore: 0.045, dalle 2 salienti al 100%. La caratteristica condivisa non saliente al 50%, posta nel riquadro accanto, riceve sinapsi negative, dell'ordine di 10^{-3} , praticamente da tutte le altre, in accordo con la sua definizione. Nel caso di soglia: 0.7, la feature al 70% diventa non saliente ed è proprio questo l'aspetto più interessante causato dall'aumento di soglia. Infatti essa riceve sinapsi negative, dell'ordine di grandezza di 10^{-3} , dalle altre proprietà cioè non viene per nulla richiamata. La proprietà c1, che ovviamente resta non saliente, in questo caso riceve sinapsi ancor più negative: dell'ordine di grandezza di 10^{-4} . Tutto questo è in linea con il fatto che si registra un indebolimento più evidente con l'aumentare del valore di soglia post-sinaptica.

Come per l'oggetto 1, effettuerò un confronto anche per l'oggetto 2 nel caso dei valori di soglia: 0.5 e 0.7.

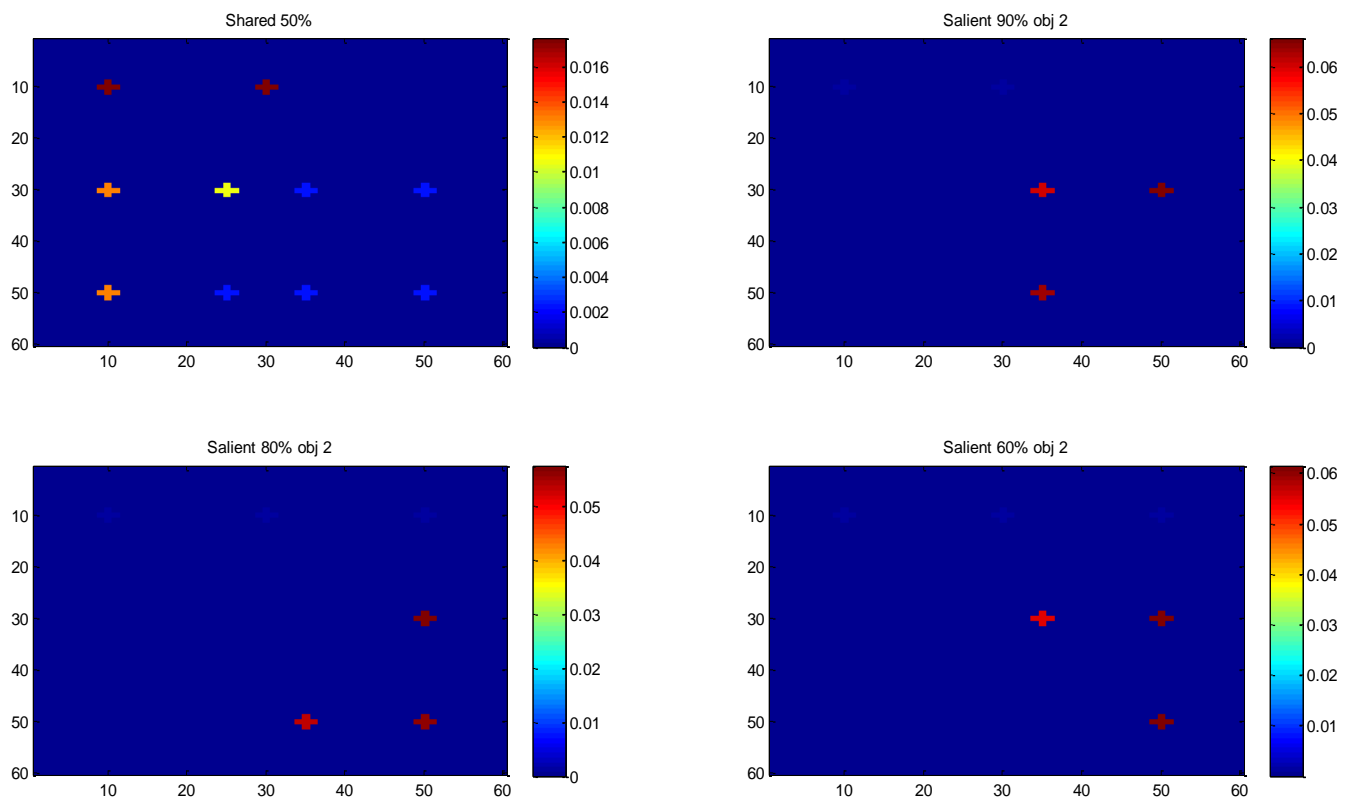


Illustrazione 11: sinapsi in ingresso alle caratteristiche dell'oggetto 2 con valore di soglia pari a 0.5. In alto è destra le sinapsi entranti alla proprietà comune al 50%; di fianco quelle dirette alla saliente al 90%; sotto a sinistra le sinapsi verso la proprietà saliente all' 80%; in basso a destra quelle direzionate alla saliente al 60%.

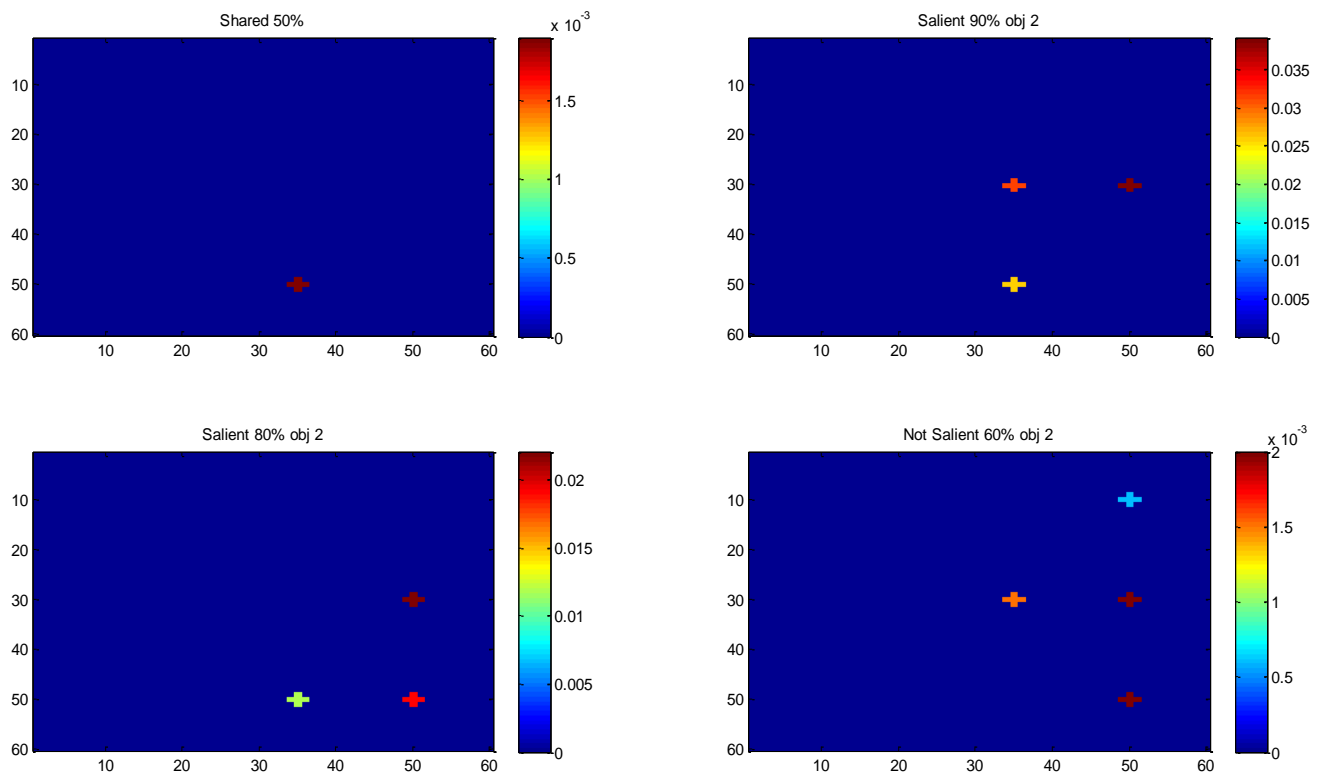


Illustrazione 12: sinapsi in ingresso alle caratteristiche dell'oggetto 2 con valore di soglia pari a 0.7. In alto è destra le sinapsi entranti alla proprietà comune al 50%; di fianco quelle dirette alla saliente al 90%; sotto a sinistra le sinapsi verso la proprietà saliente all' 80%; in basso a destra quelle direzionate alla non saliente al 60%.

Nel caso di soglia post-sinaptica pari a 0.5, la proprietà nel riquadro in basso a sinistra, saliente all'80%, riceve sinapsi deboli dalle altre salienti: circa 0.05 e debolissime, quasi nulle, dalle 3 caratteristiche condivise. Anche la feature accanto, saliente al 60%, riceve sinapsi deboli da d2, e2, f2 pari circa a 0.06 e inferiori, quasi nulle, dalle comuni. L'aumento della soglia a 0.7 provoca una riduzione ulteriore del valore delle sinapsi entranti: in ingresso alla feature all'80%, entrano sinapsi del valore pari circa a 0.02 dalle altre salienti; essa è ancora da considerarsi saliente. Viceversa, la proprietà al 60% diventa non saliente, ricevendo principalmente sinapsi negative, dell'ordine di 10^{-3} , da: c2, d2, e2, f2.

L'istogramma riportato in figura 13 mostra le sinapsi entranti alla caratteristica distintiva f1 dell'oggetto 1, saliente al 70%. La riduzione significativa del loro valore, ha portato a considerare tale proprietà da saliente, nel caso di valore 0.5, a non saliente per il valore 0.7.

Di seguito, per completezza, viene proposto un ulteriore istogramma relativo alla proprietà distintiva g2 appartenente all'oggetto 2: anch'essa, con una percentuale di salienza del 60%, nel passaggio dalla soglia 0.5 a 0.7, si deve trattare come non saliente.

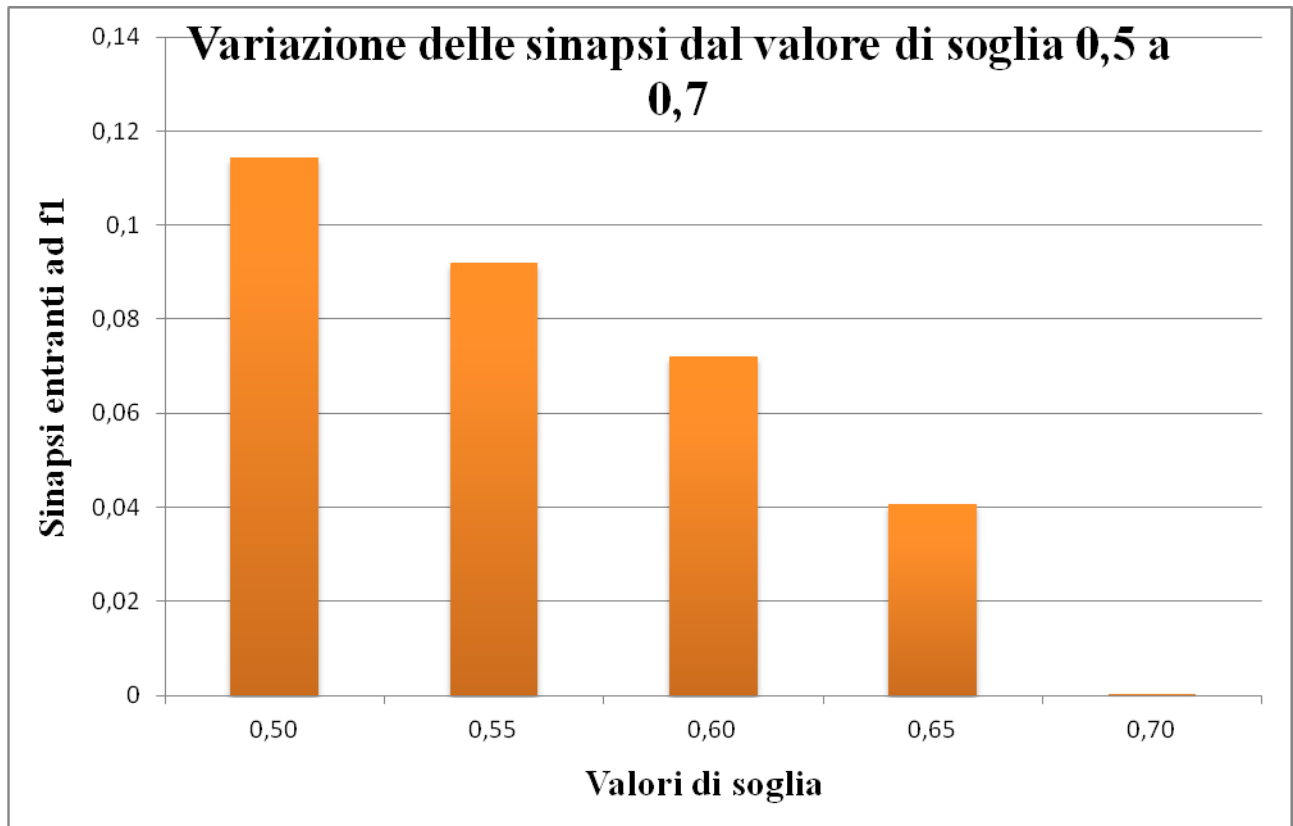


Illustrazione 13: variazione del valore delle sinapsi entranti alla proprietà f1, con percentuale di salienza del 70%, in funzione dell'aumento della soglia post-sinaptica da 0.5 a 0.7.

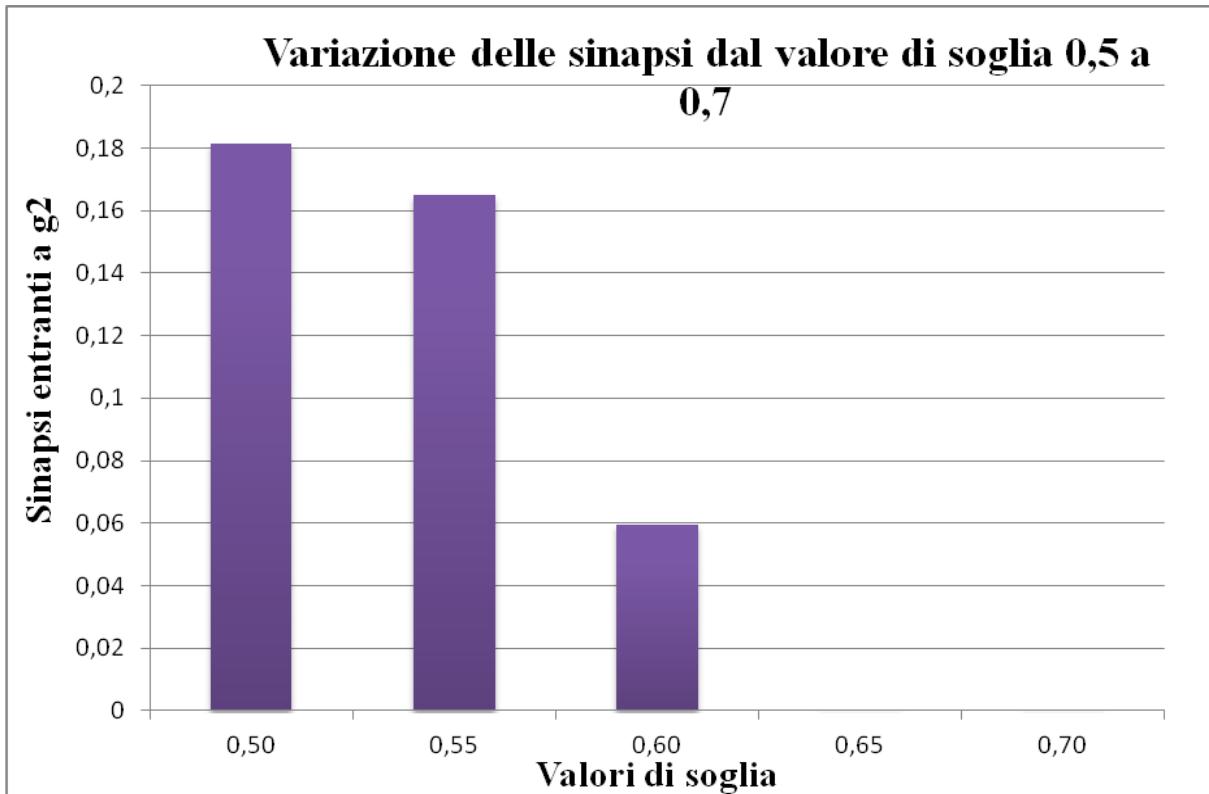


Illustrazione 14: variazione del valore delle sinapsi entranti alla proprietà g2, con percentuale di salienza del 60%, in funzione dell'aumento della soglia post-sinaptica da 0.5 a 0.7.

Il capitolo successivo tratterà i risultati, ottenuti dalle prove sulla tassonomia a 5 oggetti.

CAPITOLO 4

RISULTATI DEL MODELLO CON 5 OGGETTI

4.1 Introduzione

Finora abbiamo commentato i risultati ottenuti dalle simulazioni effettuate sul modello con 2 oggetti, caratterizzati ciascuno dallo stesso numero di proprietà: 7. Per verificare la validità degli strumenti matematici utilizzati e dimostrare la loro estendibilità anche ad altre strutture, il modello è stato utilizzato con un set più compatto di ingressi, costituito da 5 concetti. Questa volta, ognuno di essi è caratterizzato da un numero di features diverso e da differenti percentuali di salienza. Nei successivi paragrafi procederò nel farne un'analisi più approfondita.

4.2 Struttura del modello

Di seguito, l'illustrazione 15 mostra l'architettura degli input. In cima è presente la "super-categoria" o "super-concetto", condivisa da tutti gli oggetti; essa comprende 2 "sotto-categorie" o concetti, con 2 proprietà ciascuna, comuni agli elementi che ne fanno parte. La "categoria 1" è composta da 3 oggetti, la "categoria 2" ha due oggetti.

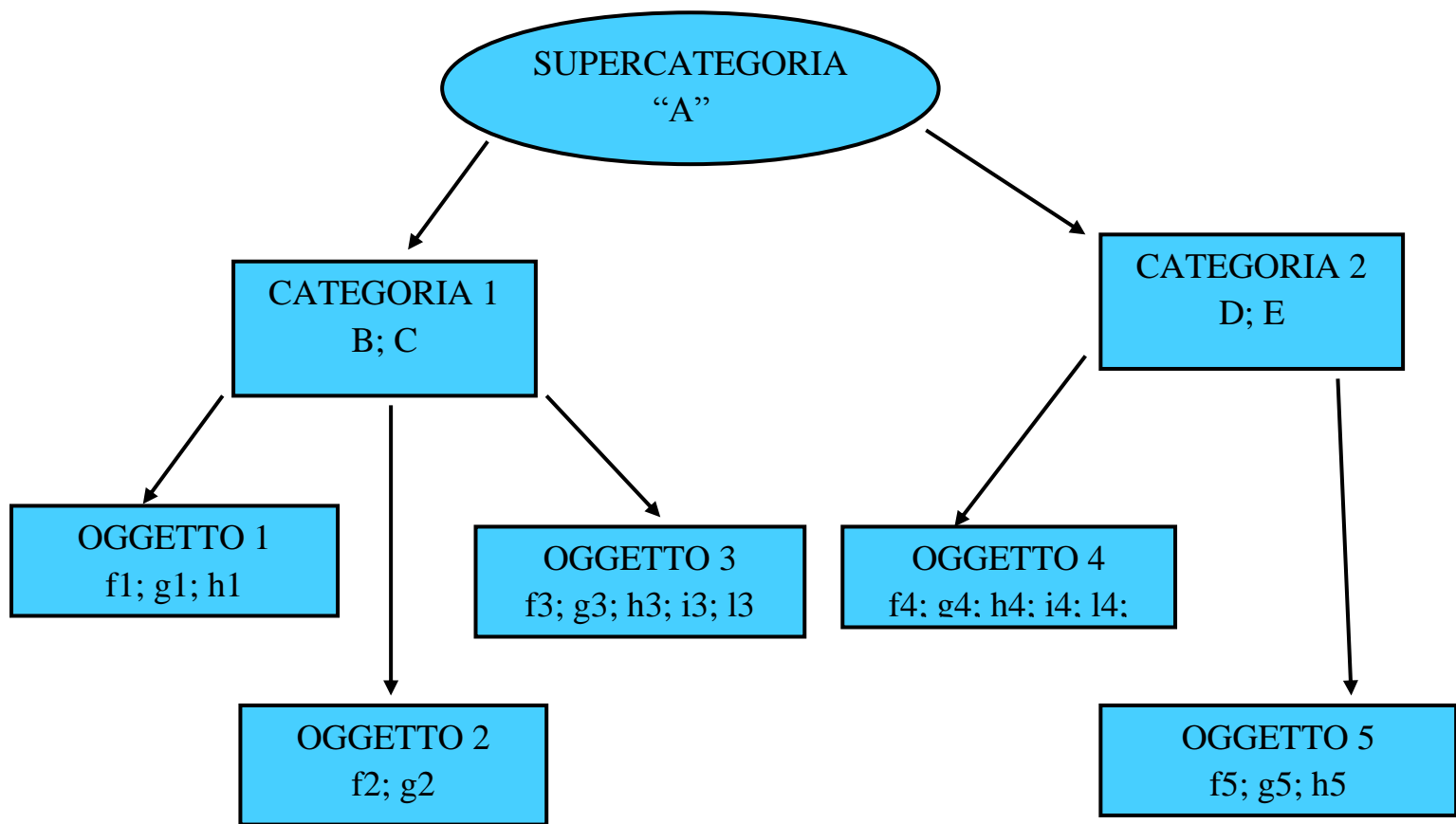


Illustrazione 15: schematizzazione del modello a 5 oggetti.

Nella tabella 7 sono mostrate le percentuali di salienza di ciascuna feature e le loro posizioni nelle varie aree che compongono il modello. In particolare la proprietà:

- A è comune a tutti gli oggetti;
- B; C sono comuni alla categoria 1;
- D; E sono comuni alla categoria 2;
- f1; f2; f3; f4; f5 appartengono alla quarta area;
- g1; g2; g3; g4; g5 appartengono alla quinta area;
- h1; h3; h4; h5 appartengono alla sesta area;
- i3; i4 appartengono alla settima area;
- l3; l4 appartengono all'ottava area;
- m4 appartengono alla nona area.

Proprietà	Percentuali	Posizioni
A	90%	[10 , 10]
B	90%	[10 , 25]
C	70%	[10 , 45]
D	90%	[10 , 35]
E	50%	[10 , 55]
f1	90%	[25 , 5]
f2	90%	[25 , 10]
f3	100%	[25 , 15]
f4	90%	[35 , 5]
f5	80%	[35 , 15]
g1	80%	[25 , 25]
g2	80%	[25 , 30]
g3	80%	[25 , 35]
g4	90%	[35 , 25]
g5	70%	[35 , 35]
h1	60%	[25 , 45]
h3	70%	[25 , 55]
h4	60%	[35 , 45]
h5	50%	[35 , 55]
i3	60%	[50 , 5]
i4	40%	[50 , 15]
l3	40%	[50 , 25]
l4	50%	[50 , 35]
m4	40%	[50 , 50]

Tabella 7: proprietà degli elementi del modello, con posizioni e relative percentuali di presenza durante l'addestramento.

4.3 Risultati

Sulla base delle simulazioni effettuate relativamente alla tassonomia con 2 oggetti, analizzata nel capitolo precedente, si è scelto di implementare la struttura più complessa a 5 oggetti, con i seguenti parametri:

- numero di prove pari a 60 , per l'addestramento della memoria lessicale;
- numero di prove pari a 100, per l'addestramento della memoria semantica;
- valore di soglia pari a 0.5, per la rete lessicale;
- valori pari a 0.5 e 0.1, rispettivamente per la soglia: post- e pre-sinaptica della rete semantica.

Dall'analisi dei grafici ottenuti dopo l'addestramento, relativi alle sinapsi entranti alle parole e agli attributi, sono emerse alcune anomalie. Per questo, di seguito, riporterò solo la tabella inerente alle simulazioni, effettuate con i parametri appena menzionati. Il termine "CASE" (caso) si riferisce all'elemento con cui si vuole lavorare, stimolando di volta in volta le features che lo riguardano. Quindi: il "CASE 1" si riferisce alla "supercategoria"; il "CASE 2" rappresenta il I° CONCETTO; il "CASE 3" riguarda il II° CONCETTO; il "CASE 4" si riferisce all'oggetto 1 e così di seguito per i restanti elementi del modello.

I comportamenti difettosi sono evidenziati con il colore giallo. In particolare: dando in ingresso la proprietà "a", la rete avrebbe dovuto evocare la parola che esprime il SUPERCONCETTO ma così non è stato. Stesso discorso vale nel caso si consideri il I° CONCETTO: somministrando la feature "a", avrebbe dovuto richiamarsi la parola corrispondente. Inoltre, esaminando il II° CONCETTO, la caratteristica "d", con una salienza elevata del

90%, avrebbe dovuto richiamare anche la feature “a”, oltre alla parola corrispondente.

CASE	STIMULATED FEATURES	OUTPUT FEATURES	OBJECT RECOGNITION
1	SUPERCONCETTO	a	
1	a		
2	I° CONCETTO	a; b; c	
2	a		
2	a; c	b	I° CONCETTO
2	b; c	a	I° CONCETTO
2	a; b	c	I° CONCETTO
3	II° CONCETTO	a; d	
3	d		II° CONCETTO
3	a; d		II° CONCETTO
3	a; e		
3	d; e	a	II° CONCETTO
4	I° PAROLA	a; b; c; f1; g1	
4	f1; g1	a; b; c	I° PAROLA
4	b; f1	a; c	
4	b; f1; h1	a; c; g1	I° PAROLA
4	a; f1	b; c	
4	g1; h1	a; b; c; f1	I° PAROLA
5	II° PAROLA	a; b; c; f2; g2	
5	f2; g2	a; b; c	II° PAROLA
5	f2		
6	III° PAROLA	a; b; c; f3; g3; h3 i3	
6	f3; g3	a; b; c; f3; g3; h3 i3	III° PAROLA
6	f3; h3	a; b; c; g3; i3	III° PAROLA
6	f3; i3	a; b; c; g3; h3; i3	III° PAROLA
6	g3; h3	a; b; c; f3; i3	III° PAROLA
6	i3; i3	a; b; c; f3; g3; h3	III° PAROLA
7	IV° PAROLA	a; d; f4; g4	
7	g4; h4	a; d; f4	IV° PAROLA
7	f4; g4	a; d	IV° PAROLA
7	f4; h4	a; d; g4	IV° PAROLA

7	f4; i4	a; d; g4	IV° PAROLA
7	f4; l4	a; d; g4	IV° PAROLA
7	f4; m4	a; d; g4	IV° PAROLA
7	h4; i4		
8	V° PAROLA	a; d; f5; g5	
8	f5; g5	a; d	V° PAROLA
8	f5; h5	a; d	
8	g5; h5	a; d; f5	V° PAROLA

Tabella 8: risultati delle simulazioni, ottenuti dall'addestramento del modello (rete semantica e rete lessicale) con un 60 prove, in riferimento alla memoria semantica.

I comportamenti poco adeguati sono stati attribuiti ad un insufficiente sviluppo delle sinapsi fra l'area semantica e quelle lessicale. Pertanto, si è scelto di aumentare il numero di prove da 60 ad 80, per l'addestramento della memoria lessicale; gli altri parametri sono rimasti invariati.

Di seguito è riportata la tabella relativa alle simulazioni che sono state effettuate con i parametri modificati. Si evince che i comportamenti anomali, riscontrati in precedenza, sono assenti. Infatti: analizzando il caso 1, dando in ingresso la proprietà "a", la rete evoca la parola corrispondente cioè il SUPERCONCETTO. Ugualmente nel secondo caso evidenziato, l'ingresso della caratteristica "a", causa l'evocazione della parola corrispondente: I° CONCETTO. In ultimo, somministrando la proprietà "d", la rete evoca anche la feature "a" oltre alla word-form relativa: II° CONCETTO. Per quello che riguarda le altre situazioni, la rete si comporta in modo adeguato.

Esaminiamo il II° CONCETTO: dando in ingresso la parola corrispondente, la rete richiama le features “a” e “d”, entrambi salienti al 90% ma non “e”, da considerarsi non saliente al 50%. Lo stesso discorso è valido anche per il caso 4: dando in ingresso la parola relativa all’oggetto1, sono evocate le proprietà: a, b, c, f1, g1, tutte considerate salienti, ma non “h1”, non saliente al 60%. Ugualmente accade quando vengono somministrate alla rete le parole relative agli oggetti 2, 4 e 5: le features con una percentuale inferiore o uguale al 60%, non sono evocate perché considerate non salienti. Viceversa, l’oggetto 3, costituito da 8 proprietà, riesce ad evocare come saliente anche la caratteristica i3 al 60%.

CASE	STIMULATED FEATURES	OUTPUT FEATURES	OBJECT RECOGNITION
1	SUPERCONCETTO	a	
1	a		SUPERCONCETTO
2	I° CONCETTO	a; b; c	
2	a	I° CONCETTO	
2	b	a; c	I° CONCETTO
2	a; c	b	I° CONCETTO
2	b; c	a	I° CONCETTO
2	a; b	c	I° CONCETTO
3	II° CONCETTO	a; d	
3	d	a	II° CONCETTO
3	a; d		II° CONCETTO
3	a; e		
3	d; e	a	II° CONCETTO
4	I° PAROLA	a; b; c; f1; g1	
4	f1; g1	a; b; c	I° PAROLA
4	b; f1	a; c	
4	b; f1; h1	a; c; g1	I° PAROLA
4	a; f1	b; c	

4	g1; h1	a; b; c; f1	I° PAROLA
4	f1		
5	II° PAROLA	a; b; c; f2; g2	
5	f2; g2	a; b; c	II° PAROLA
5	f2		
6	III° PAROLA	a; b; c; f3; g3; h3 i3	
6	f3; g3	a; b; c; f3; g3; h3 i3	III° PAROLA
6	f3; h3	a; b; c; g3; i3	III° PAROLA
6	f3; l3	a; b; c; g3; h3; i3	III° PAROLA
6	g3; h3	a; b; c; f3; i3	III° PAROLA
6	f3		
6	i3; l3	a; b; c; f3; g3; h3	III° PAROLA
7	IV° PAROLA	a; d; f4; g4	
7	g4; h4	a; d; f4	IV° PAROLA
7	f4; g4	a; d	IV° PAROLA
7	f4; h4	a; d; g4	IV° PAROLA
7	f4; i4	a; d; g4	IV° PAROLA
7	f4; l4	a; d; g4	IV° PAROLA
7	f4; m4	a; d; g4	IV° PAROLA
7	h4; i4		
8	V° PAROLA	a; d; f5; g5	
8	f5; g5	a; d	V° PAROLA
8	f5; h5	a; d	
8	g5; h5	a; d; f5	V° PAROLA
8	g5		

Tabella 9: risultati delle simulazioni, ottenuti dall'addestramento del modello (rete semantica e rete lessicale) con un 80 prove, relativamente alla memoria semantica.

La figura 16 mostra i valori delle sinapsi, WLC, entranti alle parole. La rete si comporta in modo adeguato. Per brevità di descrizione commenterò solo alcuni dei riquadri presenti. Innanzitutto la supercategoria riceve sinapsi molto forti solo dalla proprietà “A”; la categoria 2 riceve sinapsi elevate da “A”, di valore inferiore da “D” ma di valore nullo dalla caratteristica “E”, non saliente al 50%. L’oggetto 2 riceve in ingresso sinapsi da tutte le sue features, proporzionalmente alla loro percentuale di salienza. L’elemento con il maggior numero di proprietà è l’oggetto 4. Esso ammette in ingresso sinapsi forti dalle features: “A”, “D”, comuni salienti al 90% e dalle distintive salienti al 90%: “f4”, “g4”; invece le restanti caratteristiche, la comune “E” e le distintive: “h4”, saliente al 60%, “l4”, non saliente al 50%, “i4”, “m4”, non salienti al 40%, inviano sinapsi quasi nulle. Diversamente, l’oggetto 5 riceve sinapsi elevate da “A”, “D” ed “f5”, distintiva saliente all’80%, di valore inferiore da “E” e “g5”, distintiva saliente al 70% e quasi nulle da “h5”, non saliente al 50%.

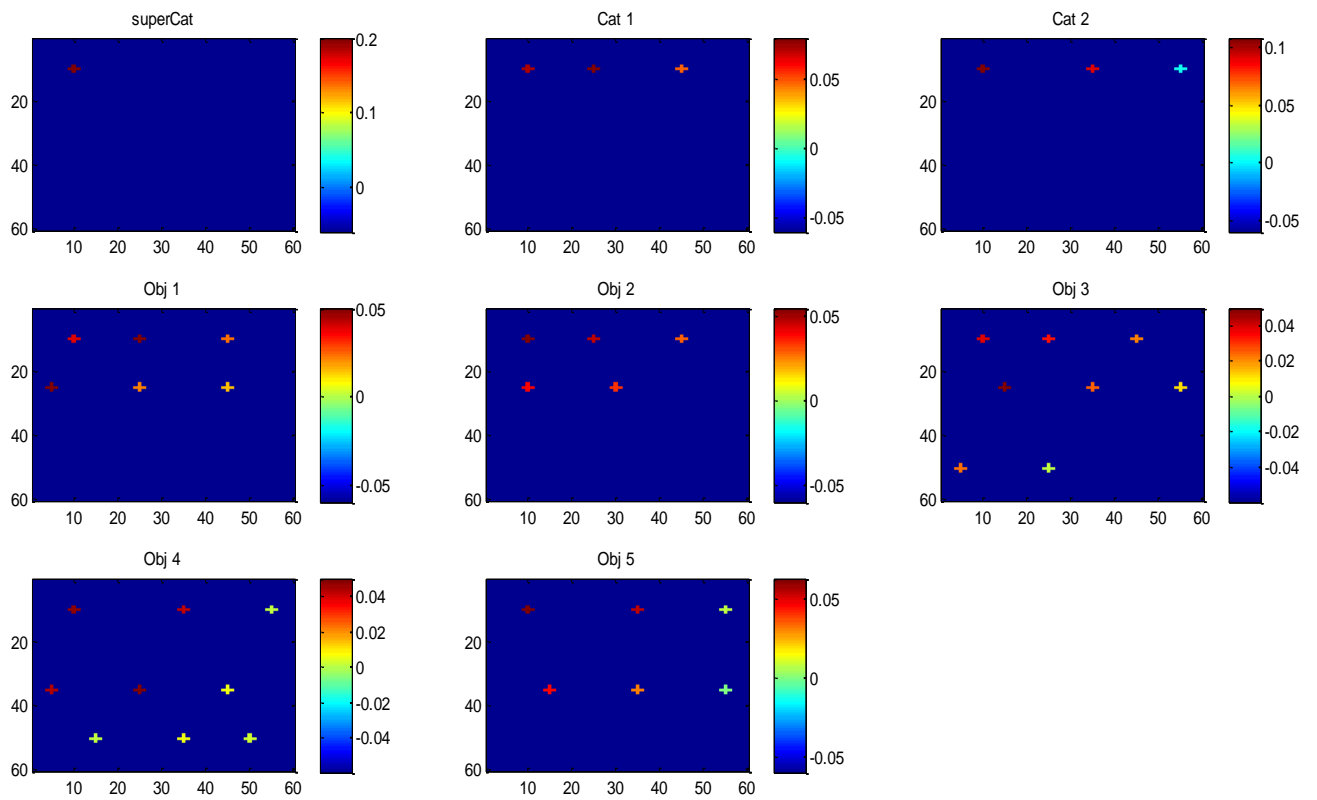


Illustrazione 16: sinapsi entranti alle parole, WLC. In alto sono riportati: la supercategoria, la categoria 1 e 2; sotto, i 5 oggetti.

La successiva figura 17 mostra le sinapsi WCL, entranti agli attributi da parte delle parole. Anche in questo caso, a favore della sintesi, commenterò solo alcuni dei riquadri rappresentati.

La feature “A”, in quanto comune agli elementi della rete, riceve sinapsi più o meno elevate da tutti, in proporzione alla maggiore o minore percentuale di salienza, propria di ciascuno di essi.

La proprietà “B”, comune saliente al 90%, riceve sinapsi forti dalla I° categoria e dagli oggetti: 1, 2, 3 ma nulle dagli altri elementi. La caratteristica “f1” riceve sinapsi positive solo dall’oggetto 1 a cui appartiene. Le proprietà “f5” e “g5”, distintive salienti rispettivamente all’80% e 70%, ricevono sinapsi nulle tranne che dall’oggetto 5. Mentre la caratteristica “m4”, in quanto non saliente al 40%, riceve sinapsi dall’oggetto 4, di valore pari circa a: -1 e di valore ancora inferiore dagli altri.

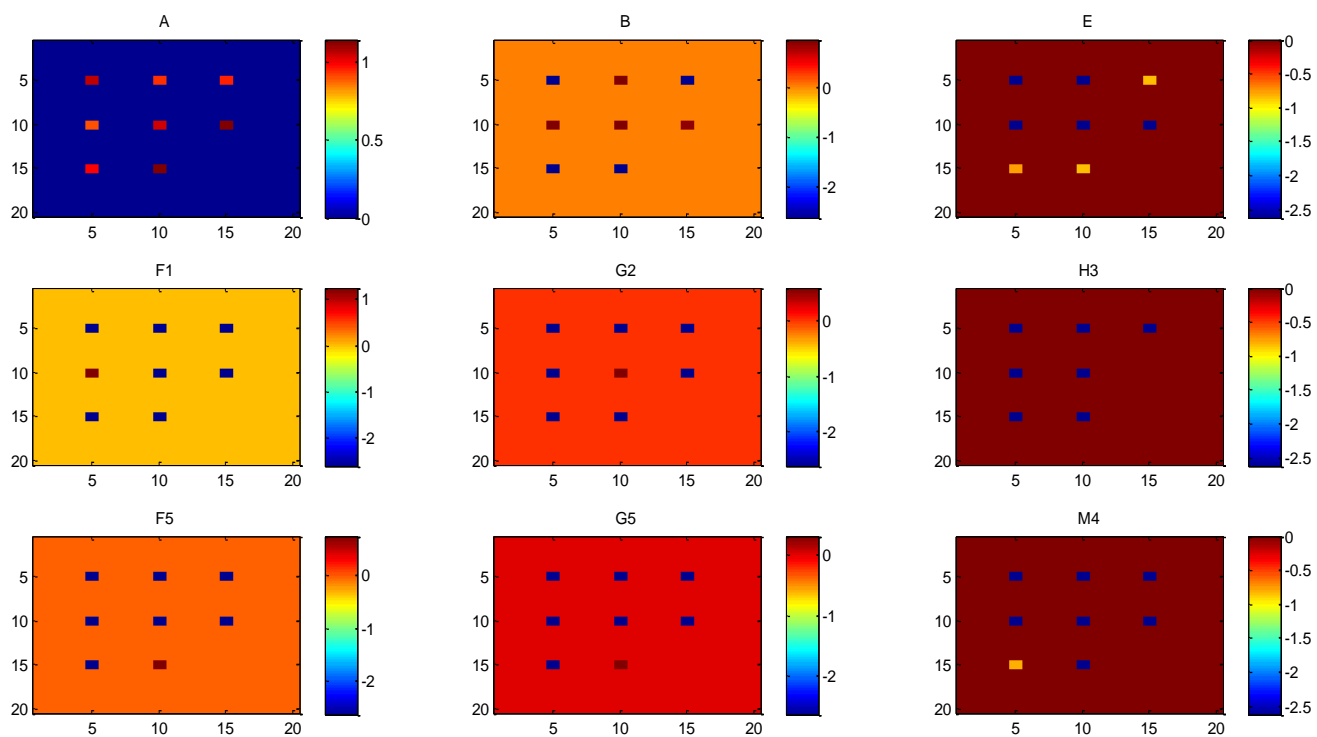


Illustrazione 17: sinapsi entranti agli attributi WCL.

Un caso anomalo da segnalare riguarda la feature “h3”, saliente al 70% ed appartenente all’oggetto 3: essa riceve in ingresso sinapsi nulle da tutti gli elementi del modello. La spiegazione è semplice: è una questione dovuta alla casualità con cui sono eseguite le permutazioni durante l’addestramento. Tuttavia, essa riceve sinapsi forti dalle altre proprietà dell’oggetto 3, all’interno della rete semantica: la successiva figura 18 mostra le suddette sinapsi. In particolare: f3, saliente al 100% e g3, saliente all’80% inviano sinapsi di valore pari a circa 0.04; di valore lievemente inferiore sono inviate da i3, saliente al 60%. La proprietà l3, non saliente al 40%, invia sinapsi all’incirca dimezzate rispetto alle altre salienti. In questo modo, la caratteristica h3 viene evocata correttamente così da partecipare alla rappresentazione semantica dell’oggetto a cui appartiene. Probabilmente questo è il motivo per cui essa riesce ad essere richiamata, sebbene sia saliente al 60%.

Analizziamo alcuni casi: la feature “A”, comune a tutti i concetti al 90%, riceve sinapsi positive dal maggior numero di attributi; quelle di valore superiore sono inviate da: B, C, D, con alta percentuale di salienza. Riducendosi quest’ultima, le altre features inviano sinapsi sempre più ridotte, fino al valore nullo per quelle con salienza inferiore o uguale al 60%, come ad esempio: h4, h5.

La proprietà “E”, in accordo con la sua definizione, riceve sinapsi molto deboli: riceve da D, le sinapsi di valore maggiore pari a: 0.025; di valore inferiore, pari a: 0.015, da f5, saliente all’80% e di valore nullo da tutte le altre. Quest’ultima è attivata solo da g5, saliente al 70% ed in misura minore da h5, non saliente al 50%.

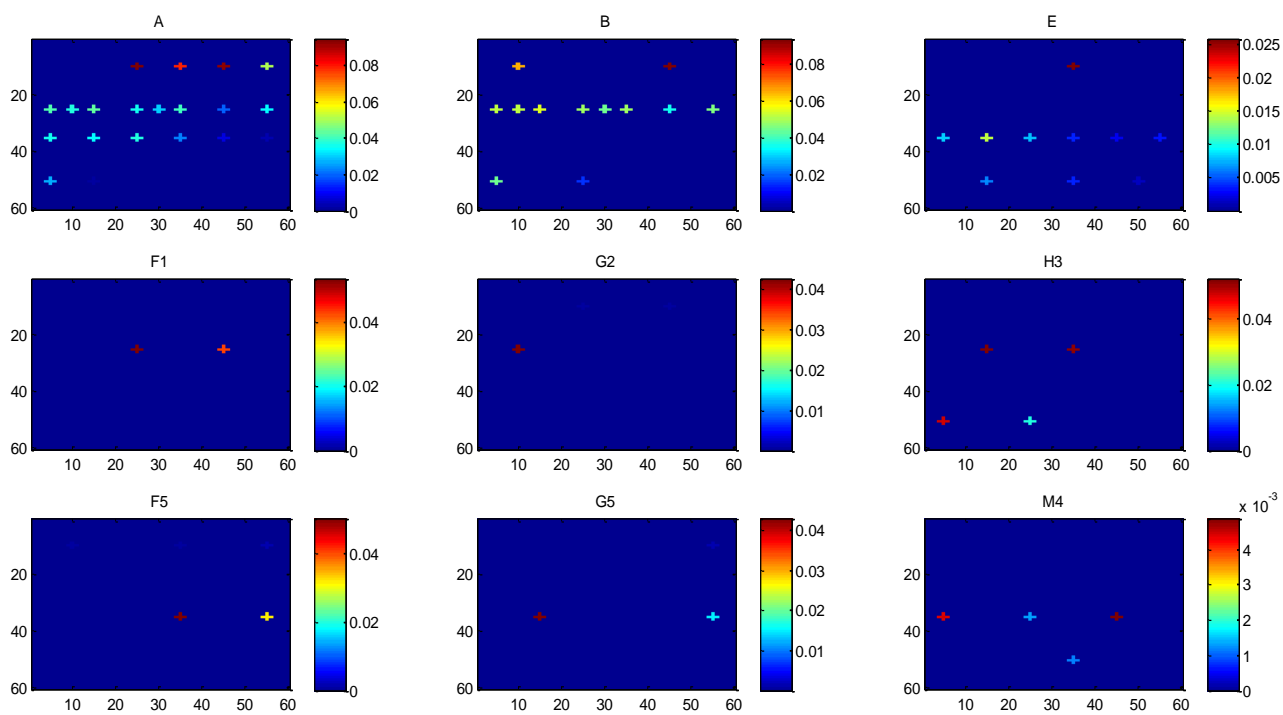


Illustrazione 18: sinapsi entranti agli attributi, all'interno della rete semantica.

Di seguito sono riportati i grafici che rappresentano la simulazione inerente all'oggetto 3, a dimostrazione di quanto appena affermato; in particolare, è evidenziata la sequenzialità temporale con cui sono richiamate le features del suddetto elemento. In essi, a sinistra è rappresentata la memoria semantica mentre a destra è posizionata la memoria lessicale. Somministro in ingresso la parola corrispondente; dalla figura 18, rappresentativa del passo 11, come è anche riportato nella tabella 9, lo stimolo consente di evocare immediatamente le proprietà: A, B, C, f3, g3, con percentuali di salienza piuttosto importanti. Successivamente, al passo 21, cominciano ad emergere anche "h3" ed "i3", salienti al 70% e 60%, lasciando fuori "l3", non saliente al 40%. Le 5 proprietà, evocate per prime, attraverso il rinforzo reciproco delle sinapsi con la parola corrispondente, cooperano nel richiamare anche le altre 2. Inoltre al passo 41, si nota come gli attributi siano circa presenti allo stesso grado di attivazione.

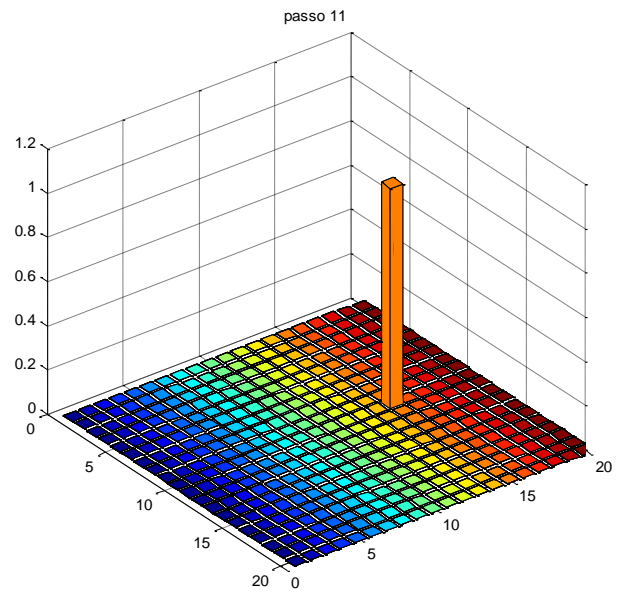
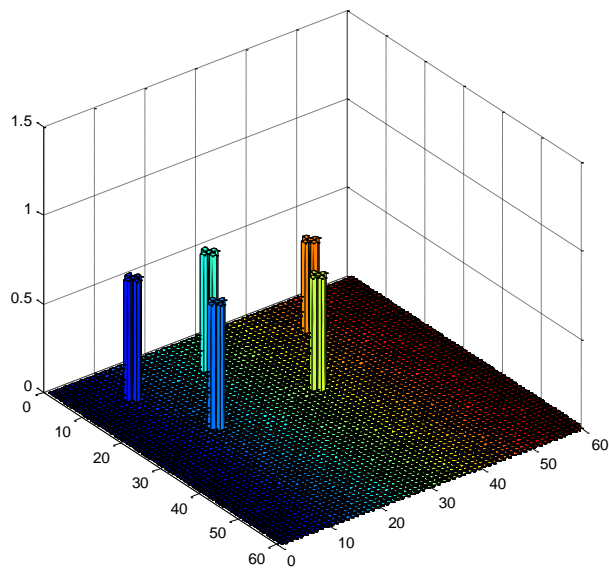


Illustrazione 18: risultati della simulazione al passo 11, dando in ingresso la parola relativa all'oggetto 3 (rappresentata nello spazio a sinistra). Subito emergono le proprietà: A, B, C, f3, g3.

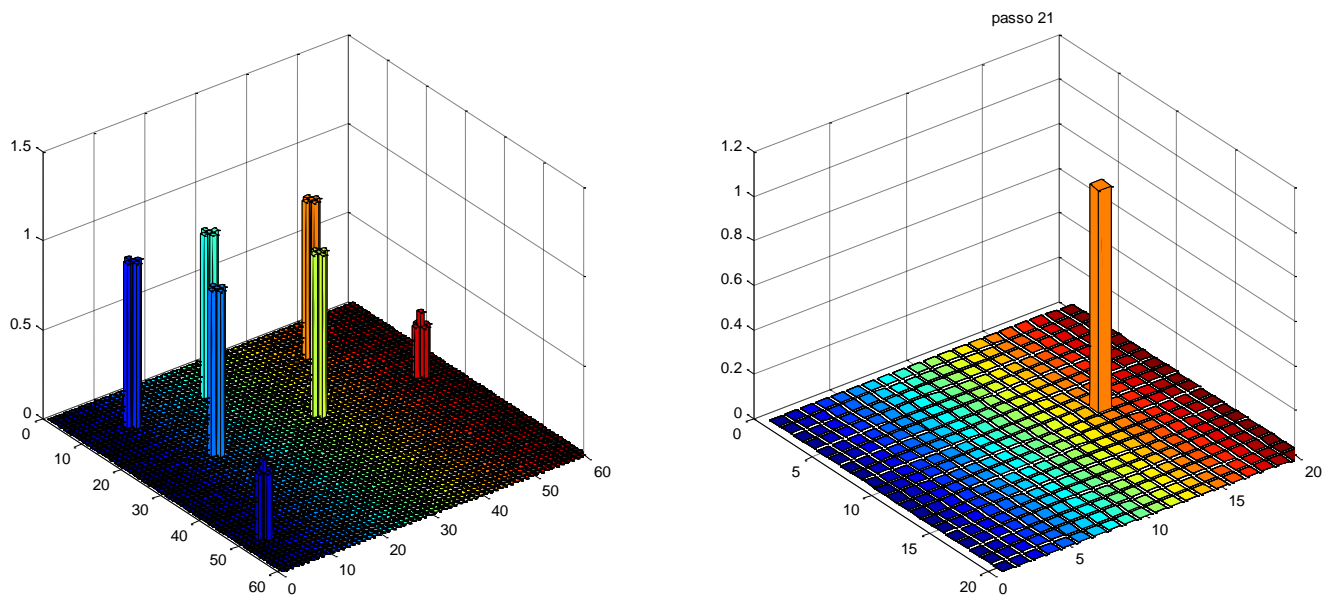


Illustrazione 19: risultati della simulazione al passo 21, dando in ingresso la parola relativa all'oggetto 3 (rappresentata nello spazio a sinistra). Vengono richiamate anche le proprietà h3 ed i3, fotografate ad un istante di attivazione inferiore rispetto alle altre.

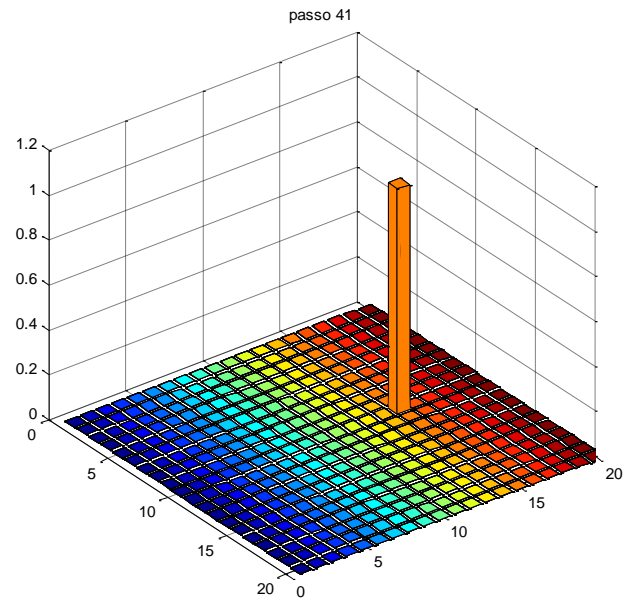
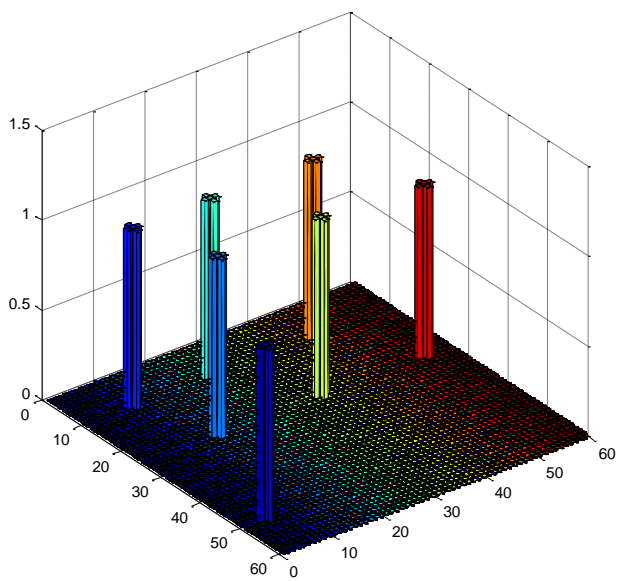


Illustrazione 20: risultati della simulazione al passo 41, dando in ingresso la parola relativa all'oggetto 3 (rappresentata nello spazio a sinistra). Le 7 proprietà evocate, sono rappresentate allo stesso stato di attivazione.

CONCLUSIONI

Lo scopo del presente lavoro è stato di valutare il ruolo della “salienda” delle caratteristiche, nel riconoscimento degli oggetti e validare una nuova rete neurale, proposta per la rappresentazione semantica e lessicale. Dapprima si è deciso di testare il modello su una struttura con due oggetti, aventi lo stesso numero di proprietà; successivamente è stata implementata ed analizzata anche una tassonomia più complessa, costituita da cinque concetti.

L’algoritmo utilizzato ha prodotto 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; le caratteristiche non salienti, in alcuni casi, sono riuscite a richiamare le salienti ma con fatica.

Nel caso del modello composto da due oggetti, le modifiche relative al valore di soglia post-sinaptica ed al numero di prove per l’addestramento della memoria lessicale, hanno prodotto i risultati sperati. L’aumento del primo parametro, ha causato un maggiore indebolimento delle sinapsi, rispetto al loro rinforzo. In particolare, tra il valore 0.5 e 0.7, le differenze sostanziali si sono riscontrate relativamente alle sinapsi entranti agli oggetti: esse andavano gradualmente riducendosi di valore. Per quanto riguarda la “salienda” delle caratteristiche, dall’analisi delle sinapsi entranti alle features, è emersa la seguente situazione: l’aumento della soglia trasformava le proprietà, anche con una percentuale di presenza medio-alta, da salienti a non salienti, nel caso di soglia pari a 0.7. Inoltre, portando il numero di simulazioni da 20 a 80 per il training della memoria lessicale, è stato possibile correggere

alcune situazioni indesiderate, simulando il progressivo addestramento della rete.

Sulla base dei risultati ottenuti con il suddetto modello, sono stati scelti i parametri con cui eseguire le simulazioni della tassonomia più complessa. Questa volta, a fare maggiormente la differenza, è stato il numero di prove scelto per l'addestramento della memoria lessicale; con 60 tentativi, sono state riscontrate delle anomalie.

Ad esempio: somministrando in ingresso la proprietà "a", la rete non riusciva ad evocare il "superconcetto" relativo. Alzando il numero di prove ad 80, lasciando invariati gli altri parametri, le questioni indesiderate sono state tutte risolte.

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

Il modello, proposto in questo lavoro di tesi, è stato addestrato usando degli oggetti semplici come ingressi, rappresentati da vettori di features. I progetti futuri riguarderanno l'analisi di un database, ricevuto dal Dipartimento di Neuropsicologia dell'ospedale San Raffaele a Milano, costituito dai dati relativi ad oggetti reali e la conseguente implementazione di un modello matematico che lo rappresenti. Per produrre tale database, un gruppo di soggetti sani si è espresso enunciando gli attributi caratteristici di alcuni oggetti, appartenenti a categorie differenti.

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, 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!

Durante la convivenza lontano dagli affetti, la serenità in casa è fondamentale per un cammino universitario e personale, sereno e produttivo. Le coinquiline di questi ultimi due anni, in particolare: Evelyn, Giulia e Valeria, sono state per me delle amiche sincere. Insieme abbiamo condiviso le gioie e i dolori dello studio, come anche i sorprendenti momenti della vita quotidiana: ci siamo sempre rispettate e soprattutto volute bene. Mi auguro che la nostra amicizia duri il più a lungo possibile, così da vedere realizzati tutti gli obiettivi ed i sogni che stiamo coltivando con tanta dedizione. Intendo inoltre ringraziare tutti gli amici ed i compagni di corso, della triennale e della magistrale, che mi sono stati accanto e non hanno esitato a tendermi la mano nei momenti di bisogno: Andrea, Ausilia, Benedetta, Chiara, Cinzia L., Cinzia N, Claudia, Domenico, Fatima, Giancarlo, Giuseppe, Marco, Mario, Matteo, Sara, Simon, Simone, Veronica.

BIBLIOGRAFIA

1. 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*.
2. Barsalou Lawrence W., Department of Psychology, Emory University, Atlanta (1999), *Perceptual symbol systems*.
3. Barsalou Lawrence W., Department of Psychology, Emory University, Atlanta, Georgia (2003), *Grounded Cognition*.
4. Barsalou Lawrence W., Department of Psychology, Emory University, Atlanta, USA (2009), *Simulation, situated conceptualization, and prediction*.
5. Barsalou Lawrence W. , Department of Psychology, Emory University, Atlanta, Georgia (2008) *Grounding Symbolic Operations in the Brain's Modal Systems*.
6. Decety Jean, Julie Grezes, Department of Psychology, The University of Chicago, 5848 S. University Avenue Chicago, IL 60637, USA, Laboratoire de Physiologie de la Perception et de l'Action, CNRS, Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France (2006), *The power of simulation: Imagining one's own and other's behavior*.
7. Floreano Dario, Stefano Nolfi, Università di Trieste e C.N.R. Roma (2000) *Reti neurali: algoritmi di apprendimento, ambiente di apprendimento, architettura*.
8. GLENBERG ARTHUR M. and MICHAEL P. KASCHAK, University of Wisconsin, Madison, Wisconsin (1997), *Grounding language in action*.

9. Goldman Mark S., Jorge Golowasch, Eve Marder, L. F. Abbott Volen Center and Department of Biology, Brandeis University, Waltham, Massachusetts , and Department of Physics, Harvard University, Cambridge, Massachusetts (2001) *Global Structure, Robustness, and Modulation of Neuronal Models*.
10. HART JOHN JR., RAKSHA ANAND, SANDRA ZOCCOLI, MANDY MAGUIRE, JACQUE GAMINO, GAIL TILLMAN, RICHARD KING, MICHAEL A. KRAUT (2007) *Neural substrates of semantic memory*.
11. Stephen M. Kosslyn, Harvard University, Cambridge, MA, USA (2005), *Mental Images and the brain*.
12. Martin Alex, Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, Maryland (2007) *The Representation of Object Concepts in the Brain*.
13. Alex Martin and Linda L. Chao (2001) *Semantic memory and the brain: structure and processes*.
14. Rizzolatti G., Craighero L., Dipartimento di Neuroscienze, Sezione di Fisiologia, Universita di Parma, Italy (2004), *The mirror-neuron system*.
15. Robbins Philip, Murat Aydede, Cambridge University (2007), *Situated Cognition*..
16. 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*.
17. 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.

18. 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.*