

Alma Mater Studiorum - Università di Bologna

Campus di Cesena

Scuola di Ingegneria e Architettura

Corso di Laurea Magistrale in Ingegneria Biomedica

Titolo della tesi:

MODELLI NEURO-COMPUTAZIONALI DI MEMORIA
SEMANTICA: ANALISI DELL'APPRENDIMENTO
DIPENDENTE DAL CONTESTO E SINCRONISMO NEURALE

Tesi in:

SISTEMI NEURALI LM

Relatore:

Prof. Mauro Ursino

Presentato da:

Alice Monti

II Sessione

Anno Accademico 2018-2019

INDICE

Introduzione	1
-------------------------------	----------

Capitolo 1

MEMORIA SEMANTICA	5
------------------------------------	----------

1.1 Memoria semantica e rappresentazione concettuale	5
1.2 Descrittori fondamentali di <i>features</i>	7
1.3 <i>Database</i> e tassonomie semantiche	10
1.3.1 Tassonomia di Animali	12
1.3.2 Tassonomia di Oggetti	15

Capitolo 2

MODELLO DI MEMORIA SEMANTICA	17
---	-----------

2.1 Modello della rete semantica	17
2.2 Formulazione matematica	18
2.3 Addestramento delle sinapsi	20
2.3.1 Apprendimento con la regola di Hebb	20
2.3.2 Procedura di addestramento	22
2.3.3 Requisiti della rete semantica	23
2.3.4 Problema della soglia post-sinaptica	24
2.3.5 Assegnazione dei parametri	27
2.4 Risultati	28
2.4.1 Risultati dell'addestramento sinaptico	28
2.4.2 Risultati delle simulazioni	33

Capitolo 3

MODELLO DI MEMORIA SEMANTICA CON CONTESTO39

3.1 Contesto	39
3.1.1 Influenza del contesto sui concetti	39
3.1.2 <i>Features</i> contesto-dipendenti e indipendenti	40
3.2 Formulazione matematica del contesto	42
3.3 Addestramento delle sinapsi con contesto	44
3.3.1 Procedura di addestramento	44
3.3.2 Assegnazione dei parametri	47
3.4 Risultati	48
3.4.1 Risultati dell'addestramento sinaptico	48
3.4.2 Risultati delle simulazioni.	51

Capitolo 4

MODELLO DI MEMORIA SEMANTICA CON OSCILLATORI 59

4.1 Sincronismo in banda gamma e oscillatori	59
4.2 Formulazione matematica degli oscillatori	61
4.3 Addestramento delle sinapsi	65
4.4 Assegnazione dei parametri	66
4.5 Risultati	68
4.5.1 Risultati delle simulazioni del Modello I	69
4.5.2 Risultati delle simulazioni del Modello II	74

Conclusioni79

Appendice85

Bibliografia91

Introduzione

L'organizzazione e i meccanismi neurali della memoria semantica sono una questione centrale nell'ambito delle neuroscienze cognitive, non solo per il loro interesse teorico, ma anche per capire *deficit* di conoscenza in disturbi neurologici quali Alzheimer e demenza semantica. Negli ultimi decenni, a proposito di questo argomento, sono stati condotti molteplici studi neuropsicologici e di neuroimaging e sono stati sviluppati modelli matematici e teorici con lo scopo di analizzare e studiare i meccanismi alla base della memoria semantica. Grazie a questi modelli neuro-computazionali si può cercare di comprendere come la conoscenza del significato dei concetti venga immagazzinata all'interno del nostro complesso sistema nervoso e successivamente utilizzata. Un tipo di rete neurale largamente impiegata come memoria è la rete basata su dinamiche di attrattori: in tale rete le informazioni che caratterizzano i singoli concetti vengono espresse dall'attivazione simultanea di gruppi di neuroni (nodi), che codificano per le proprietà caratterizzanti. Il principio alla base di tale modellazione è che la semantica di un concetto sia descritta attraverso una collezione di *features* (caratteristiche) che sintetizzano la conoscenza completa del concetto e che sono codificate nelle diverse regioni corticali. Attraverso l'apprendimento sinaptico basato sulla regola di Hebb, la rappresentazione del singolo concetto viene memorizzata come punto di equilibrio della rete; l'informazione completa può essere successivamente ripristinata a partire da un punto iniziale parziale (cioè dall'attivazione di pochi nodi) sfruttando le dinamiche attrattive del modello che converge verso l'equilibrio.

Le reti di attrattori basate su caratteristiche forniscono intuizioni significative su diversi fenomeni comportamentali riguardanti la memoria semantica in condizioni normali ma anche in condizioni patologiche. In particolare, i modelli basati su attrattori sono in grado di spiegare il diverso ruolo svolto dalle *features* nella rappresentazione semantica concettuale e la distinzione tra concetti subordinati e sovraordinati.

Nonostante questi importanti risultati, diversi aspetti organizzativi della memoria semantica non sono ancora sufficientemente noti e molteplici comportamenti in alcune patologie neurodegenerative richiedono ancora un'adeguata formalizzazione nell'ambito delle reti neurali.

In accordo con una organizzazione della memoria semantica basata su attributi, il mio lavoro di tesi è finalizzato ad indagare, attraverso una rete attrattiva, il ruolo di diversi tipi di caratteristiche nella rappresentazione dei concetti, presupponendo che esse possano avere un ruolo differente nell'archiviazione e nel ripristino degli oggetti e che, basandosi sulla somiglianza, un oggetto possa essere riconosciuto anche se alcune delle sue proprietà sono mancanti. In particolare, l'obiettivo di questo studio è quello di estendere e ottimizzare un modello di rete semantica ad una dimensione, sviluppato negli anni precedenti dagli studi del Prof. Ursino M. Attraverso tale modello si cercherà di: indagare i meccanismi che portano alla formazione di categorie, analizzare il ruolo delle diverse proprietà nella costruzione di un concetto e comprendere come le connessioni tra queste proprietà siano indotte dall'esperienza. In effetti, non tutte le *features* hanno la stessa importanza nel riconoscimento di un concetto, né sono tutte spontaneamente evocate quando si pensa ad un concetto.

Una ipotesi originale del presente lavoro è che la semantica di un oggetto non sia fissa, bensì possa dipendere dal particolare contesto in cui l'oggetto è utilizzato o si trova. Per simulare questo fenomeno, per ogni concetto sono state selezionate alcune sue proprietà che tendono a verificarsi spesso insieme, individuando quindi un contesto particolare e specifico.

Si è voluto inoltre studiare un modello estremamente semplificato, basato sulla teoria del sincronismo oscillatorio in banda gamma, che prevede l'utilizzo di oscillatori per mantenere in memoria contemporaneamente differenti rappresentazioni semantiche. Si presume infatti che la sincronizzazione delle oscillazioni consenta il ripristino simultaneo in memoria di concetti differenti realizzando una corretta soluzione del problema di "*binding & segmentation*".

In generale, le simulazioni ottenute a seguito dell'addestramento mostrano come, evocando alcune *features*, siano richiamate le altre caratteristiche salienti; inoltre tali risultati possono fornire indicazioni su quali meccanismi neurali possono essere sfruttati per formare robuste categorie tra concetti. In particolare, la rete addestrata è in grado di risolvere efficacemente compiti di riconoscimento di concetti, mantenendo una distinzione tra le categorie e i suoi membri, e fornendo un diverso ruolo per le *features* salienti rispetto alle proprietà marginali.

La tesi è articolata nel modo seguente. Nel primo capitolo, dopo una breve panoramica sul concetto di memoria semantica e sulla problematica della rappresentazione concettuale, si passerà ad esaminare i descrittori fondamentali delle *features* e si descriveranno nel dettaglio le due ampie tassonomie sulle quali sono realizzati i modelli utilizzati nel presente lavoro. Il secondo capitolo prevede invece l'analisi completa, sia a livello qualitativo sia a livello matematico, del modello di rete semantica basato sulle dinamiche attrattive. Il terzo capitolo si concentra sulla trattazione relativa al modello implementato tenendo conto del contesto, individuando la principale differenza tra caratteristiche dipendenti e indipendenti dal contesto. Per questi due modelli si riportano i risultati relativi alla fase di addestramento sinaptico e alla successiva attività di riconoscimento dei concetti. Infine, nel quarto capitolo si prenderà in considerazione una semplice rete basata su oscillatori, che consente il mantenimento in memoria di più concetti, distinti tra loro grazie alla desincronizzazione delle oscillazioni in banda gamma.

Capitolo 1

MEMORIA SEMANTICA

1.1 Memoria semantica e rappresentazione concettuale

Il dominio della memoria semantica comprende informazioni riguardanti le proprietà dei concetti e tutti i processi che consentono in maniera efficiente di recuperare, utilizzare e mettere le informazioni al servizio del pensiero e del linguaggio. La memoria semantica riguarda le conoscenze generali sul mondo, non è personale bensì comune a tutti; è il magazzino della memoria a lungo termine specializzato nell'acquisizione di informazioni su fatti e conoscenze astratte. Tale memoria ha il preciso scopo di memorizzare ciò che tanti episodi hanno in comune, al fine di creare uno stereotipo del concetto che prescinde dai particolari. Una tipica prestazione della memoria semantica è, ad esempio, la capacità di descrivere verbalmente le caratteristiche di un oggetto conosciuto o di un luogo in cui si è svolta una certa esperienza. La memoria semantica si differenzia da quella episodica, anche se entrambe sono memorie dichiarative in quanto disponibili come conoscenza consapevole in cui il soggetto è cosciente ed in grado di esprimere ciò che ha memorizzato. La memoria episodica invece è personale, autobiografica, legata fortemente al contesto e non condivisa con altri. Essa si occupa della memorizzazione dei particolari di un episodio (ad esempio il ricordo individuale di un gatto in un preciso momento) ed è di solito acquisita attraverso un'unica esperienza [8].

Il modo in cui i concetti vengono organizzati e archiviati in memoria risulta un aspetto cruciale per molteplici processi della cognizione umana quali il riconoscimento di oggetti, la pianificazione di azioni, il linguaggio, la memoria episodica e la formazione di idee astratte. L'organizzazione delle informazioni concettuali influenza inevitabilmente il comportamento umano, dal momento che governa sia la modalità attraverso cui il mondo esterno determina i concetti sia il modo in cui i tali concetti influenzano l'azione, trasmettendo informazioni tra diverse regioni del nostro cervello.

Per questa ragione, lo studio della memoria semantica attraverso modelli teorici svolge un ruolo fondamentale nelle moderne neuroscienze cognitive. Diverse problematiche riguardo ai meccanismi di funzionamento della memoria semantica meritano particolare attenzione e, ad oggi, tali meccanismi non sono stati compresi ancora del tutto. Non è chiaro come avvenga la formazione di categorie a partire dalla individuale esperienza senso-motoria o come avvenga la distinzione tra concetti sovraordinati (cioè le categorie) e concetti subordinati (ovvero i singoli membri di una data categoria). Inoltre, nell'ambito del processo di riconoscimento, ci si interroga sul ruolo dei diversi attributi che costituiscono un oggetto [2,10].

Sebbene gli studiosi si siano arrovellati per secoli circa la natura dei concetti, la memoria semantica è formalmente diventata un argomento di studio nelle scienze cognitive solo di recente. Questo fatto risulta notevole, considerato che la memoria semantica è uno dei tratti umani maggiormente definiti, il quale comprende la maggior parte delle conoscenze che acquisiamo sul mondo attorno a noi. Notevole è anche la varietà di attività cognitive di tutti i giorni che dipendono da questo vasto archivio di conoscenze. L'uomo utilizza la conoscenza concettuale non solo per una semplice e quotidiana interazione con gli oggetti; l'intero bagaglio culturale dell'uomo, tra cui la scienza, la letteratura, le istituzioni sociali, si basa ed è costruito sulla conoscenza concettuale. Non si potrebbe ragionare, pianificare il futuro e nemmeno ricordare il passato senza conoscenza concettuale. Tutte queste attività dipendono dall'attivazione di concetti memorizzati nel vasto archivio della memoria semantica [2].

Diverse teorie riguardanti la struttura della memoria semantica cercano di dare risposta a tutti questi interrogativi. La maggior parte dei modelli teorici connessionisti, ormai ad oggi molto affermati, presuppone che gli oggetti concreti siano rappresentati all'interno di un magazzino semantico come una collezione di "*features*" (caratteristiche o proprietà), a-modali o modalità-specifiche, codificate in diverse regioni del cervello e reciprocamente interconnesse in una flessibile e altamente plastica rete distribuita.

Secondo tale visione, i concetti vengono rappresentati in memoria come una raccolta di caratteristiche che includono più tipi di informazioni, a somma delle nostre precedenti esperienze sensoriali, motorie, emotive, funzionali, enciclopediche e la manipolazione di tali proprietà sembra essere alla base dell'elaborazione concettuale [9,11].

All'interno di questo ambito, molte questioni controverse circondano ancora il ruolo delle *features*: la vera natura della loro organizzazione neurale, la struttura computazionale, che consente di tradurre la nostra esperienza in concetti, e il modo in cui esse influenzino il comportamento umano [10].

1.2 Descrittori fondamentali di *features*

In un modello connessionista basato su caratteristiche predomina una importante questione teorica ovvero, quale sia il ruolo assunto dai diversi attributi che concorrono a formare concetti. Si può affermare infatti che non tutte le *features* acquistino la stessa "importanza" nell'ambito del riconoscimento di oggetti, né tutte sono evocate spontaneamente quando si pensa ad un concetto [10]. Una banale intuizione è che l'idea di "cane" possa essere più facilmente identificata dalla caratteristica "abbaia" piuttosto che dalla proprietà "ha quattro zampe". Numerosi autori hanno quindi sostenuto che non tutte le *features* semantiche possiedano una rilevanza tra loro simile per la rappresentazione concettuale.

Per prima cosa sono state proposte molteplici distinzioni tra proprietà, in particolare è stato evidenziato che le caratteristiche sensoriali svolgono un incarico principale nella rappresentazione di entità viventi mentre le caratteristiche funzionali sono fondamentali per l'espressione di entità non viventi. In generale le *features* differiscono non solo per il loro contenuto, ma anche in termini di informazioni (possono essere sensoriali, motorie, enciclopediche) o rilevanza sull'entità che descrivono [4,9].

L'idea che i concetti siano una raccolta di attributi derivanti dalla nostra percezione ed esperienza ha portato all'introduzione di svariati descrittori per poter distinguere e classificare le *features* semantiche.

Tali descrittori sono non solo di tipo binario (“true” o “false” booleani), bensì possono assumere valori graduali, possono cioè essere associati ad un valore numerico intero. Un primo importante descrittore è la distintività; una caratteristica si definisce distintiva se risulta importante al fine di distinguere o discriminare un dato concetto da altri. Nel caso estremo, una caratteristica appartenente solamente ad un unico concetto (“abbaia” per l’idea di “cane”) è perfettamente distintiva [10].

Definiamo invece una *feature* come condivisa se essa appartiene a più concetti contemporaneamente (“ha il pelo corto” vale sia per il concetto di “cane” sia per quello di “gatto”). Idea comune è che le caratteristiche condivise contribuiscano alla formazione di concetti sovraordinati ovvero alla formazione di categorie (ad esempio “ha le ali” e “ha le piume” per la categoria “UCCELLO”), mentre le caratteristiche distintive consentono ai singoli membri di poter essere riconosciuti. Per quanto riguarda le *features* condivise, alcune di queste sono solo parzialmente condivise, in quanto caratterizzano solo alcuni membri di una categoria (ma non tutti). Quindi, non determinano la semantica della categoria generale, né possono essere utilizzate per identificare i singoli membri. Esse giocano però un ruolo centrale nell’assegnazione di singoli membri a una categoria specifica [9].

Una misura comune dell’importanza di una *feature* è la sua dominanza, che quantifica la frequenza della produzione della caratteristica stessa: indica cioè la percentuale di partecipanti che elencano un particolare attributo per un dato concetto in un’attività di elencazione. Si prevede che una *feature* che si verifichi più frequentemente venga memorizzata fortemente nella memoria semantica, di conseguenza acquisirà una dominanza maggiore. In questo caso si parla di attributo saliente (o ad elevata dominanza) il quale è evocato spontaneamente quando si pensa a un oggetto. All’altro estremo, una *feature* che emerge raramente, caratterizzata quindi da una bassa dominanza, viene considerata marginale per un dato concetto. La marginalità presuppone che, sebbene un attributo contribuisca al riconoscimento di un concetto, esso venga in mente meno frequentemente.

La *feature* “usato per mandare messaggi” è distintiva in quanto è elencata solo per il concetto di “piccione”. Allo stesso tempo questa proprietà non risulta saliente, poiché raramente viene riportata in compiti di elencazione di caratteristiche, presumibilmente perché non svolge una funzione estremamente importante nella rappresentazione del concetto “piccione”. L’attributo “tuba”, d’altro canto, è distintivo poiché permette l’immediato riconoscimento del concetto “piccione” ma anche saliente poiché riportato da molti soggetti in compiti elencazione di *features* [4,9].

In secondo luogo, una caratteristica può avere un grande impatto emotivo (ad esempio “fa paura” per il concetto “pistola”) o può avere una grande utilità funzionale (come “usata per cucinare” per il concetto di “pentola”) che potrebbe favorire la sua memorizzazione nel magazzino semantico. Si ritiene infatti che anche questi due importanti aspetti di una *feature* possano influenzarne la dominanza. Di contro, i concetti sono più difficili da identificare se hanno molti vicini semantici (cioè concetti che condividono un numero sostanziale di caratteristiche), dal momento che possono essere confusi tra loro (come nel caso dei due oggetti “divano” e “poltrona”). Diversi studi hanno infatti identificato la distanza semantica, descrittore che misura la somiglianza tra rappresentazioni, come un altro fattore molto influente, il quale tiene conto della specificità di categoria [4].

A questo punto abbiamo due dimensioni principali per quantificare la mansione delle *features* all’interno di una struttura semantica: una dimensione globale che misura l’importanza di una caratteristica per tutti i concetti all’interno di un *database* (distintività contro condivisione) ed una dimensione locale che misura l’importanza della *feature* semantica per il concetto (dominanza contro marginalità). Infine, da svariati studi è emerso che molteplici descrittori di proprietà semantiche, tra i quali alcuni appena citati (carattere distintivo e dominanza) e altri come la rilevanza semantica (calcolata come combinazione non lineare tra dominanza e carattere distintivo) e la distanza semantica, sembrano spiegare una varietà di fenomeni comportamentali, sia in soggetti sani sia in pazienti neurologici.

Riassumendo, le seguenti definizioni riprendono gli aspetti fondamentali delle *features* che verranno poi utilizzate nei capitoli seguenti:

1. **Condivisa:** è una caratteristica che appartiene a diversi concetti.
 - **Totalmente condivisa:** è quella caratteristica appartenente a tutti i membri di una categoria.
 - **Parzialmente condivisa:** è quella caratteristica appartenente solo ad alcuni membri di una categoria (più di un singolo membro).
2. **Distintiva:** è una caratteristica che appartiene ad un solo concetto;
 - **Saliente distintiva:** è quella caratteristica distintiva che viene evocata spontaneamente quando si pensa all'oggetto, cioè un attributo frequentemente usato per definire un dato concetto.
 - **Marginale distintiva:** è quella caratteristica che non viene frequentemente evocata nel definire un concetto.

1.3 Database e tassonomie semantiche

Lo scopo del mio lavoro di tesi è stato quello di realizzare un modello neuro-computazionale basato su dinamiche di attrattori che prevede l'utilizzo di tassonomie semantiche e che si propone di riprodurre il comportamento della memoria semantica. Le tassonomie alla base del modello sono state modificate a partire da un *set* di dati raccolto dalla Scuola Universitaria Superiore IUSS di Pavia (Dott.ssa E. Catricalà) [4]. Il *set* di dati originale conteneva ottantadue concetti, suddivisi in elementi viventi ("uccelli", "animali terrestri", "ortaggi", ecc.) e non viventi ("mobili", "utensili", "articoli da cucina", ecc.). A livello pratico, le *features* semantiche sono state raccolte mediante attività di elencazione, nelle quali ai partecipanti era stato chiesto di produrre elenchi di attributi per singoli concetti di interesse. Le descrizioni fornite dai partecipanti sono state poi elaborate utilizzando aspetti quali la categoria, le proprietà percettive (aspetto, suoni, odori, sensazioni, sapori), le proprietà funzionali (a cosa serve, dove/quando viene utilizzato) ma anche fatti enciclopedici (da dove proviene) o eventi storici.

Non è stato imposto alcun limite di tempo per l'attività di elencazione e l'ordine di presentazione degli esemplari è stato randomizzato su partecipanti e categorie. I dati ottenuti sono stati poi analizzati ed elaborati, nello specifico: i sinonimi sono stati registrati una singola volta, i quantificatori ridondanti sono stati rimossi, le *features* aggettivo-sostantivo (“ha quattro zampe”) sono state divise in due singoli attributi (“ha le zampe” e “ha quattro zampe”), quelle disgiuntive (“è marrone o grigio”) sono state separate in due diverse caratteristiche (“è marrone” ed “è grigio”) mentre quelle congiuntive sono state divise quando le parti contenevano informazioni differenti (ad esempio “ha le orecchie” e “ha orecchie appuntite”). A partire da questi dati, è stato poi possibile ottenere diversi indici sia a livello di *features* (carattere distintivo, dominanza, rilevanza e distanza semantica) sia a livello di concetto e raccogliere statistiche sulla distribuzione di svariati tipi di attributi [4,9].

La mia analisi non ha previsto l'utilizzo dell'intero *database* originale bensì solo di una precisa selezione di concetti ai quali si è cercato di associare un numero più o meno uniforme di proprietà; in alcuni casi non ho utilizzato l'intero numero di caratteristiche presenti nel *database* di partenza. Per addestrare il modello di rete semantica che verrà trattato in questo elaborato ci si è limitati a costruire due tassonomie che caratterizzano molteplici concetti per ANIMALI e OGGETTI. I concetti contenevano tutti i tipi di caratteristiche di interesse, ovvero *features* totalmente condivise, parzialmente condivise, distintive con alta dominanza (salienti) e distintive con bassa dominanza (marginali), permettendo così di verificare diverse ipotesi di lavoro che verranno chiarite in seguito. È infine necessaria una precisazione riguardo all'attribuzione di salienza e marginalità delle *features* nelle tassonomie utilizzate, in quanto si discosta dai criteri adottati in precedenza [10]. Innanzitutto, sono state considerate unicamente le caratteristiche distintive e parzialmente condivise di ciascun concetto e tali attributi sono stati ordinati secondo valori crescenti di dominanza (da tre a venti). A partire da tale elenco ordinato, per ciascun concetto si è deciso di considerare un rapporto pari al 50% tra *features* marginali e salienti.

In accordo con questo criterio emerge che ogni singolo concetto risulta costituito per metà da attributi salienti e per metà da attributi marginali. Occorre tenere presente che si è voluto di proposito escludere le proprietà totalmente condivise in quanto esse sono indispensabili per la formazione categoriale e dunque risultano necessariamente sempre salienti. Ciò ha permesso di ottenere una uniformità nelle due tassonomie implementate di animali e di oggetti sia in termini di tipologie di attributi sia in termini di una successiva ottimizzazione dei parametri della rete semantica.

1.3.1 Tassonomia di Animali

Questa prima tassonomia è costituita da 16 animali, ciascuno dei quali è mediamente caratterizzato da un numero di *features* compreso tra 25 e 35. In particolare, la tassonomia è realizzata da un totale di 231 proprietà di cui: 26 sono totalmente condivise (cioè che permettono la formazione di una categoria, riportate in Tabella 1), 159 sono distintive (riportate in Tabella 2) e 46 sono parzialmente condivise (riportate in Tabella 3). A partire dalla super categoria “ANIMALE”, identificata da attributi di categoria come “mangia”, “ha le zampe”, “si riproduce”, si differenziano le due principali macro categorie: “MAMMIFERO” (“ha quattro zampe”, “ha il pelo”, “ha la coda”) e “UCCELLO” (“ha due zampe”, “ha le piume”, “ha le ali”). La suddivisione ulteriore all’interno della categoria “UCCELLO” prevede la distinzione tra gli uccelli “VOLATILI” (identificati dalle *features* di categoria “vola” e “fa il nido”) ai quali appartengono i singoli concetti di “oca”, “pappagallo”, “piccione” e “gufo” e quelli “NON VOLATILI” (caratterizzati da “svolazza” e “ha la cresta”) che comprendono i due soli concetti di “gallina” e “gallo”. Più complesse e articolate risultano le categorie minori derivanti dalla macro categoria “MAMMIFERO”, in particolare si distinguono subito i mammiferi “DOMESTICI”, i quali presentano “da compagnia” come attributo di categoria. Di contro i mammiferi appartenenti alla categoria dei “SELVATICI” risultano ulteriormente disgiunti nelle sottoclassi “ERBIVORI” (con “mangia erba”) e “CARNIVORI” (“mangia la carne” e “ha gli artigli”).

Tra questi ultimi sono stati scelti i concetti di “orso” e “leone” mentre la categoria degli erbivori prevede una ultima discriminazione tra erbivori “AFRICANI”, contraddistinti dalle *features* categoriali “vive in Africa” e “vive nella savana” ed erbivori “DA ALLEVAMENTO”, riconoscibili grazie a “si ricava la carne” e “vive nella fattoria”. Tra questi ultimi concetti di animale troviamo “giraffa”, “zebra” ed “elefante” per la prima delle due categorie appena citate e “pecora”, “mucca”, “cavallo” per gli erbivori da allevamento.

pos	concetto	features di categoria
1	cane	2,3,4,5,6,7,8,9,10,11
2	gatto	2,3,4,5,6,7,8,9,10,11
3	orso	2,3,4,5,6,7,8,9,34,35,36
4	leone	2,3,4,5,6,7,8,9,34,35,36
5	giraffa	2,3,4,5,6,7,8,9,34,54,55,56
6	zebra	2,3,4,5,6,7,8,9,34,54,55,56
7	elefante	2,3,4,5,6,7,8,9,34,54,55,56
8	pecora	2,3,4,5,6,7,8,9,34,54,84,85
9	mucca	2,3,4,5,6,7,8,9,34,54,84,85
10	cavallo	2,3,4,5,6,7,8,9,34,54,84,85
11	gallina	2,3,4,5,113,114,115,116,117,118
12	gallo	2,3,4,5,113,114,115,116,117,118
13	oca	2,3,4,5,113,114,115,116,139,140
14	pappagallo	2,3,4,5,113,114,115,116,139,140
15	piccione	2,3,4,5,113,114,115,116,139,140
16	gufo	2,3,4,5,113,114,115,116,139,140

Tabella 1 – Nella prima colonna si riportano le posizioni associate al singolo concetto; nella seconda si riportano i nomi dei concetti della tassonomia. Nella terza colonna sono elencate le *features* condivise costituenti le categorie, indicate come posizione assunta all’interno della rete. Il significato delle *features* è riportato nella Figura A1 dell’Appendice.

La tassonomia degli animali risulta possedere membri semanticamente molto vicini tra loro, grazie alla presenza di un numero elevato di *features* condivise, consentendo quindi una ampia e dettagliata suddivisione categoriale. In generale, i concetti riguardanti gli esseri viventi sono semanticamente più vicini tra loro rispetto a quelli non viventi, perciò all’interno di questa tassonomia si riesce più facilmente ad operare una suddivisione in categorie. Occorre però tenere presente che la vicinanza semantica rende ancora più difficile il compito alla rete, cioè il riconoscimento di un concetto.

Per maggiore completezza si faccia riferimento alla Figura A1 nella sezione Appendice, in cui viene riportato lo schema completo della tassonomia di animali, con l'indicazione di tutte le *features* utilizzate.

pos	concetto	features distintive
1	cane	12,13,14,15,16,17,18,19,20,21,22,23
2	gatto	24,25,26,27,28, 29,30,31,32,33
3	orso	37,38,39,40,41,42,43,44,45
4	leone	46,47,48,49,50,51,52,53
5	giraffa	57,58,59,60,61,62,63,64,65
6	zebra	66,67,68,69,70,71,72
7	elefante	73,74,75,76,77,78,79,80,81,82,83
8	pecora	86,87,88,89,90,91,92,93,94
9	mucca	95,96,97,98,99,100,101,102,103
10	cavallo	104,105,106,107,108,109,110,111,112
11	gallina	119,120,121,122,123,124,125,126,127
12	gallo	128,129,130,131,132,133,134,135,136,137,138
13	oca	141,142,143,144,145,146,147,148,149,150,151,152,153
14	pappagallo	154,155,156,157,158,159,160,161,162,163,164,165
15	piccione	166,167,168,169,170,171,172,173,174,175,176
16	gufo	177,178,179,180,181,182,183,184,185,186

Tabella 2 – Nella prima colonna si riportano le posizioni associate al singolo concetto; nella seconda si riportano i nomi dei concetti della tassonomia. Nella terza colonna sono elencate le *features* distintive associate ai concetti, indicate come posizione assunta all'interno della rete. Il significato delle *features* è riportato nella Figura A1 dell'Appendice.

pos	concetto	features parzialmente condivise
1	cane	187,188,196,204,206,207,208
2	gatto	187,188,194,196,197,198,202,205,206,207,208,210,211,222
3	orso	189,190,193,195,199,200,204,208,209,212,229,231
4	leone	189,191,193,200,202,205,211,213,214,223,232
5	giraffa	189,191,195,197,208,209,210,215,216,217,229
6	zebra	203,214,218,219,223,229,232
7	elefante	192,193,195,199,208,212,229,230
8	pecora	190,196,198,201,207,208,219,224,227,228
9	mucca	189,193,196,201,206,208,211,213,215,216,224,227,228
10	cavallo	187,189,190,201,203,206,211,214,217,218,219,223,228
11	gallina	190,201,210,220,225,226,228
12	gallo	201,220,225,228
13	oca	190,201,226,228
14	pappagallo	188,221,226,229
15	piccione	190,192,194,210,230
16	gufo	189,192,205,207,210,221,222,231

Tabella 3 – Nella prima colonna si riportano le posizioni associate al singolo concetto; nella seconda si riportano i nomi dei concetti della tassonomia. Nella terza colonna sono elencate le *features* parzialmente condivise associate ai concetti, indicate come posizione assunta all'interno della rete. Il significato delle *features* è riportato nella Figura A1 dell'Appendice.

1.3.2 Tassonomia di Oggetti

Questa seconda tassonomia è costituita da 12 oggetti e ciascuno è mediamente caratterizzato da un numero di *features* compreso tra 15 e 20. In particolare, la tassonomia è realizzata da un totale di 165 caratteristiche di cui: 3 sono totalmente condivise (realizzano le categorie) e 28 sono parzialmente condivise (entrambe riportate nella Tabella 4) mentre 134 sono distintive (riportate nella Tabella 5). In questa tassonomia non sono presenti molte *features* totalmente condivise per cui non è possibile una suddivisione in molteplici categorie, come invece accadeva per la tassonomia degli animali. In questo caso, a partire dalla super categoria “OGGETTO”, identificata appunto dall’unico attributo condiviso da tutti “è un oggetto”, si possono distinguere solamente due modeste sotto categorie relative agli oggetti di “ARREDO” e agli “UTENSILI DA CUCINA”. Alla prima appartengono i concetti di “tavolo”, “divano” e “libreria” che condividono la *feature* “è un mobile” mentre alla seconda appartengono i concetti di “pentola”, “forchetta” e “tazzina” contraddistinti dall’attributo “si trova in cucina”. I rimanenti sei concetti di questa tassonomia (ovvero “bottiglia”, “lampada”, “sveglia”, “pistola”, “martello” e “ombrello”) non creano nessuna ulteriore o particolare categoria.

La tassonomia di oggetti, se confrontata con quella degli animali descritta in precedenza, mostra una minore vicinanza semantica tra i concetti che la compongono. Di conseguenza, avendo meno proprietà condivise tra i vari concetti, si è riusciti ad avere una suddivisione categoriale piuttosto limitata.

Ancora una volta, per approfondire e completare la trattazione, si può fare riferimento alla Figura A2, riportata nella sezione conclusiva Appendice, che presenta lo schema completo della tassonomia degli oggetti, con l’indicazione di tutte le *features* utilizzate.

pos	concetto	f. di categoria	features parzialmente condivise
1	tavolo	2,3	139,140,141,142,146,147,152,161
2	divano	2,3	142,145,148
3	libreria	2,3	139,143,147,149,163
4	pentola	2,4	140,143,150,154,158,159,163,165
5	forchetta	2,4	147,149,150,155,159,161,165
6	tazzina	2,4	139,141,144,146,152,154,163,164,165
7	bottiglia	2	146,151,153,155,164
8	lampada	2	139,144,145,151,158
9	sveglia	2	139,144,146,149,158,162
10	pistola	2	149,156,157,160,165
11	ombrello	2	145,146,148,159,162,165,166
12	martello	2	156,157,159,160,162,166

Tabella 4 – Nella prima colonna si riportano le posizioni associate al singolo concetto; nella seconda si riportano i nomi dei concetti della tassonomia. Nella terza colonna sono elencate le features condivise costituenti le categorie mentre nella quarta le features parzialmente condivise, indicate come posizione assunta all'interno della rete. Il significato delle features è riportato nella Figura A2 dell'Appendice.

pos	concetto	features distintive
1	tavolo	5,6,7,8,9,10,11,12,13
2	divano	14,15,16,17,18,19,20,21,22,23,24,25,26,27
3	libreria	28,29,30,31,32,33,34,35,36,37
4	pentola	38,39,40,41,42,43,44,45,46
5	forchetta	47,48,49,50,51,52,53,54,55
6	tazzina	56,57,58,59,60,61,62,63,64,65,66
7	bottiglia	67,68,69,70,71,72,73,74,75
8	lampada	76,77,78,79,80,81,82,83,84,85,86,87
9	sveglia	88,89,90,91,92,93,94,95,96,97,98,99,100
10	pistola	101,102,103,104,105,106,107,108,109,110,111,112,113,114,115
11	ombrello	116,117,118,119,120,121,122,123,124,125,126,127,128
12	martello	129,130,131,132,133,134,135,136,137,138

Tabella 5 – Nella prima colonna si riportano le posizioni associate al singolo concetto; nella seconda si riportano i nomi dei concetti della tassonomia. Nella terza colonna sono elencate le features distintive associate ai concetti, indicate come posizione assunta all'interno della rete. Il significato delle features è riportato nella Figura A2 dell'Appendice.

Capitolo 2

MODELLO DI MEMORIA SEMANTICA

2.1 Modello della rete semantica

Il ruolo di diversi tipi di caratteristiche nell'ambito del riconoscimento dei concetti può trarre vantaggio dall'uso di una classe di modelli matematici connessionisti, chiamati "reti attrattive". In queste reti, un concetto viene rappresentato come un insieme distribuito di nodi attivi (che simula la codifica dei gruppi neurali per le *features*) mentre le regolarità statistiche del verificarsi di tali caratteristiche vengono codificate attraverso connessioni tra i nodi (cioè le sinapsi), realizzando una rete auto-associativa. Le reti auto-associative hanno la peculiarità di essere essenzialmente costituite da un solo strato neuronale, in cui tutti i neuroni fungono sia da input che da output ed ognuno alimenta sé stesso oltre che tutti gli altri ad esso vicini. Ogni episodio viene memorizzato come punto di equilibrio stabile, con un bacino di attrazione circostante, e può essere poi evocato sulla base del contenuto, dando cioè in pasto alla rete (a seguito della fase di addestramento) un episodio simile a quello memorizzato [8].

Il modello trattato in questo lavoro di tesi è una versione monodimensionale e semplificata di un modello più complesso a due dimensioni, sviluppato in recenti studi dal Prof. Ursino (2018) [9]. In particolare, il modello da me analizzato è costituito da una rete di neuroni rappresentativa della rete semantica dedicata alla descrizione di concetti, i quali sono considerati come collezione di *features* sensoriali-motorie. All'interno di questa rete (schema in Figura 1) i concetti vengono rappresentati come un insieme di proprietà raccolte in un vettore di lunghezza M (numero di attributi utilizzati). Ogni unità neurale codifica per una proprietà e riceve sinapsi da tutte le altre (ma non da sé stessa) formando una matrice di sinapsi di dimensione $M \times M$. Le sinapsi tra neuroni (unicamente eccitatorie nel nostro modello) vengono create sulla base dell'esperienza passata della rappresentazione del concetto e sono addestrate con il paradigma hebbiano.

Sono stati realizzati due modelli distinti, sulla base delle due tassonomie di animali e oggetti descritte nel capitolo precedente, le quali vengono descritte attraverso un vettore di *features*; alcune proprietà sono comuni così da realizzare semplici categorie mentre altre sono proprietà distintive. La rete addestrata può risolvere semplici compiti di riconoscimento concettuale di animali/oggetti mantenendo una distinzione tra categorie e suoi membri e assegnando un ruolo differente alle caratteristiche dominanti rispetto a quelle marginali [9].

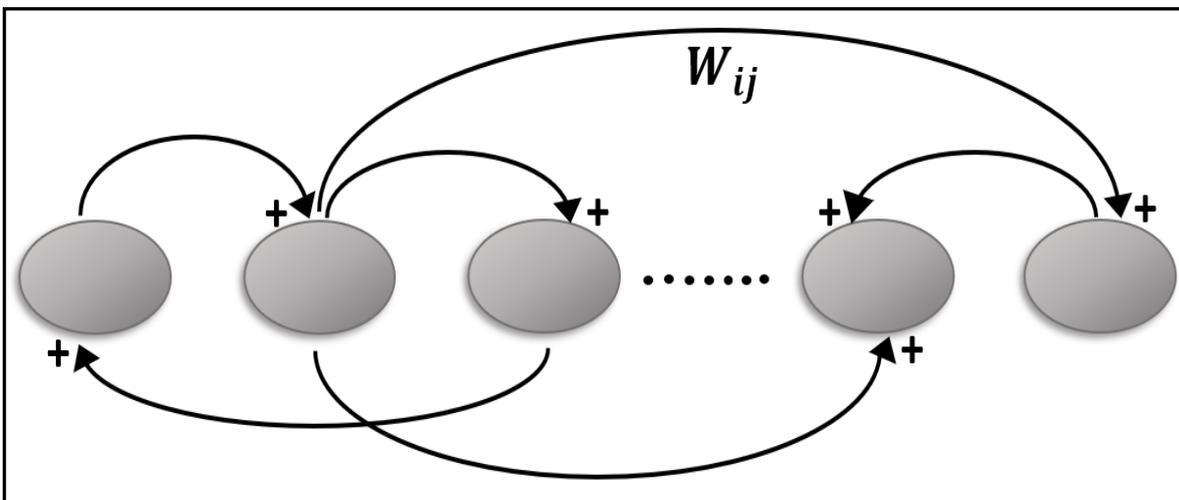


Figura 1 – Rappresentazione schematica del modello di rete semantica.

2.2 Formulazione matematica

Ogni unità nella rete semantica riceve un input globale che verrà di seguito definito u e produce un'attività di output indicata come x . La relazione ingresso-uscita può essere descritta attraverso: l'espressione di un filtro passa-basso del primo ordine con costante di tempo τ (Equazione 1), che tiene conto delle proprietà integrative del neurone ed allo stesso tempo riproduce il comportamento della risposta nel tempo, e da una funzione sigmoidale (Equazione 2), la quale rappresenta una soglia (inferiore e superiore) di saturazione dei neuroni. La relazione ingresso-uscita risulta:

$$(1) \quad \tau \frac{d}{dt} x_i(t) = -x_i(t) + H(u_i(t))$$

dove τ è la costante di tempo che determina la velocità della risposta allo stimolo mentre $H(u_i(t))$ è una funzione di attivazione sigmoideale che posso descrivere con la seguente espressione:

$$(2) \quad H(u_i(t)) = \frac{1}{1+e^{-(u_i(t)-\varphi)/T}}$$

dove T è un parametro che determina la pendenza centrale del sigmoide mentre il parametro φ stabilisce la posizione centrale della sigmoide. Dall'espressione della funzione sigmoideale deriva che il valore di saturazione superiore è pari ad uno; questo indica che tutte le attivazioni neurali sono normalizzate al valore massimo. L'ingresso alla rete semantica, u_i , può essere espresso come somma di due diversi contributi (Equazione 3):

$$(3) \quad u_i(t) = I_i(t) + E_i(t)$$

dove I_i rappresenta l'ingresso esterno per l'unità neurale situata in posizione i -esima, proveniente da una catena di elaborazione sensoriale-motoria che estrae le caratteristiche. Tale ingresso assume il valore zero nel caso di assenza di determinate *features* oppure il valore unitario in caso di presenza di proprietà. E_i identifica invece un termine di accoppiamento eccitatorio proveniente dalle unità neurali che codificano una differente proprietà e che quindi si trovano in un'altra posizione della rete semantica. L'espressione del termine di accoppiamento eccitatorio è (Equazione 4):

$$(4) \quad E_i = \sum_j W_{ij} x_j$$

dove il pedice i denota la posizione del neurone post-sinaptico (cioè il neurone bersaglio) mentre il pedice j indica la posizione del neurone pre-sinaptico. La sommatoria si estende a tutti i neuroni pre-sinaptici della rete semantica e W_{ij} rappresentano le sinapsi inter-area eccitatorie, le quali sono soggette ad apprendimento. Considero infine che le *features* che non si verificano insieme risultano semplicemente non connesse, perciò si è esclusa la presenza di sinapsi inibitorie; inoltre non si prevedono auto-anelli eccitatori di una *features* verso sé stessa (i sempre diverso da j) [9,10].

2.3 Addestramento delle sinapsi

2.3.1 Apprendimento con la regola di Hebb

Un aspetto fondamentale per la costruzione di una rete neurale è sicuramente l'utilizzo di una regola di apprendimento che permetta di modificare le connessioni sinaptiche sulla base dell'esperienza. I comportamenti utili dal punto di vista cognitivo emergono proprio in virtù della plasticità sinaptica e del suo rapporto con l'esperienza. Per questa ragione, prima di descrivere dettagliatamente la procedura di addestramento, è opportuno definire il concetto alla base della regola di apprendimento hebbiano.

Donald Hebb nel 1949 formula un importante principio, in riferimento all'addestramento sinaptico, il quale stabilisce che una sinapsi si rinforza con l'uso affermando quanto segue:

“quando un neurone A contribuisce ripetutamente e in maniera duratura alla eccitazione di un neurone B, allora la forza dell'azione del neurone A sul neurone B aumenta”

È facile capire come la regola di Hebb implichi l'idea di connessionismo, teoria secondo la quale il cervello lavora connettendo tra loro stati mentali. Di conseguenza la percezione di un oggetto potrebbe essere spiegata come attività contemporanea di gruppi di neuroni interconnessi, ognuno dei quali risponde ad una proprietà dell'oggetto. Ovvero, l'eccitazione di un determinato gruppo neuronale corrisponde alla rappresentazione mentale di quella specifica proprietà. Hebb, inoltre, discute la possibilità che la memoria possa essere realizzata in maniera concreta dalle connessioni inter-neurali e che esse si formino grazie all'esperienza, cioè alla possibilità di variare i pesi sinaptici; in questo modo si può dunque spiegare l'apprendimento [8].

La forma più semplice della regola di Hebb prevede che la sinapsi si aggiorni sulla base della correlazione tra attività pre-sinaptica e attività post-sinaptica.

Indicando con ΔW_{ij} la variazione delle sinapsi, e con γ un coefficiente di apprendimento (positivo), maggiore sarà γ , più velocemente verrà appresa la variazione della sinapsi (Equazione 5). La regola di Hebb assume la forma seguente:

$$(5) \quad \Delta W_{ij} = \gamma X_j X_i$$

dove X_j identifica l'attività del neurone pre-sinaptico mentre X_i quella del neurone post-sinaptico (per semplicità si assume attività pari a zero o uno).

- Se entrambi i neuroni sono inibiti non si ottiene variazione sinaptica;
- Se entrambi sono eccitati la sinapsi da j ad i cresce di un valore γ .
- Se uno dei due neuroni è inattivo la variazione della sinapsi risulta nulla.

Questa regola, nella sua formulazione originaria presenta un forte limite: prevede unicamente il rinforzo della sinapsi. Se così fosse, si assisterebbe sempre e solo ad un aumento delle sinapsi eccitatorie, con il risultato di una eccessiva connettività cerebrale a causa di una smisurata eccitazione. Per ottenere un addestramento efficiente che implichi una maggiore valenza biologica, si deve quindi prevedere la possibilità di indebolimento della sinapsi nel caso in cui essa non venga utilizzata [8].

Si modifica dunque la regola di Hebb (Equazione 5) immaginando che il rinforzo o l'indebolimento avvengano confrontando l'uscita del neurone con un valore di soglia (nel caso di utilizzo classico della regola). La regola di Hebb modificata (Equazione 6) si può esprimere come segue:

$$(6) \quad \Delta W_{ij} = \gamma_{ij} (X_i - \vartheta_{post})(X_j - \vartheta_{pre})$$

dove valori uguali per le soglie ϑ_{post} e ϑ_{pre} implicano la formazione di sinapsi simmetriche. Tali sinapsi simmetriche risultano tuttavia inadeguate per relazionare tra loro concetti in un modello di memoria semantica. Per questo motivo per il nostro modello si è reso necessario utilizzare valori diversi per le soglie pre e post-sinaptiche, così da creare sinapsi asimmetriche che siano in grado di discriminare tra proprietà salienti e marginali e ancora tra caratteristiche condivise e distintive [9].

2.3.2 Procedura di addestramento

Prima di procedere con l'addestramento della rete, tutte le sinapsi eccitatorie vanno poste a zero. La procedura di addestramento prevede che i concetti, descritti tramite *features*, vengano presentati singolarmente alla rete: ciò corrisponde ad una fase in cui il soggetto sperimenta, fa esperienza con i concetti e apprende la loro semantica. Inoltre, non tutte le caratteristiche sono sempre presenti in ingresso, bensì hanno una diversa frequenza di occorrenza (cioè una diversa probabilità di verificarsi); in particolare un'alta frequenza rende salienti le *features*. Si consideri la caratteristica n -esima e si denoti con P_{kn} la probabilità che tale proprietà venga presentata come input per il k -esimo concetto. Nel nostro addestramento le *features* salienti sono contraddistinte da una probabilità di ricevere in ingresso un input elevato (sufficiente per portare il neurone corrispondente a saturazione, cioè la proprietà viene percepita) pari a $P_{kn(s)}$, mentre le proprietà marginali hanno una probabilità pari a $P_{kn(m)}$. La probabilità di ricevere in ingresso un input basso sarà invece data dall'espressione $(1 - P_{kn})$, la quale risulta insufficiente per l'eccitazione del corrispondente neurone (cioè la *feature* non viene percepita).

Nello specifico, alle caratteristiche salienti è stata associata un'alta probabilità pari al 70% (cioè $P_{kn(s)} = 0.7$, uguale a quella utilizzata nei precedenti lavori) mentre alle proprietà marginali è stata associata una bassa probabilità pari 35% (cioè $P_{kn(m)} = 0.35$), più bassa rispetto ai lavori precedenti in cui era stata fissata pari al 40%, in modo da rendere più forte la distinzione tra proprietà salienti e marginali [10].

L'addestramento semantico complessivo consiste in mille epoche consecutive: tutti i concetti sono stati dati in pasto alla rete separatamente e permutati in ordine casuale (ogni concetto è presentato una volta in ogni epoca). Le sinapsi eccitatorie all'interno della rete semantica sono state addestrate con il meccanismo hebbiano, modificando il peso sinaptico sulla base della correlazione tra attività pre e post-sinaptica.

Per tenere conto non solo del potenziamento a lungo termine ma anche del depotenziamento, si considera il confronto di queste attività con un valore di soglia. Un basso livello di attività in un neurone provoca un depotenziamento della forza sinaptica se accompagnato da un alto livello di attività nell'altro neurone (questo avviene nel caso in cui le attività sono negativamente correlate). Viceversa, quando entrambi i neuroni hanno un'attività elevata, la sinapsi si rafforza.

L'Equazione 6 richiede tuttavia alcune restrizioni per essere realistica da un punto di vista fisiologico. Per prima cosa, il fattore di apprendimento deve diminuire quando la sinapsi si avvicina a un livello massimo di saturazione (W_{max}), cioè una sinapsi non può crescere indefinitamente (Equazione 7):

$$(7) \quad \gamma_{ij} = \frac{\gamma_0}{W_{max}} (W_{max} - W_{ij})$$

dove γ_0 è il tasso di apprendimento quando la sinapsi è nulla.

In secondo luogo, quando entrambe le attività pre e post-sinaptiche sono basse, non dovrebbero verificarsi cambiamenti di peso (Equazione 8). Quindi:

$$(8) \quad \text{se } x_i < \vartheta_{post} \text{ e } x_j < \vartheta_{pre} \text{ allora } \Delta W_{ij} = 0$$

Infine, una sinapsi non può diventare negativa altrimenti l'eccitazione verrebbe convertita in inibizione (non è fisiologicamente accettabile). Quando si calcola il nuovo valore della sinapsi si avrà (Equazione 9):

$$(9) \quad W_{ij} \leftarrow \text{Max}\{W_{ij} + \Delta W_{ij}, 0\}$$

2.3.3 Requisiti della rete semantica

Al fine di un corretto processo di addestramento occorre tenere presente alcuni requisiti che la rete deve necessariamente soddisfare [9,11].

- *Features* salienti devono essere evocate da tutte le altre proprietà (salienti e marginali) del concetto. Devono ricevere forti sinapsi in input ma inviare forti sinapsi di output solo alle altre caratteristiche salienti.

- *Features* marginali non devono essere evocate dalle altre proprietà del concetto quindi devono ricevere sinapsi in input molto deboli. Poiché devono però favorire la ricostruzione del concetto, esse devono inviare forti sinapsi di output solo verso le caratteristiche salienti.
- *Features* distintive devono richiamare non solo altri attributi distintivi salienti dello stesso concetto ma anche caratteristiche condivise (“abbaia” deve richiamare tutti gli attributi di “cane”, tra cui “ha il pelo corto”, “ha le orecchie” che sono condivise con altri animali).
- *Features* condivise da concetti diversi in una categoria devono richiamare le altre caratteristiche totalmente condivise nella stessa categoria ma non devono richiamare quelle distintive dei singoli membri, né quelle parzialmente condivise (“ha quattro zampe” richiama attributi comuni di categoria come “ha il pelo”, “ha la coda”, “è un mammifero” ma non attributi dei singoli membri come “abbaia” o “miagola”).

2.3.4 Problema della soglia post-sinaptica

Il nostro modello di rete semantica ha come obiettivo di base quello di riuscire ad avere una chiara distinzione sul ruolo delle caratteristiche condivise e distintive e sul ruolo delle caratteristiche salienti e marginali. Per raggiungere questo intento le sinapsi devono essere fortemente asimmetriche: ciò si realizza considerando valori diversi di soglie per le attività pre e post-sinaptica.

In particolare, il comportamento desiderato delle sinapsi, descritto nel paragrafo precedente, può essere ottenuto considerando un valore elevato di soglia per l'attività post-sinaptica (ϑ_{post}) e un valore basso di soglia (cioè vicino all'inibizione) per l'attività pre-sinaptica ϑ_{pre} ; quest'ultimo è stato inserito in modo da evitare un rinforzo sinaptico indesiderato dovuto ad un'attività neuronale residua. Quindi:

- Se entrambi i neuroni, pre-sinaptico e post-sinaptico, sono attivi allora le sinapsi si rafforzano;

- ii. Se il neurone post-sinaptico è inibito mentre il neurone pre-sinaptico è eccitato allora la sinapsi si indebolisce;
- iii. Se il neurone post-sinaptico è eccitato mentre il neurone pre-sinaptico è inibito allora la sinapsi mostra solo un moderato indebolimento.

Nel nostro modello, l'unico aspetto che determina la salienza di una *feature* è la frequenza di occorrenza. Studi precedenti hanno dimostrato che, utilizzando la regola di Hebb (Equazione 9), una caratteristica diventa saliente se la sua probabilità P_{kn} è superiore alla soglia post-sinaptica ϑ_{post} [10]. Se ciò si verifica infatti, il potenziamento supera il depotenziamento e le sinapsi entranti nella *feature* raggiungono il valore di saturazione (dopo un numero sufficiente di epoche). Viceversa, se la frequenza di un attributo è inferiore alla soglia post-sinaptica, il depotenziamento supera il potenziamento e le sinapsi entranti tendono ad annullarsi (si tratta di una *feature* che non può essere evocata spontaneamente ed è quindi marginale).

In conclusione, si può affermare che il livello di salienza è strettamente correlato al valore basale utilizzato per la soglia post-sinaptica ϑ_{post} , in particolare: maggiore è questa soglia, maggiore è il livello di frequenza richiesto per avere salienza. Nel nostro caso specifico ϑ_{post} è pari a 0.67, questo significa che una caratteristica deve essere percepita in più del 67 % delle presentazioni per diventare saliente.

Utilizzando la regola di Hebb con diversi valori per le soglie pre e post-sinaptiche si riesce bene a discriminare tra caratteristiche marginali e salienti, e tra distintive e totalmente condivise, in base alla loro frequenza di occorrenza. Il comportamento della rete, tuttavia, risulta inadeguato per affrontare la salienza di una caratteristica parzialmente condivisa [9,10].

A titolo di esempio si considerino le due *features* “vola” e “ha le piume” per gli uccelli nella tassonomia di animali: il primo attributo è parzialmente condiviso (4 su 6) mentre il secondo è totalmente condiviso (6 su 6).

Se la soglia post-sinaptica resta fissa, la *feature* “ha le piume” diventa saliente per tutti e sei gli uccelli e, dopo l’addestramento, viene sempre evocata. L’attributo “vola” invece non è saliente per “gallina” e “gallo”, bensì si verifica con alta frequenza (70%) solo per gli uccelli volatili (“oca”, “pappagallo”, “piccione” e “gufo”). Si osservi ora la connessione tra le due *feature* “vola” e “ha le piume”. A seguito dell’addestramento, la proprietà “ha le piume” si verifica nel 100% degli uccelli poiché saliente per tutti, di contro l’attributo “vola” si verifica solamente nel 66% degli uccelli. Utilizzando quindi una soglia ϑ_{post} fissa, si crea progressivamente una sinapsi non solo da “vola” verso “ha le piume” ma anche una sinapsi da “ha le piume” verso “vola” e quest’ultima diventa saliente per tutti gli uccelli. Ciò significa che, all’interno della tassonomia di animali, qualsiasi animale che vola ha le piume (corretto), ma anche che qualsiasi animale che ha le piume è in grado di volare. È immediato notare che questa ultima affermazione è del tutto scorretta in quanto non regge né per il concetto di “gallo” né per quello di “gallina”.

Per risolvere questo conflitto, si assume allora una soglia post-sinaptica variabile che deve aumentare quando si collegano le *features* in una categoria, mentre deve rimanere bassa quando si collegano caratteristiche distintive dei singoli membri. In altre parole, si assume che la soglia cambi con l’esperienza, in particolare che essa aumenti con il numero di occorrenze della caratteristica pre-sinaptica [10]. In questo modo, non solo una *feature* condivisa in una categoria non evoca una proprietà distintiva di un singolo membro, ma anche un attributo totalmente condiviso (“ha le piume”) non evoca una caratteristica parzialmente condivisa (“vola”).

Si consideri dunque la sinapsi W_{ij} che collega la i -esima caratteristica post-sinaptica alla j -esima *feature* pre-sinaptica. Nel modello si è assunto che la soglia post-sinaptica dipenda dal numero di occorrenze della proprietà pre-sinaptica, quindi che risulti funzione di j e la regola Hebb si può esprimere come segue (Equazione 10):

$$(10) \quad \Delta W_{ij} = \gamma_{ij} (x_i - \vartheta_{(post,j)}) (x_j - \vartheta_{pre})$$

dove il pedice j nella soglia post-sinaptica specifica la sua dipendenza dalla caratteristica pre-sinaptica. Indicando con N_j il numero medio di occorrenze di una *feature* j -esima in un'epoca, a seguito dell'addestramento tale numero sarà più grande per le proprietà distintive salienti e ancora maggiore per quelle condivise. A questo punto l'espressione della soglia post-sinaptica variabile finale risulta (Equazione 11):

$$(11) \quad \vartheta_{post,j} = \begin{cases} \vartheta_{post-base} + \Delta\vartheta_{post}(N_j - 1) & \text{se } N_j > 1 \\ \vartheta_{post-base} & \text{se } N_j < 1 \end{cases}$$

dove $\vartheta_{post-base}$ e $\Delta\vartheta_{post}$ sono parametri che impostano rispettivamente il valore basale (assunto uguale a ϑ_{post}) e la velocità di incremento della soglia post-sinaptica. Infine, si suppone che la soglia post-sinaptica non possa mai superare un livello di saturazione massimo identificato da $\vartheta_{post-sat}$.

2.3.5 Assegnazione dei parametri

I parametri utilizzati per l'addestramento sinaptico sono elencati nella Tabella 6. Alla costante di tempo (τ) è stato assegnato un valore (pochi millisecondi) normalmente usato nelle equazioni di campo medio deterministico. Alla pendenza della sigmoide (T) è stato assegnato un valore molto basso per avere un transizione veloce da inibizione a saturazione mentre l'ascissa centrale della sigmoide (φ) assume un valore tale per cui l'attività dei neuroni sia trascurabile in presenza di input nullo. Si riportano inoltre i valori delle soglie pre-sinaptica e post-sinaptica (in particolare il valore basale, quello corrispondente alla saturazione e alla velocità di incremento) di cui è già stato commentato il significato. Infine, sono presenti anche il rateo di apprendimento γ_0 , il massimo valore sinaptico W_{max} e le frequenze di occorrenza scelte rispettivamente per le *features* salienti $P_{kn(s)}$ e marginali $P_{kn(m)}$.

Significato	Simbolo	Valore
Costante di tempo	τ	3 ms
Pendenza della sigmoide	T	0.01
Posizione della sigmoide	φ	0.55
Soglia pre-sinaptica	ϑ_{pre}	0.05
Soglia post-sinaptica (valore basale)	ϑ_{post}	0.67
Saturazione della soglia post-sinaptica	$\vartheta_{post-sat}$	0.95
Velocità di incremento della soglia post-sinaptica	$\Delta\vartheta_{post}$	0.5
Rateo di apprendimento	γ_0	0.02
Massima forza sinaptica	W_{max}	0.8
Probabilità associata a <i>features</i> salienti	$P_{kn(s)}$	0.7
Probabilità associata a <i>features</i> marginali	$P_{kn(m)}$	0.35

Tabella 6 – Parametri utilizzati per l'addestramento sinaptico.

2.4 Risultati

Di seguito si effettua l'analisi dei *pattern* di sinapsi ottenuti a seguito della fase di addestramento. I risultati ottenuti permetteranno di evidenziare il diverso ruolo delle *features* condivise rispetto a quelle distintive e di quelle salienti rispetto alle marginali. Verranno successivamente presentati i risultati di alcune simulazioni riguardanti una attività di riconoscimento di concetti. In questa sezione si prende in considerazione la rete semantica descritta in precedenza implementata con la tassonomia di animali; risultati del tutto analoghi valgono equivalentemente per la tassonomia di oggetti.

2.4.1 Risultati dell'addestramento sinaptico

Si ricorda che l'addestramento consiste in mille epoche dove vengono presentati singolarmente tutti gli animali in ordine random. I risultati ottenuti dopo l'addestramento sono riportati nei pannelli sottostanti, rappresentando la forza delle sinapsi entranti in una *feature* e provenienti da altre proprietà della rete semantica. In ascissa sono presenti le proprietà della tassonomia, riportate in termini di posizione assunta all'interno della rete, mentre in ordinata si trova il peso sinaptico (il valore massimo è W_{max}).

Occorre anche ricordare che le sinapsi possono essere solamente di tipo eccitatorio. Per commentare con chiarezza i risultati ottenuti si riportano, a titolo di esempio, i pannelli delle sinapsi di sole cinque diverse *features* della tassonomia. In particolare: una caratteristica totalmente condivisa che concorre a formare la super categoria “ANIMALE” (“mangia”), un attributo che partecipa alla costituzione della macro categoria “MAMMIFERO” (“ha il pelo”) e una caratteristica della sottocategoria “CARNIVORO” (“mangia la carne”). Infine, sono presenti le sinapsi relative a due *features* appartenenti al concetto di “leone”, nello specifico: una distintiva saliente (“ruggisce”) ed una distintiva marginale (“difende il territorio”).

Le sinapsi in input alla *feature* “mangia” (posizione 3) sono riportate in Figura 2. Tale caratteristica è saliente e comune a tutti gli animali della tassonomia (quindi è totalmente condivisa), concorre a formare la super categoria “ANIMALE”. Essa riceve quindi sinapsi da tutte le proprietà degli animali, tranne che da sé stessa; le sinapsi che provengono dalle caratteristiche salienti hanno una forza maggiore mentre quelle provenienti dalle *features* marginali sono più deboli perché capitano più raramente.

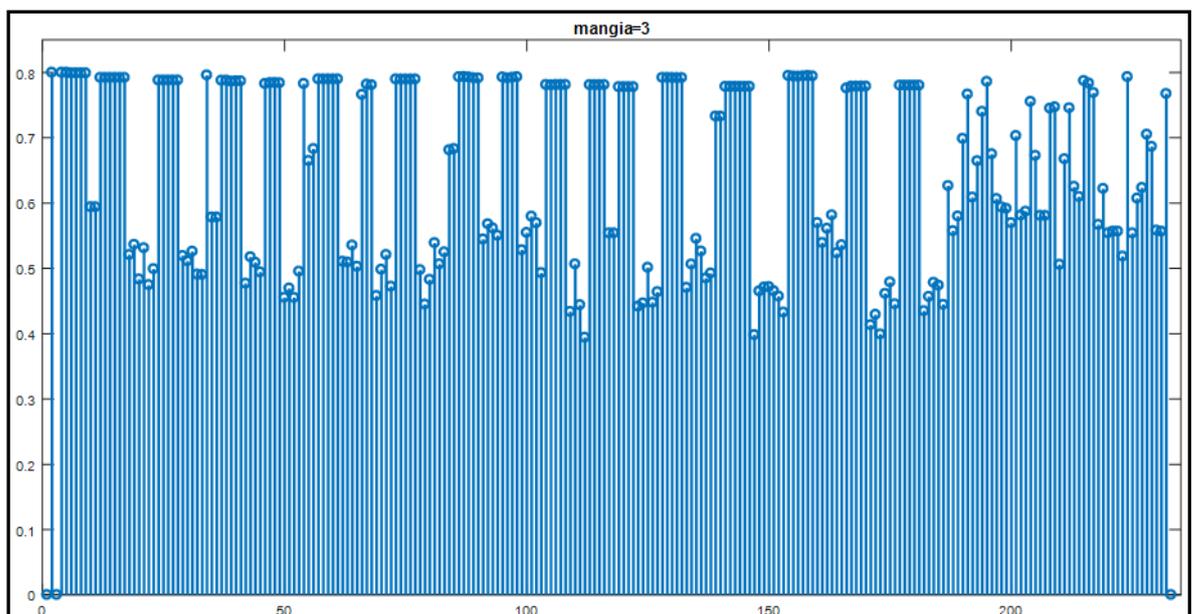


Figura 2 – Forza delle sinapsi entranti nella *feature* “mangia” (3).

Le sinapsi per la caratteristica “ha il pelo” si osservano in Figura 3. Questa *feature*, collocata in posizione 8 nella rete semantica, risulta saliente e comune solo a tutti i mammiferi determinando, insieme ad altre, la macro categoria “MAMMIFERO”. Le sinapsi entranti questa volta non provengono da tutte le proprietà degli animali bensì solo da quelle dei mammiferi e la forza sinaptica è sempre pesata dalla salienza. Non si hanno, ad esempio, sinapsi inviate dalle *features* degli “UCCELLI” (dalla 113 alla 186) e neppure dalle proprietà che appartengono alla categoria gerarchicamente superiore (cioè della super categoria “ANIMALE” nelle posizioni 2, 3, 4, 5) perché infatti non tutti gli animali che mangiano hanno il pelo.

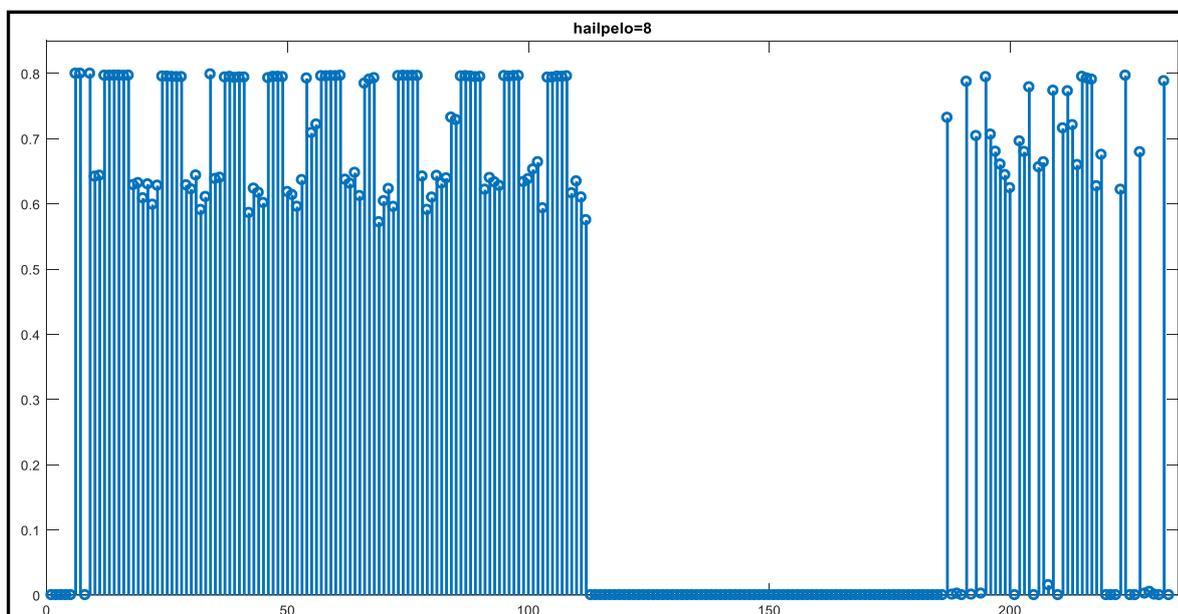


Figura 3 – Forza delle sinapsi entranti nella *feature* “ha il pelo” (8).

Le sinapsi entranti nella *feature* “mangia la carne” (posizione 36) sono mostrate in Figura 4 (nella pagina seguente). Tale proprietà è saliente ma comune ai soli carnivori (“leone” e “orso”) dalle cui proprietà riceve sinapsi pesate dalla salienza. Non riceve da attributi appartenenti a categorie gerarchicamente superiori (da 2 a 9 e 34); di contro riceve una sinapsi forte dalla proprietà “ha gli artigli” (in posizione 35) insieme alla quale costituisce la sotto categoria “CARNIVORO”.

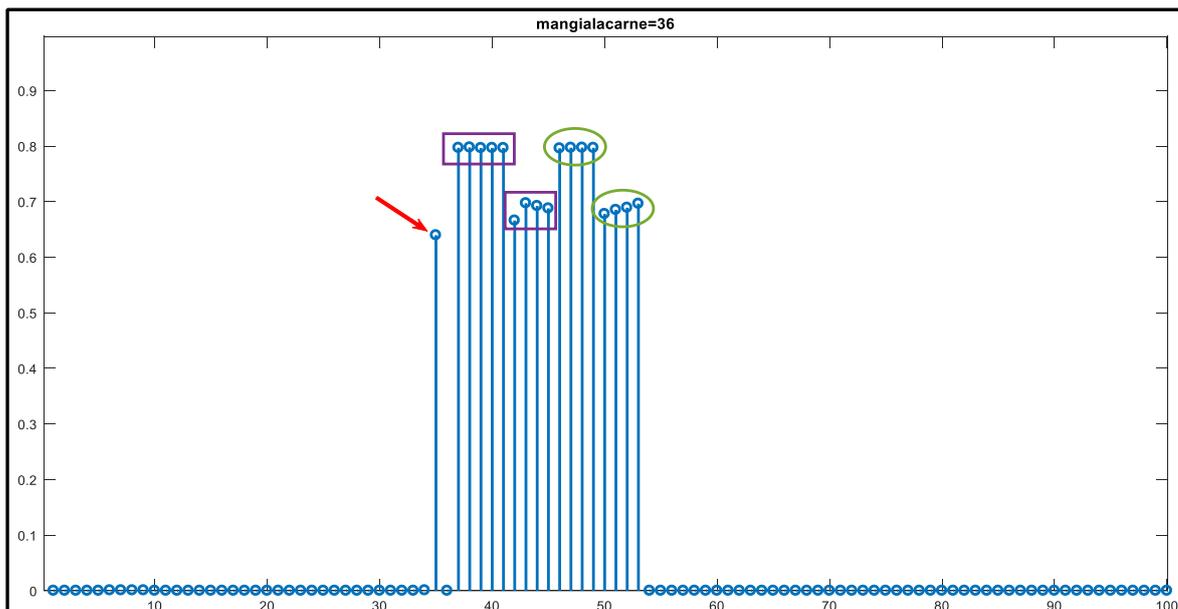


Figura 4 – Forza delle sinapsi entranti nella feature “mangia la carne” (36). I rettangoli in viola discriminano le sinapsi provenienti dalle features di “orso” mentre gli ellissoidi in verde identificano le sinapsi dalle features di “leone”. L’attributo 36 indicato dalla freccia in rosso concorre alla formazione della categoria “CARNIVORO”.

Si discutono ora le figure relative alle sinapsi di due caratteristiche appartenenti al singolo concetto di “leone”.

Come proprietà distintiva saliente si considera la *feature* “ruggisce” che occupa la posizione 46 nella rete semantica. Come evidente dalla figura nella pagina seguente (Figura 5), essa riceve sinapsi forti da tutte le altre caratteristiche distintive (salienti e marginali) dello stesso animale. Le sinapsi provenienti dagli altri attributi salienti di “leone” sono prossime alla saturazione massima. Nello specifico si individuano in ordine contiguo le *features*: “re della foresta” (47), “ha denti affilati” (48) e “ha la criniera folta” (49). Poco meno potenti sono invece quelle provenienti dalle proprietà marginali collocate nelle posizioni 50, 51, 52, 53 (rispettivamente “difende il territorio”, “è forte”, “simbolo di coraggio”, “è maestoso”). La *feature* distintiva saliente non riceve però sinapsi né da caratteristiche di altri animali, né da caratteristiche da lei parzialmente condivise (come ad esempio “è un felino” che condivide con “gatto”).

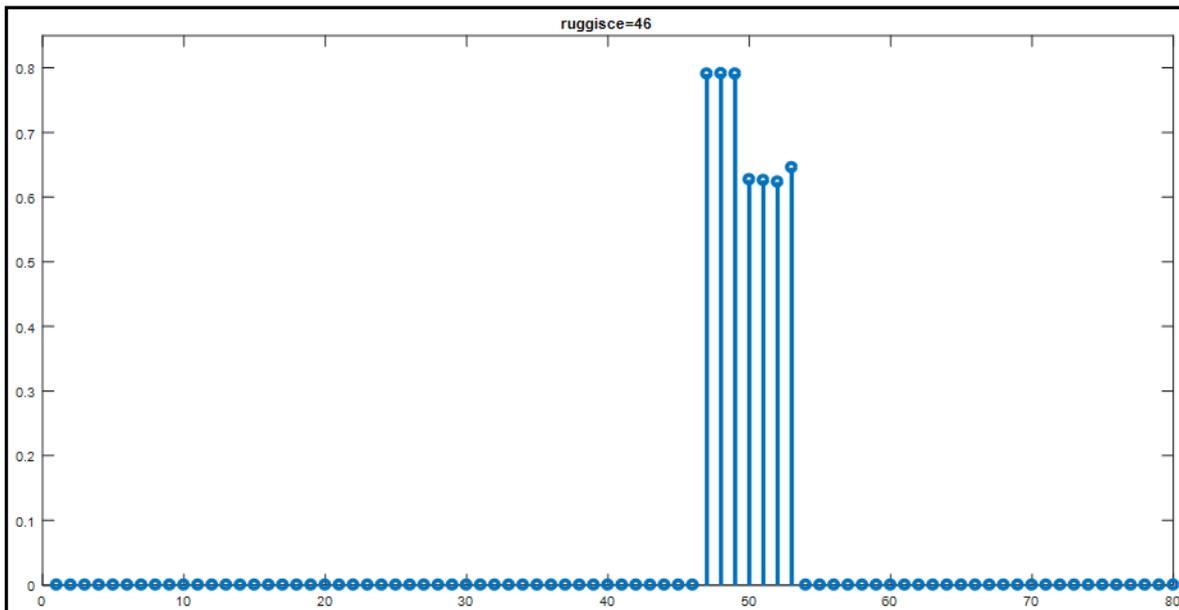


Figura 5 – Forza delle sinapsi entranti nella feature “ruggisce” (46).

Infine, per quanto riguarda la proprietà distintiva marginale “difende il territorio”, localizzata in posizione 50 all’interno della rete, si faccia riferimento alle sinapsi riportate nella Figura 6 sottostante. In questo caso si possono osservare debolissime sinapsi provenire dalle altre *features* distintive di “leone” ma, poiché praticamente nulle, sono insufficienti ad evocare tale attributo.

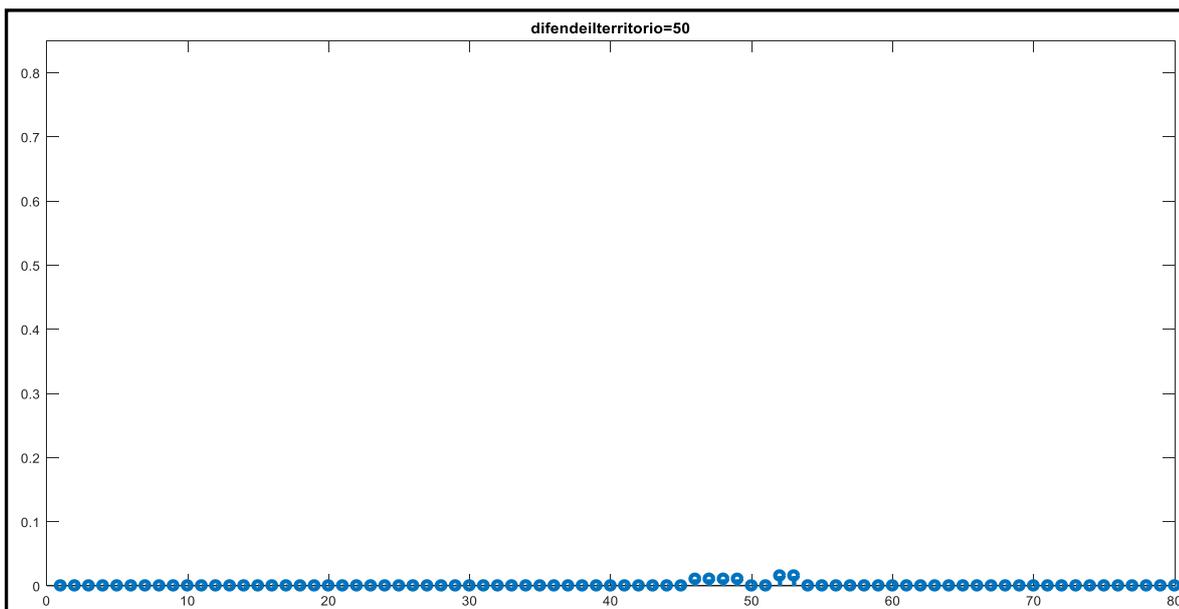


Figura 6 – Forza delle sinapsi entranti nella feature “difende il territorio” (50).

2.4.2 Risultati delle simulazioni

Dal momento che le sinapsi ottenute dall'addestramento riescono abbastanza ragionevolmente a discriminare tra *features* condivise e distintive e tra salienti e marginali, si è deciso di testare il modello con compiti di riconoscimento di concetti. Alla rete vengono fornite in input alcune *features* appartenenti ad un animale e si analizza il comportamento della rete verificando quali proprietà vengono evocate spontaneamente. Il compito di riconoscimento si considera ben svolto se il modello riesce ad evocare tutte le caratteristiche salienti del concetto di interesse.

Per osservare come evolve il comportamento della rete nel compito di riconoscimento si riportano pannelli che catturano diverse istantanee (rispettivamente agli *step* 10, 20, 40) della fase di simulazione. In ascissa sono presenti le 231 *features* della tassonomia (come posizione assunta all'interno della rete) e in ordinata vi è l'attività dei neuroni normalizzata. Occorre notare che, durante la simulazione, le proprietà richiamate in input tendono a raggiungere velocemente il livello massimo di attività neurale (già verso il ventesimo *step*) mentre le caratteristiche che devono essere evocate dalla rete giungono a saturazione circa dopo una quarantina di *step*.

Per commentare con chiarezza i risultati ottenuti si considera, a titolo di esempio, il compito di riconoscimento del concetto "leone", di cui sono già state riportate le sinapsi nel paragrafo precedente.

Se una caratteristica distintiva (saliente o marginale) è eccitata, la rete evoca tutte le proprietà salienti del concetto (senza evocarne nessuna marginale a parte quella eventualmente data come input). Quindi se l'input è la *feature* distintiva saliente "ruggisce" (46), il modello evoca correttamente tutte le proprietà salienti di "leone" (come si vede in Figura 7) in particolare: quelle totalmente condivise (dalla 2 alla 9 e dalla 34 alla 36), parzialmente condivise (193, 200, 202, 205, 213) e le distintive salienti (47, 48, 49).

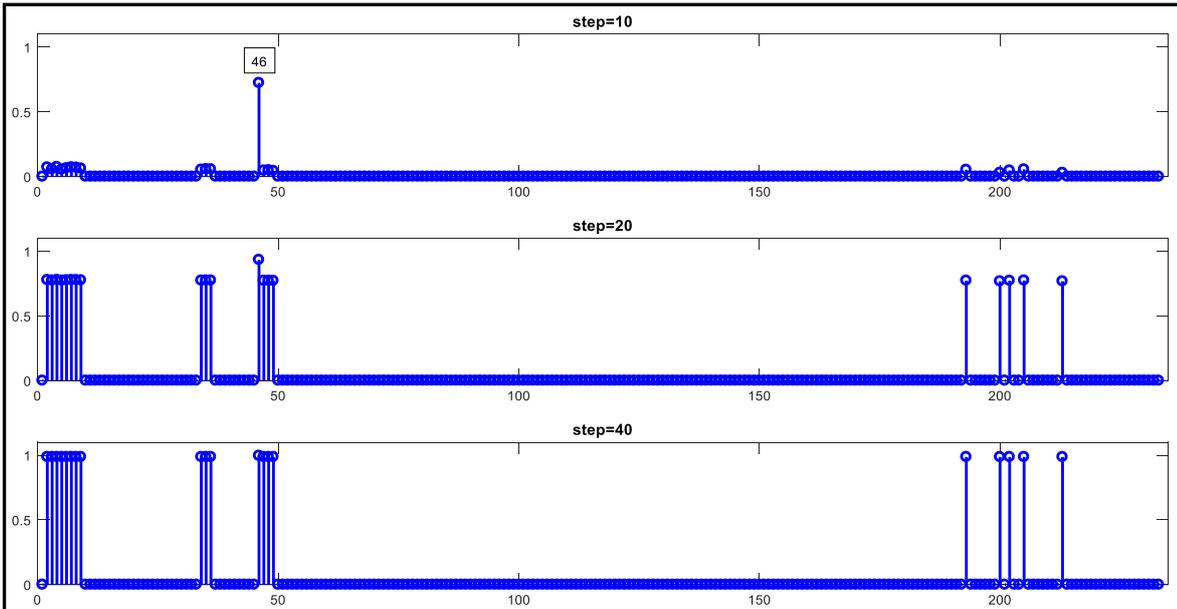


Figura 7 – Riconoscimento del concetto di “leone”, in input si considera la *feature* distintiva saliente “ruggisce” (46).

Il corretto riconoscimento del concetto “leone” avviene anche nel caso in cui l’input sia la *feature* distintiva marginale “è maestoso” (53); tale comportamento è riportato in Figura 8 e la rete riesce ad evocare tutte le caratteristiche salienti citate nell’esempio precedente.

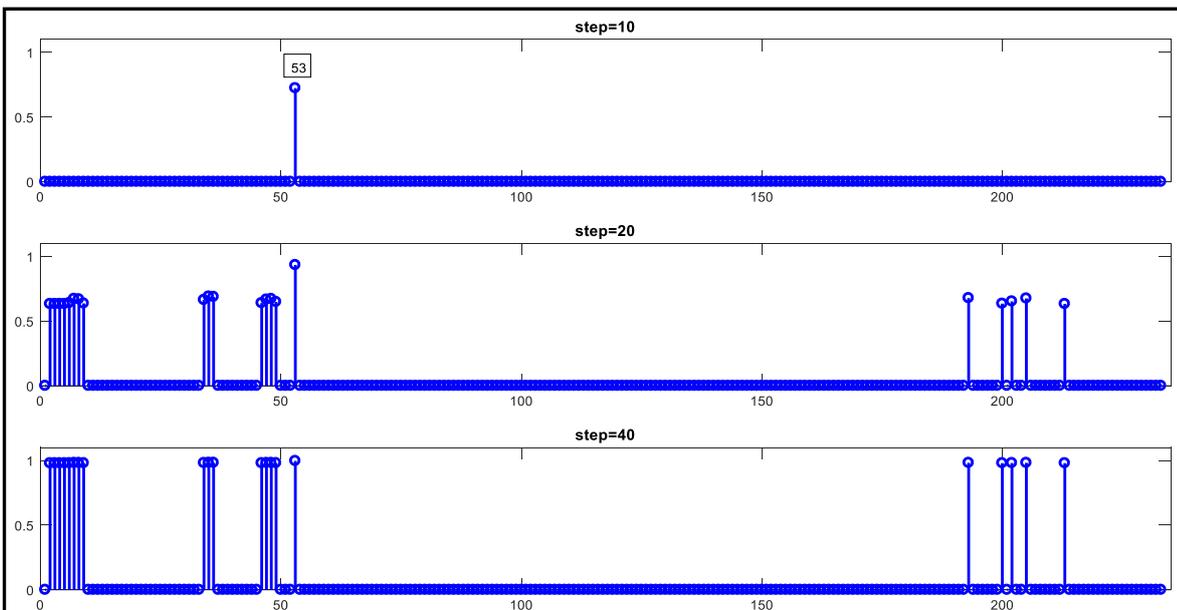


Figura 8 – Riconoscimento del concetto di “leone”, in input si considera la *feature* distintiva marginale “è maestoso” (53).

Il riconoscimento corretto si può individuare grazie al confronto degli ultimi due pannelli delle Figure 7 e 8 relativi al quarantesimo *step*, i quali differiscono solamente per la proprietà marginale 53, data in input.

Se si eccita una caratteristica totalmente condivisa (quindi che concorre a definire una data categoria), la rete evoca tutte le *features* salienti totalmente condivise di quella categoria e delle categorie gerarchicamente ad essa superiori. Proprietà condivise solo da alcuni membri (ma non da tutti) non vengono invece evocate.

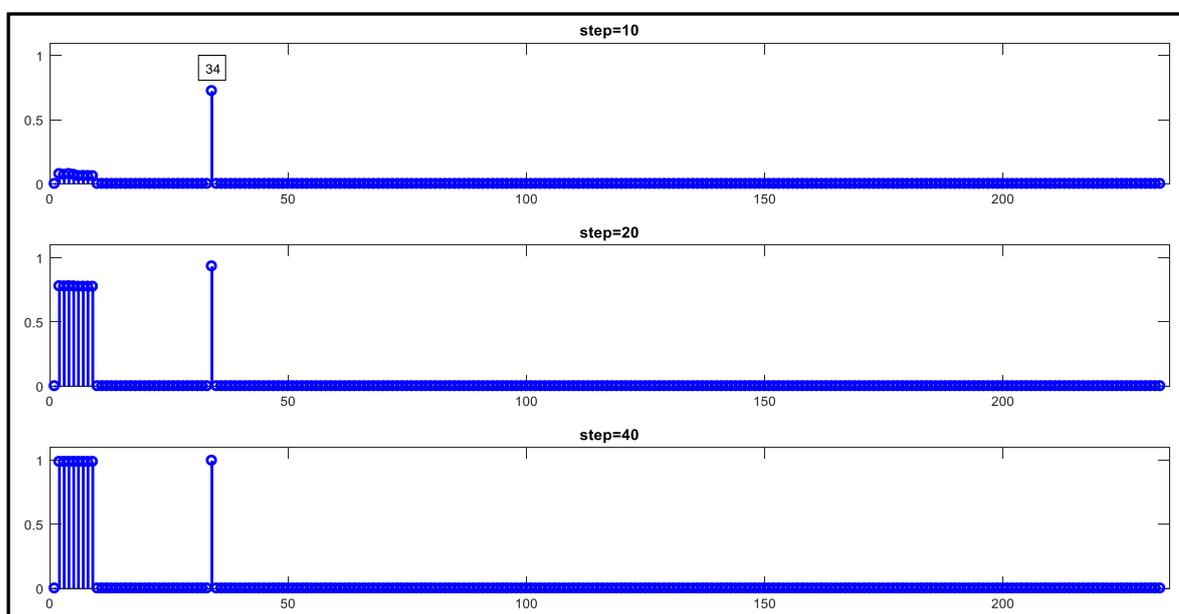


Figura 9 – Riconoscimento del concetto di animale selvatico, in input si considera la feature totalmente condivisa “è selvatico” (34).

Un esempio di questo caso (riportato in Figura 9) si presenta se in input è fornita la *feature* “è selvatico” (posizione 34 nella rete) che costituisce appunto la categoria di un animale “SELVATICO”. La rete evoca in maniera spontanea solamente tutte le proprietà totalmente condivise della macro categoria “MAMMIFERO” (cioè le *features* 6, 7, 8, 9) e della super categoria “ANIMALE” (ovvero le caratteristiche 2, 3, 4, 5). Infine, se una caratteristica parzialmente condivisa (saliente o marginale che sia) viene utilizzata come input, la rete evoca solamente le *features* totalmente condivise delle categorie che accomunano i concetti che condividono la proprietà comune.

Per chiarire, se si eccita la proprietà parzialmente condivisa “ha la coda con ciuffo” (213) del concetto di “leone” (in Figura 10), presente anche per il concetto di “mucca”, la rete evoca correttamente le caratteristiche totalmente condivise appartenenti alle categorie superiori di “ANIMALE” (2, 3, 4, 5), “MAMMIFEERO” (6, 7, 8, 9), e “SELVATICO” (34). Tuttavia, non riesce ad andare oltre nella discriminazione tra “leone” e “mucca”.

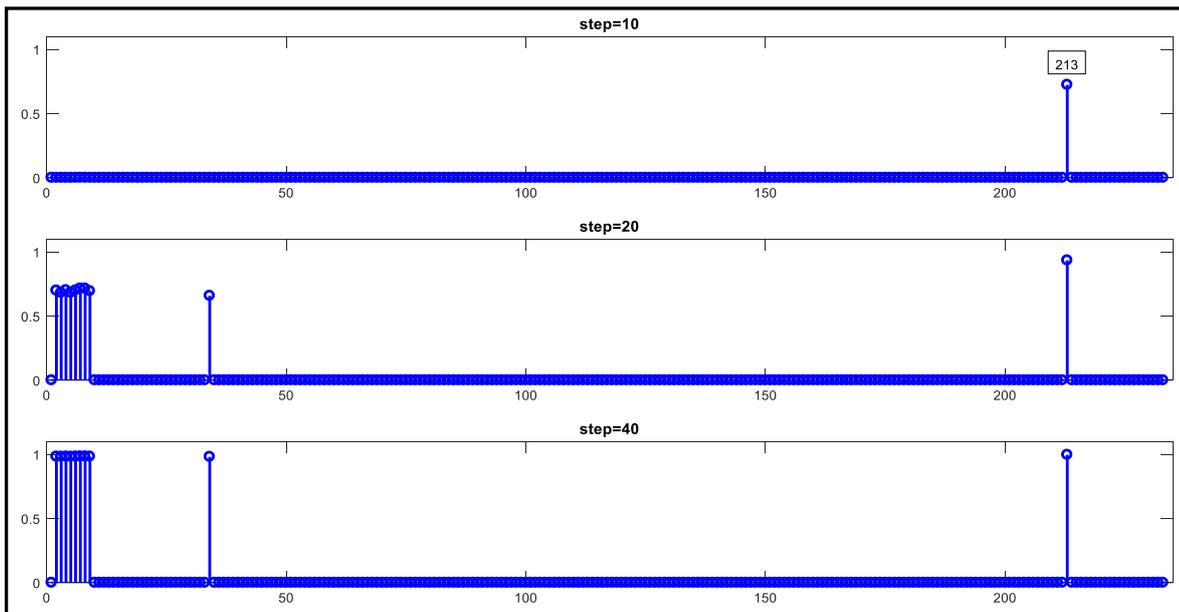


Figura 10 – Riconoscimento del concetto di “leone”, in input si considera la feature parzialmente condivisa “ha la coda con ciuffo” (213).

Si vuole di seguito riportare una simulazione di particolare interesse, la quale è in grado di mostrare il corretto funzionamento della rete reso possibile grazie all’implementazione di una soglia post-sinaptica variabile. Come già descritto in dettaglio nel paragrafo 2.3.4, l’utilizzo di una soglia post-sinaptica fissa risultava inadeguato per affrontare la salienza di una caratteristica parzialmente condivisa. Ad esempio, rendeva saliente la *feature* “vola” (139) (alta frequenza di occorrenza solo per quattro uccelli su sei) anche per i due uccelli non volatili “gallo” e “gallina”. La simulazione riportata in Figura 11 mostra come questo conflitto sia stato risolto grazie all’utilizzo della soglia post sinaptica che cambia con l’esperienza, la quale consente ora di operare il corretto riconoscimento del concetto senza evocare *features* errate.

Si nota infatti che fornendo in input alla rete una proprietà saliente del concetto di “gallina”, in particolare l’attributo “fa i pulcini” (120), la rete evoca spontaneamente tutte le *features* salienti del concetto di interesse senza che la proprietà “vola” venga erroneamente richiamata. Le *features* che emergono dal corretto riconoscimento risultano infatti quelle costituenti la super categoria “ANIMALE” (2, 3, 4, 5), la macro categoria “UCCELLO” (da 113 a 116) e ancora la categoria degli uccelli “NON VOLATILI” (117 e 118), oltre a tutte le caratteristiche salienti distintive (da 119 a 122) e parzialmente condivise (201, 210, 225, 226) di “gallina”.

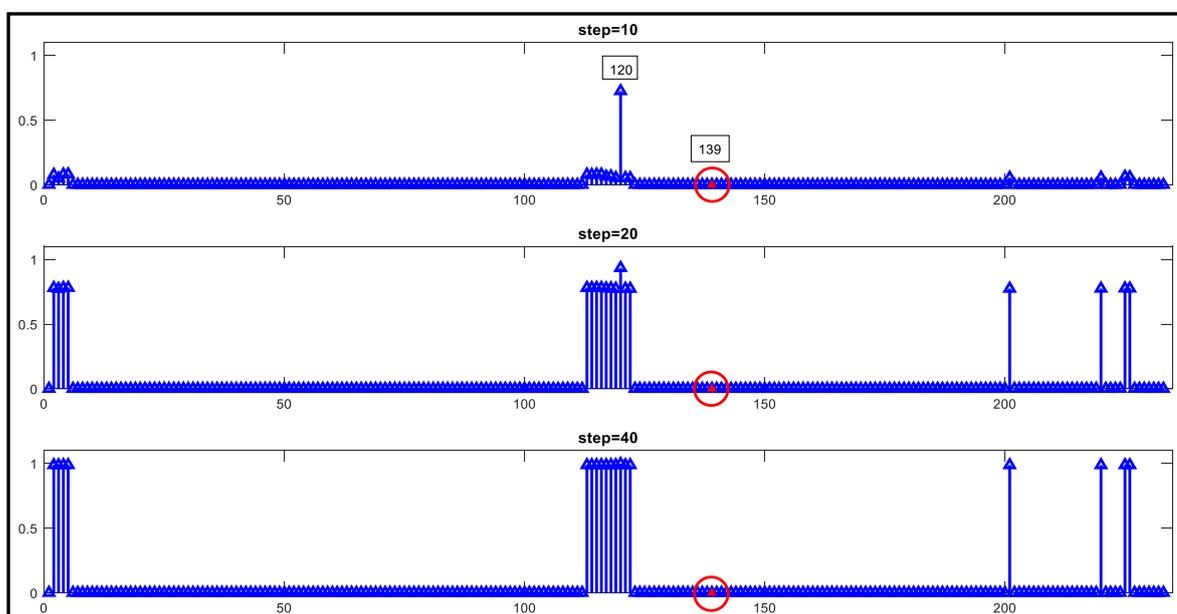


Figura 11 – Riconoscimento del concetto di “gallina”, in input si considera la feature distintiva saliente “fa i pulcini” (120). Non viene erroneamente evocata la feature “vola” (139).

Discussione

Le simulazioni per il compito di riconoscimento del concetto sono state ripetute per tutti i singoli animali e oggetti che realizzano le due tassonomie, in particolare variando la scelta delle *features* e il numero di attributi forniti in input e si è generalmente osservato un corretto funzionamento della rete. Occorre però precisare un limite della rete che si è osservato in un numero estremamente limitato di casi.

Nello specifico, è accaduto che una *feature* marginale (“si trova allo zoo”, collocata in posizione 229 nella rete) sia stata evocata spontaneamente come saliente durante il compito di riconoscimento del concetto (“zebra”). Questo fenomeno non costituisce un forte ostacolo al corretto funzionamento generale della rete, inoltre si pensa che sia dovuto principalmente ad un effetto di randomizzazione degli input.

Riassumendo, le prestazioni complessive del modello, in termini di rapporto “numero di successi” rispetto al “numero totale di prove eseguite”, possono essere sintetizzate come segue:

- i. Il riconoscimento dei singoli concetti (animali) avviene se sono stimulate le caratteristiche distintive: la rete riconosce l’animale di interesse senza evocare le caratteristiche degli altri (cioè, non si è mai osservata interferenze tra un animale e l’altro). Inoltre è stato verificato che una sola caratteristica distintiva data in input (marginale o saliente) risulta sufficiente per evocare tutte le proprietà salienti di un singolo concetto.
- ii. La formazione di una categoria avviene solo se sono utilizzate come input le *features* condivise: la rete riconosce correttamente la categoria corrispondente e non evoca mai le caratteristiche distintive dei singoli concetti.

Capitolo 3

MODELLO DI MEMORIA SEMANTICA CON CONTESTO

3.1 Contesto

3.1.1 Influenza del contesto sui concetti

A prima vista, le rappresentazioni concettuali, ad esempio la nostra nozione interna del concetto di “arancia”, sembrano statiche: si ha infatti l'impressione che il concetto “arancia” rappresenti qualcosa di costante (cioè un agrume acido, di colore arancione e forma sferica) e che questo significato non subisca variazioni. La ricerca, nell’ambito degli studi riguardanti la memoria semantica, in passato ha tradizionalmente adottato questa prospettiva statica della rappresentazione concettuale, ma più di recente si è iniziato a domandarsi quanto siano effettivamente statiche le rappresentazioni dei concetti e ancora quanto esse siano costanti tra persone diverse e nel tempo. Si potrebbe supporre che la memoria semantica sia essenzialmente *context-free* (sicuramente una arancia possiede determinate proprietà invarianti) indipendentemente da qualsiasi evento specifico e condivisa. Tuttavia, per comprendere il sistema semantico e i dati sperimentali derivanti dagli studi in questo campo, le teorie che cercano di descrivere la rappresentazione dei concetti devono inevitabilmente incorporare spiegazioni relative ad aspetti della memoria semantica che non sono né liberi dal contesto, né indipendenti da eventi specifici, né condivisi. Le attuali spiegazioni riguardanti le conoscenze concettuali sono dunque a favore della possibilità che le rappresentazioni dei concetti siano più dinamiche. Secondo questi modelli, i concetti sono rappresentazioni multidimensionali distribuite (in parte) tra le regioni del cervello che sono alla base delle funzioni sensoriali, percettive e motorie. Pertanto, il significato sensoriale-motorio di un semplice concetto come quello di “arancia” non è un tutto indivisibile, bensì è distribuito attraverso una gamma di *features* [5].

A causa di questa architettura distribuita, non vi è alcun impedimento per determinate caratteristiche di un concetto ad essere più attive di altre (ad esempio, a seconda delle circostanze). Inoltre, un'architettura distribuita consente alle rappresentazioni concettuali di essere dinamicamente sensibili al contesto.

Collettivamente gli studi principali suggeriscono che le rappresentazioni concettuali siano fluide, cambiando non solo in funzione del contesto, in riferimento alla modalità e al compito dello stimolo, ma anche in funzione del contesto portato da un particolare individuo (attraverso esperienza personale, degenerazione neuronale, preferenze di elaborazione o abilità). Ci sono prove infatti a sostegno della tesi che anche la modalità di esperienza avuta con un oggetto (ad esempio, quale mano viene generalmente utilizzata per interagire con esso) influisce sulla sua rappresentazione concettuale. Infine, dato che individui diversi hanno svariate esperienze, le rappresentazioni concettuali di uno stesso concetto risultano differire tra gli individui. Per chiarire, un carpentiere avrà una rappresentazione concettuale del concetto di “martello”, inevitabilmente legata al contesto lavorativo in cui opera tutti i giorni, che risulta diverso rispetto all’idea che possiede una casalinga del medesimo utensile [5,6].

3.1.2 Features contesto-dipendenti e indipendenti

Alcuni studi di ricerca di psicologia cognitiva e di semantica lessicale hanno evidenziato che i concetti non sono delle rappresentazioni stabili, ma la loro struttura varia considerevolmente anche a seconda del contesto. Bersalou (*“Context-independent and context-dependent information in concepts”*, 1982) propone una distinzione tra le svariate proprietà che costituiscono i concetti, in particolare discrimina tra proprietà indipendenti dal contesto (CI) e proprietà dipendenti dal contesto (CD). Le prime si attivano in tutte le occasioni in cui si pensa ad un determinato concetto (ad esempio “serpente a sonagli” solitamente evoca la caratteristica “velenoso”) e generalmente sono *features* che vanno a costituire i significati fondamentali dei concetti.

Le caratteristiche *context-dependent* invece vengono attivate solo in contesti particolari e rappresentano una fonte di variabilità della codifica semantica. Tali proprietà sembrano infatti venire in mente raramente, e quando lo fanno, è solo in contesti rilevanti. Se si pensa al concetto di “pianoforte”, ad esempio, vengono spontaneamente evocate *features* CI quali “ha i tasti” o “si usa per suonare” mentre la proprietà “pesante” viene citata saltuariamente. Lo stesso concetto di “pianoforte” però, richiamato nell’ambito di un trasloco, evocherà più facilmente la *feature* “pesante” poiché legata al contesto di una possibile esperienza passata. Nei casi in cui sia presente una grande quantità di informazioni o di molteplici significati associati ad uno stesso concetto, i processi di controllo cognitivi svolgono un ruolo cruciale nel garantire che vengano richiamate solo le giuste informazioni. Senza le informazioni contestuali la comprensione potrebbe risultare meno adeguata alle particolari esigenze di volta in volta presentate dall’ambiente circostante [1].

La visione che prevede la separazione tra *features* CD e CI, si colloca a metà strada tra due opposte teorie riguardanti il significato concettuale: una sostiene che tutte le proprietà di un concetto siano attive in tutte le occasioni; l'altra afferma che le proprietà attive siano completamente determinate dal contesto [6]. Barsalou e i suoi collaboratori hanno proposto una teoria secondo la quale le *features* vengono evocate automaticamente dopo essere state frequentemente associate ad un dato concetto; inoltre frequenti accoppiamenti tra concetto e proprietà fanno sì che venga stabilita in memoria una relazione automatizzata. Di conseguenza si ritiene che le proprietà CD siano generalmente inattive (quindi non vengano evocate in maniera spontanea) poiché solo raramente elaborate contemporaneamente ai rispettivi concetti. Può accadere infine che una proprietà CD possa cambiare stato e diventare *feature* indipendente dal contesto per un dato concetto (a riprova della dinamicità della rappresentazione concettuale). Ad esempio, la *feature* “indossa ferri di cavallo” può essere CD per persone che sporadicamente si trovano a contatto con cavalli, potrebbe diventare CI per qualcuno durante un corso di apprendimento per diventare maniscalco [1].

3.2 Formulazione matematica del contesto

Nel lavoro di tesi da me svolto la formulazione matematica che esprime la dipendenza di un dato concetto dal contesto si basa sulla teoria della probabilità, in particolare sul concetto di probabilità condizionata [7].

La probabilità, secondo il classico approccio frequentista, è definita come rapporto tra “numero di casi favorevoli” e “numero di casi possibili” e, più in generale, è numero che esprime il grado di verosimiglianza con cui un evento è destinato a verificarsi. Si definisce probabilità marginale, la probabilità del verificarsi di un evento semplice E , indicata come $P(E)$. Si parla invece di probabilità congiunta quando ci si riferisce al verificarsi simultaneo di due o più eventi $P(E_1 \cap E_2)$. A partire da questa ultima definizione, la probabilità marginale può essere interpretata a partire da un insieme di probabilità congiunte. Sotto certe condizioni e considerando due soli eventi E_1 ed E_2 , la probabilità marginale dell'evento E_1 , cioè $P(E_1)$, può essere calcolata sommando la probabilità congiunta di E_1 e di E_2 alla probabilità congiunta di E_1 e \bar{E}_2 , come specificato di seguito (Equazione 12):

$$(12) \quad P(E_1) = P(E_1 \cap E_2) + P(E_1 \cap \bar{E}_2)$$

dove E_2 e \bar{E}_2 sono eventi incompatibili e collettivamente esaustivi [7].

Il calcolo di una probabilità si modifica quando si dispone di una certa informazione sugli eventi coinvolti ed in questo caso si parla di probabilità condizionata. Il condizionamento è utile quando si vuole analizzare un certo evento A , definito evento condizionato, a partire dal possesso di una certa informazione su un evento B , chiamato invece condizionante. Si definisce quindi la probabilità condizionata, indicata con il simbolo $P(A|B)$, quella probabilità per cui si verifichi un certo evento A , sapendo che l'evento B si è verificato. Nella pratica questa probabilità (Equazione 13) si ottiene rapportando la probabilità congiunta $P(A \cap B)$ alla probabilità marginale del singolo evento B che si è verificato:

$$(13) \quad P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Si noti che deve sempre valere $P(B) > 0$ altrimenti la probabilità condizionata non avrebbe senso (non è possibile definire la probabilità condizionata rispetto ad un evento impossibile). Non tutti i possibili eventi sono direttamente osservabili. Se però l'evento A è in qualche modo legato ad un secondo evento B , che invece possiamo osservare, la probabilità condizionata $P(A|B)$ prende il nome di probabilità “a posteriori”, poiché indica un valore di probabilità valutata dopo la conoscenza di B . Viceversa, la probabilità marginale dell'evento A , cioè $P(A)$, viene definita probabilità “a priori”, dal momento che non si conosce nulla a suo riguardo [7].

Obiettivo di questa breve trattazione è quello di riuscire a determinare una espressione della probabilità a priori (probabilità dell'evento condizionato). A partire dalla Equazione 13, si può scrivere la probabilità congiunta come prodotto tra probabilità condizionata e marginale, in particolare vale:

$$(14a) \quad P(A \cap B) = P(B)P(A|B)$$

$$(14b) \quad P(A \cap \bar{B}) = P(\bar{B})P(A|\bar{B})$$

Riprendendo l'espressione data dall'Equazione 12, $P(A)$ risulta la somma delle probabilità congiunte quindi posso esprimerla come segue:

$$(15) \quad P(A) = P(A \cap B) + P(A \cap \bar{B})$$

Sostituendo infine le due espressioni 14a e 14b delle probabilità congiunte nella Equazione 15 e ricordando che vale la relazione $P(\bar{B}) = 1 - P(B)$, si ottiene l'espressione finale (Equazione 16) della probabilità a priori, che verrà utilizzata per implementare l'addestramento con contesto.

$$(16) \quad P(A) = P(B)P(A|B) + (1 - P(B))P(A|\bar{B})$$

L'idea fondamentale espressa da tale equazione considera che la $P(A|B)$ sia molto maggiore rispetto alla $P(A|\bar{B})$ e che quindi l'evento A si verifichi molto spesso insieme all'evento B . Questa equazione dunque implica il concetto di contesto dal momento che si va creare una sinapsi che richiama A ogni volta che è presente B , anche se A è un evento marginale.

3.3 Addestramento delle sinapsi con contesto

3.3.1 Procedura di addestramento

L'addestramento delle sinapsi del modello con contesto non si discosta di molto rispetto a quello del modello di base. La differenza sostanziale rispetto al modello trattato nel capitolo precedente consiste nell'utilizzo di un addestramento più complesso che fa uso della probabilità condizionata per caratterizzare le *feature* dipendenti dal contesto [7].

Complessivamente si implementa un addestramento di mille epoche consecutive e le sinapsi eccitatorie all'interno della rete vengono addestrate con il meccanismo hebbiano a soglia variabile per risolvere il problema della salienza delle caratteristiche parzialmente condivise. Unico requisito aggiuntivo che caratterizza la rete semantica con contesto (per gli altri requisiti si faccia riferimento al paragrafo 2.3.3) è che le *features context-dependent* devono ricevere sinapsi in input solamente dalle altre proprietà appartenenti allo stesso specifico contesto. Come le proprietà marginali della rete di base (che sono invece CI), devono però inviare forti sinapsi di output verso le *features* salienti per favorire la ricostruzione del concetto.

In accordo con il modello di base privo di contesto, alle caratteristiche salienti è stata associata un'alta probabilità di verificarsi, pari al 70% (cioè $P_{kn(s)} = 0.7$), mentre alle caratteristiche marginali indipendenti dal contesto è stata associata una bassa probabilità pari 35% (cioè $P_{kn(m)} = 0.35$).

Per quanto riguarda le probabilità associate alle caratteristiche dipendenti dal contesto occorre fare alcune precisazioni. Per semplicità della trattazione considero il caso di un contesto realizzato da due sole *features context-dependent*: alla proprietà condizionante ($f1$) viene assegnata la medesima probabilità che contraddistingue le caratteristiche marginali ($perc_f1 = 0.35$) mentre la probabilità assegnata alla proprietà condizionata ($f2$) viene calcolata attraverso l'espressione finale della probabilità a priori (Equazione 16).

Il comportamento che si intende ottenere è il seguente:

- Quando si verifica la prima proprietà $f1$, la caratteristica $f2$ deve avere una elevata probabilità di verificarsi. In questo modo le *features* $f1$ e $f2$ si scambiano forti sinapsi evocandosi quindi a vicenda.
- Quando invece la proprietà $f1$ non si verifica, allora la caratteristica $f2$ deve avere basse probabilità di verificarsi. Raramente dovrà capitare che la *feature* $f2$ si verifichi senza $f1$; le due proprietà non devono scambiarsi forti sinapsi in modo tale da non evocarsi a vicenda.

Per far in modo che ciò avvenga devono essere assegnati opportuni valori alle due probabilità condizionate. In particolare, la probabilità che la *feature* condizionata $f2$ si presenti dal momento che si è verificata $f1$, indicata come $P(f2|f1)$, deve essere elevata ed è stata settata ad un valore massimo pari a $P_{max} = 0.85$. Di contro la probabilità che la $f2$ si presenti dal momento che non si è verificata la $f1$, che rappresenta invece la probabilità $P(f2|\overline{f1})$, deve essere molto bassa e si è scelto un valore pari a $P_{min} = 0.05$. Di seguito si riporta il codice *matlab* utilizzato nella procedura di addestramento che realizza tale comportamento:

```
I(f1)=1*(rand(1,1)<perc_f1);  
I(f2)=I(f1)*(rand(1,1)<0.85)+(1-I(f1))*(rand(1,1)<0.05);
```

Per disporre di un quadro esaustivo in grado di spiegare l'influenza del contesto, l'addestramento delle sinapsi è stato ripetuto per diverse configurazioni relative alle *features context-dependent* legate tra loro in svariati modi al fine di implementare differenti tipologie di contesto.

Un primo addestramento ha previsto lo schema che relaziona tra loro tre *features* dipendenti dal contesto in cui la prima caratteristica $f1$ risulta condizionante per le altre due proprietà $f2$ ed $f3$ (Figura 12). Il codice corrispondente per tale l'addestramento contestuale risulta:

```
I(f1)=1*(rand(1,1)<perc_f1);  
I(f2)=I(f1)*(rand(1,1)<0.85)+(1-I(f1))*(rand(1,1)<0.05);  
I(f3)=I(f1)*(rand(1,1)<0.85)+(1-I(f1))*(rand(1,1)<0.05);
```

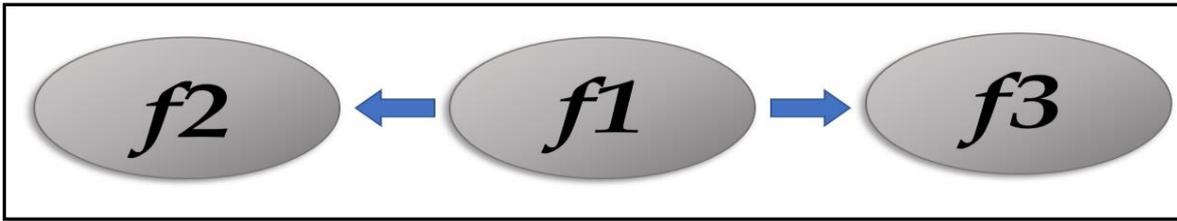


Figura 12 – Schema di contesto in cui la feature $f1$ è condizionante sia per $f2$ sia per $f3$.

È stato poi implementato un secondo addestramento contestuale, il cui schema si può osservare nella figura sottostante (Figura 13). Questa modalità prevede che la *feature* $f1$ condizioni la caratteristica $f2$, la quale a sua volta risulta condizionante per la proprietà $f3$. Il codice relativo a questo specifico addestramento del contesto risulta:

```
I (f1)=1* (rand(1,1)<perc_f1) ;
I (f2)=I (f1) * (rand(1,1)<0.85)+(1-I (f1)) * (rand(1,1)<0.05) ;
I (f3)=I (f2) * (rand(1,1)<0.85)+(1-I (f2)) * (rand(1,1)<0.05) ;
```

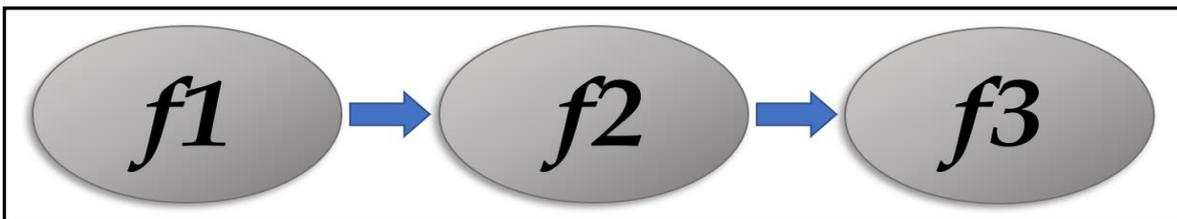


Figura 13 – Schema di contesto in cui la feature $f1$ è condizionante per $f2$ che a sua volta è condizionante per $f3$.

Infine, un'ultima configurazione per l'addestramento del contesto è stata realizzata considerando un totale di quattro *features context-dependent* ($f1$, $f2$, $f3$, $f4$) ognuna delle quali condiziona la caratteristica successiva come si può vedere dallo schema riportato in Figura 14.

```
I (f1)=1* (rand(1,1)<perc_f1) ;
I (f2)=I (f1) * (rand(1,1)<0.85)+(1-I (f1)) * (rand(1,1)<0.05) ;
I (f3)=I (f2) * (rand(1,1)<0.85)+(1-I (f2)) * (rand(1,1)<0.05) ;
I (f4)=I (f3) * (rand(1,1)<0.85)+(1-I (f3)) * (rand(1,1)<0.05) ;
```

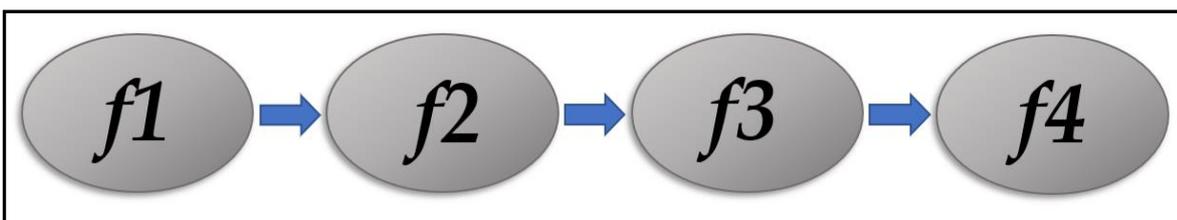


Figura 14 – Schema di contesto in cui ogni feature è condizionante per la successiva.

Si noti che, in tutti questi schemi, indipendentemente dalla loro particolare realizzazione, l'idea di base è che le proprietà (tre o quattro) abbiano una elevata probabilità di verificarsi tutte contemporaneamente (contesto) e quindi si connettano reciprocamente con forti sinapsi eccitatorie. Di conseguenza, dopo l'addestramento, anche solo la presenza di una di queste *features* è in grado di rievocare le altre.

3.3.2 Assegnazione dei parametri

I parametri utilizzati per la fase di addestramento sinaptico con contesto sono elencati nella tabella sottostante (Tabella 7) e risultano praticamente identici a quelli utilizzati per addestrare la rete di base. Dapprima sono elencati i parametri necessari all'implementazione delle equazioni del modello, poi i valori che consentono di realizzare l'apprendimento hebbiano, infine le probabilità determinanti salienza e marginalità delle *features*. Si aggiungono in questo caso i due soli parametri che esprimono le probabilità condizionate, il cui significato è già stato discusso nel paragrafo precedente.

Significato	Simbolo	Valore
Costante di tempo	τ	3 ms
Pendenza della sigmoide	T	0.01
Posizione della sigmoide	φ	0.55
Soglia pre-sinaptica	ϑ_{pre}	0.05
Soglia post-sinaptica (valore basale)	ϑ_{post}	0.67
Saturazione della soglia post-sinaptica	$\vartheta_{post-sat}$	0.95
Velocità di incremento della soglia post-sinaptica	$\Delta\vartheta_{post}$	0.5
Rateo di apprendimento	γ_0	0.02
Massima forza sinaptica	W_{max}	0.8
Probabilità associata a <i>features</i> salienti	$P_{kn(s)}$	0.7
Probabilità associata a <i>features</i> marginali	$P_{kn(m)}$	0.35
Probabilità condizionata massima	P_{max}	0.85
Probabilità condizionata minima	P_{min}	0.05

Tabella 7 – Parametri utilizzati per l'addestramento sinaptico con contesto.

3.4 Risultati

Di seguito si effettua l'analisi dei *pattern* di sinapsi ottenuti a seguito della fase di addestramento, tenuto conto, questa volta, della presenza di *features* legate dal contesto. Verranno successivamente presentati i risultati di alcune simulazioni riguardanti attività di riconoscimento di concetti, ponendo particolare attenzione ai *test* che prevedono in input proprietà *context-dependent*. Anche in questa sezione, si prende in considerazione la rete semantica descritta in precedenza realizzata a partire dalla tassonomia di animali, medesimi risultati si ottengono con la tassonomia di oggetti.

3.4.1 Risultati dell'addestramento sinaptico

Di seguito si riportano, a titolo di esempio, le sinapsi di alcune *features* della tassonomia di animali ottenute a seguito dell'addestramento. Per le specifiche posizioni assunte dalle caratteristiche utilizzate si faccia sempre riferimento allo schema riportato in Figura A1 nella sezione Appendice. Dapprima si riassume brevemente il comportamento delle sinapsi di: una *feature* condivisa che concorre a formare la super categoria "ANIMALE" ("si riproduce"), un attributo che partecipa alla costituzione della macro categoria "UCCELLO" ("ha le ali") e una proprietà distintiva saliente ("canta al mattino") del concetto di "gallo" (schema in Figura 15).

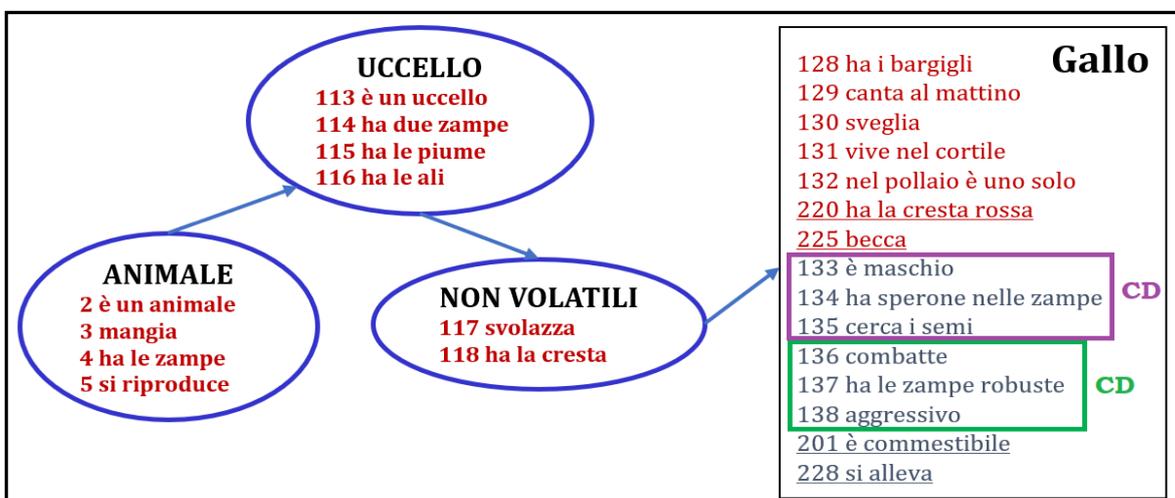


Figura 15 – Schema relativo al concetto di "gallo" e possibili features CD in evidenza.

Particolare attenzione si presterà alle sinapsi di tre *features* distintive marginali di “gallo”, ovvero le proprietà “combatte” (136), “ha zampe robuste” (137), “è aggressivo” (138) selezionate per individuare uno specifico contesto.

La caratteristica “si riproduce” (in posizione 5) è saliente e comune a tutti gli animali: riceve sinapsi da tutte le proprietà degli animali, tranne che da sé stessa. Inoltre, si ricorda che la forza sinaptica è pesata dalla salienza: maggiore per le *features* salienti e minore per quelle marginali. La proprietà “ha le ali” (116) risulta saliente e comune solo a tutti gli uccelli, di conseguenza le sinapsi entranti provengono solo dagli animali appartenenti a tale categoria (da 113 alla 186). Infine, un attributo distintivo saliente per il concetto di “gallo” è “canta al mattino” (129), il quale riceve sinapsi forti da tutte le altre caratteristiche distintive (salienti e marginali) dello stesso animale. Le sinapsi relative a queste tre *features* appena citate (5, 116, 129) mostrano un comportamento pienamente concorde con risultati di un corretto addestramento sinaptico.

Si ponga ora attenzione alle tre caratteristiche *context-dependent* collocate nelle posizioni 136, 137, 138 della rete (rispettivamente “combatte”, “ha le zampe robuste”, “è aggressivo”). Esse, a differenza delle semplici *features* distintive marginali (che hanno sinapsi entranti praticamente inesistenti), ricevono sinapsi dalle altre proprietà legate al contesto di cui fanno parte (riquadro in verde in Figura 15) ma non ricevono sinapsi da sé stesse. Inoltre, tali sinapsi, dal momento che le proprietà di contesto sono in ogni caso marginali per il concetto, non raggiungono mai il massimo valore di saturazione W_{max} . In aggiunta occorre notare che queste tre *features* non ricevono sinapsi dalle altre proprietà del concetto, quindi non sono evocate al di fuori di quel particolare contesto. In riferimento all’esempio pratico riportato in Figura 16 (pagina seguente), si osserva che la *feature* di contesto “combatte” (136) riceve sinapsi solamente dalla caratteristica “ha le zampe robuste” (137) e dalla proprietà “è aggressivo” (138).

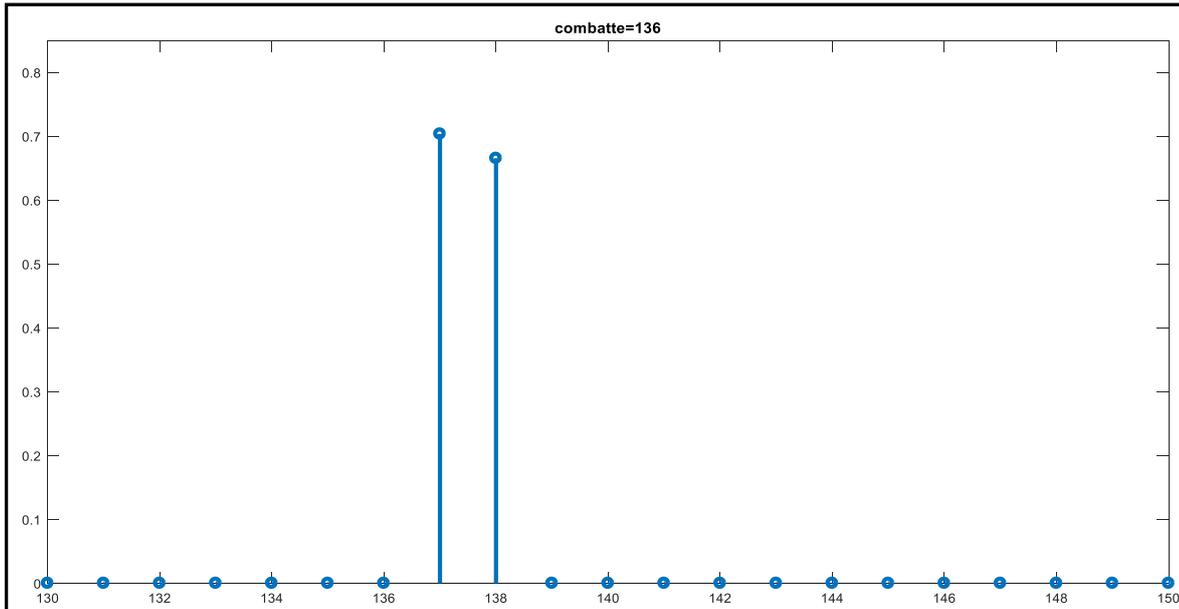


Figura 16 – Forza delle sinapsi entranti nella feature CD “combatte” (136).

Ragionamento analogo vale per le altre due proprietà CD, le cui sinapsi sono riportate nelle Figure 17 e 18. Un comportamento delle sinapsi del tutto equivalente a quello appena descritto si ottiene per altre tre *features* scelte per creare un differente contesto (in viola in Figura 15): si tratta delle proprietà “è maschio” (133), “ha lo sperone nelle zampe” (134) e “cerca i semi” (135). Per brevità non si riportano le figure corrispondenti.

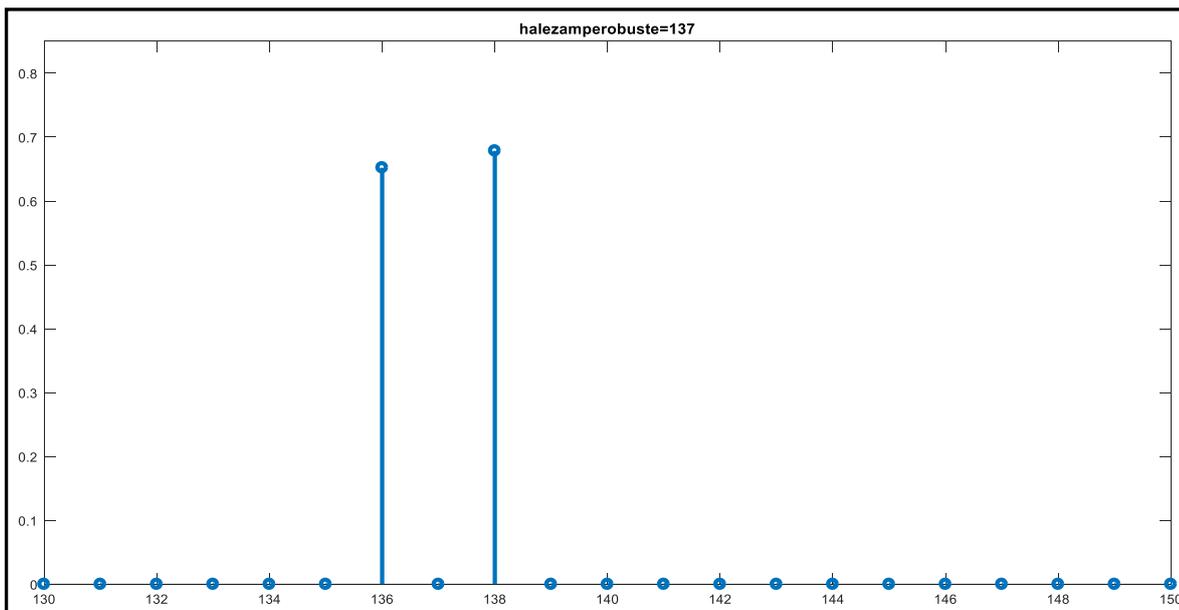


Figura 17 – Forza delle sinapsi entranti nella feature CD “ha le zampe robuste” (137).

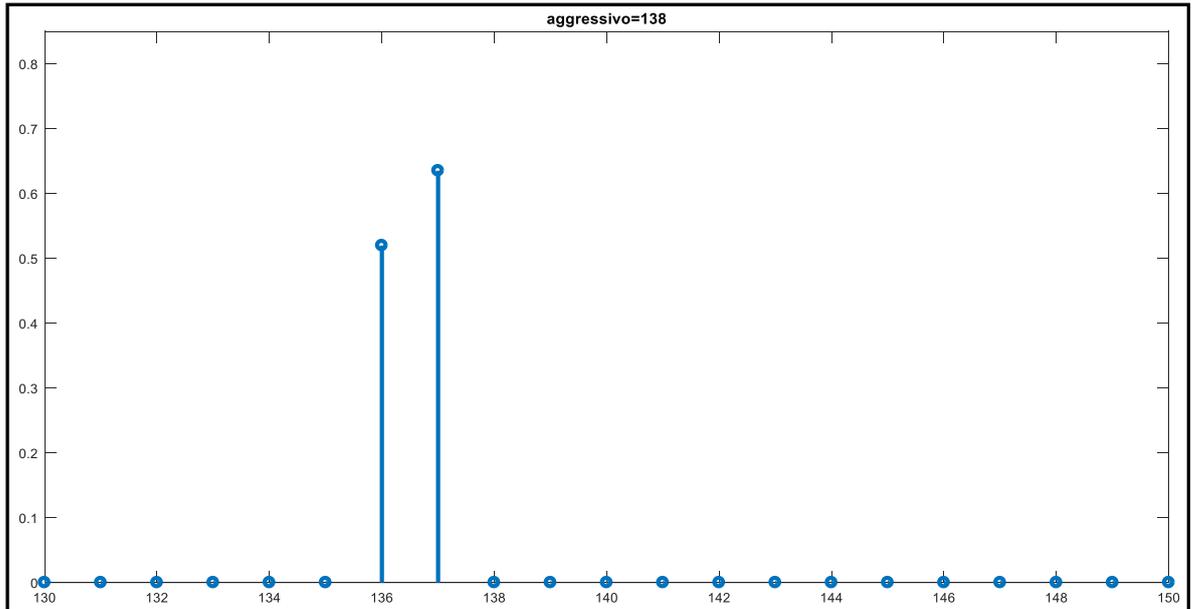


Figura 18 – Forza delle sinapsi entranti nella feature CD “è aggressivo” (138).

3.4.2 Risultati delle simulazioni

Anche per il modello con contesto, le sinapsi ottenute dall’addestramento risultano abbastanza soddisfacenti per cui, dopo la fase di addestramento, si è proceduto con i *test* per simulare compiti di riconoscimento di concetti. Non verranno di seguito considerate tutte possibili tipologie di *features* come possibili input bensì ci si limita ad analizzare il comportamento della rete a fronte di ingressi costituiti da proprietà dipendenti dal contesto.

Dato un input costituito da una proprietà CD, questa volta il riconoscimento si considera efficace se il modello riesce ad evocare tutte le caratteristiche salienti del concetto ed allo stesso tempo tutte quelle appartenenti al medesimo contesto. Se invece alla rete viene data in ingresso una *feature* indipendente dal contesto (condivisa, marginale o saliente) il comportamento del modello attuale risulta del tutto equivalente a quello del modello base (devono essere evocate solo le proprietà salienti). L’evoluzione della simulazione viene sempre controllata catturando diverse istantanee (ancora una volta agli *step* 10, 20, 40). Si riportano di seguito le simulazioni dei compiti di riconoscimento a fronte di un addestramento sinaptico nel quale sono state utilizzate tre *features* CD che si richiamano a vicenda.

Per semplicità, lo schema di addestramento scelto per discutere i risultati è così costituito: la proprietà $f1$ è condizionante per la $feature$ $f2$ e allo stesso tempo è condizionante anche per la caratteristica $f3$ (si faccia riferimento alla Figura 12). Per commentare con chiarezza i risultati ottenuti si consideri, a titolo di esempio, il compito di riconoscimento del concetto di “gallo” ed in particolare le caratteristiche *context-dependent* “combatte”, “ha le zampe robuste”, “è aggressivo”. La prima simulazione (Figura 19), che ha previsto in input la caratteristica “combatte” (136), ha evidenziato un corretto riconoscimento del concetto: la rete è riuscita infatti ad evocare spontaneamente tutte le $features$ salienti di “gallo” e allo stesso tempo anche le altre due proprietà CD (137, 138) accomunate da sinapsi ora molto forti. In particolare, tra le caratteristiche salienti sono emerse quelle che costituiscono le categorie “ANIMALE” (2, 3, 4, 5), “UCCELLO” (113, 114, 115, 116) e “NON VOLATILE” (117, 118), quelle salienti distintive (da 128 a 132) ed infine quelle parzialmente condivise ma salienti (220 e 225). Il valido riconoscimento del concetto di interesse è stato ottenuto anche per gli altri due input (“ha le zampe robuste” e “combatte”) di cui si possono osservare le simulazioni nelle Figure 20 e 21 nella pagina seguente.

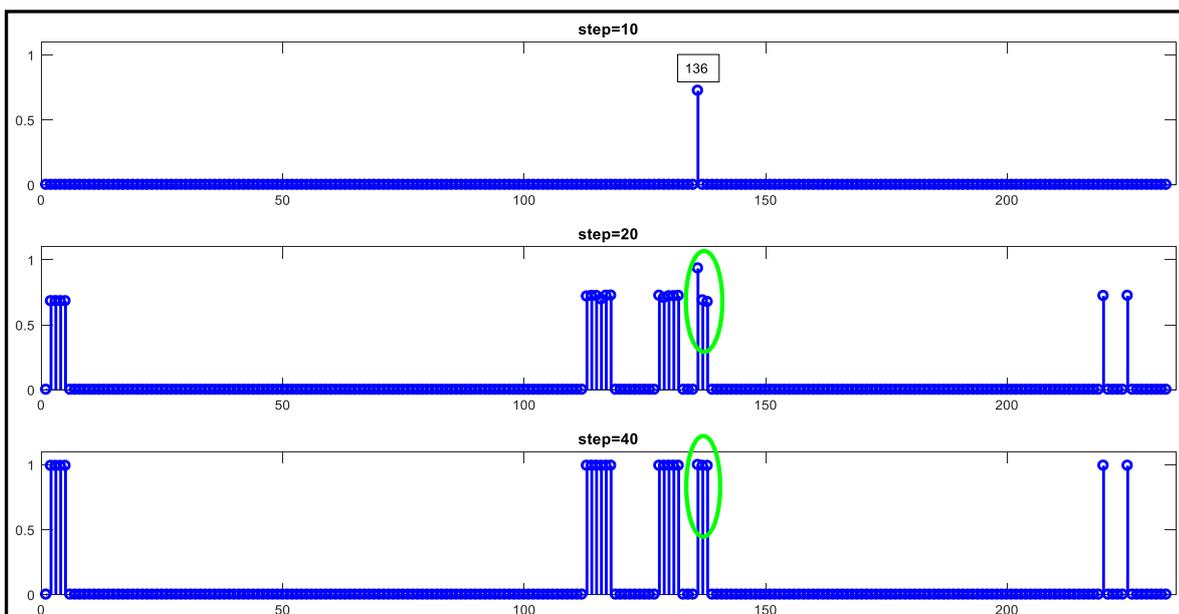


Figura 19 – Riconoscimento del concetto di “gallo”, in input si considera la feature dipendente dal contesto “combatte” (136).

Anche per la rete con contesto si nota come le *features* chiamate in input tendano a raggiungere rapidamente il livello di saturazione (ventesimo *step*) mentre le proprietà evocate giungono a saturazione più lentamente, in particolare quelle CD che non sono fornite in input. Questo aspetto può essere indice di una maggiore difficoltà della rete nel far emergere le *features* di contesto, le quali restano pur sempre caratterizzate da marginalità.

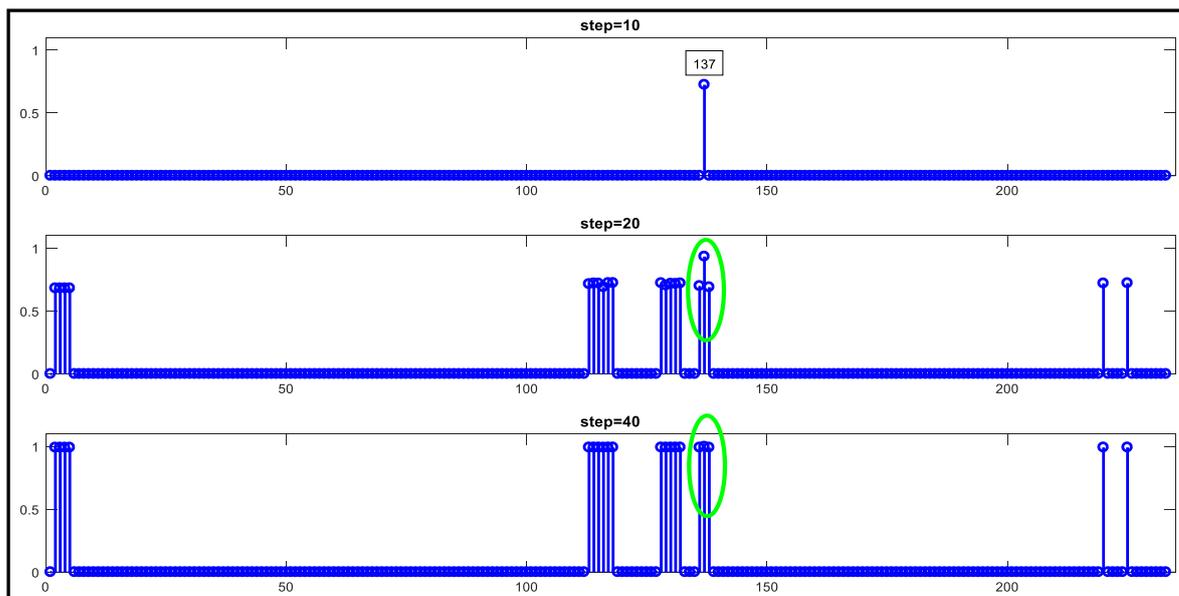


Figura 20 – Riconoscimento del concetto di “gallo”, in input si considera la feature dipendente dal contesto “ha le zampe robuste” (137).

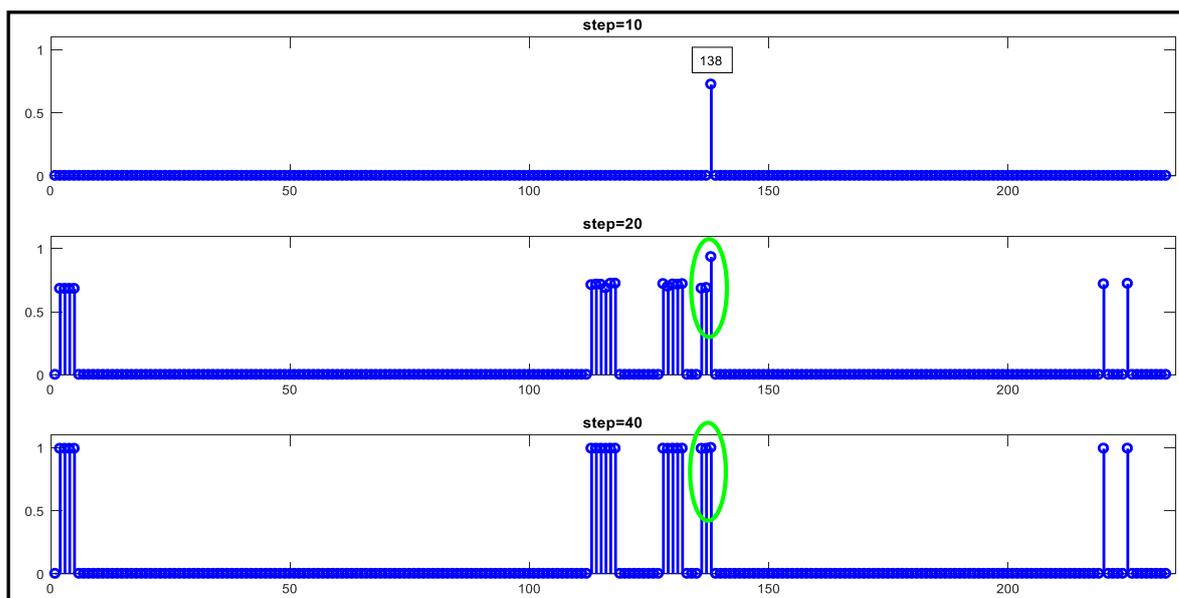


Figura 21 – Riconoscimento del concetto di “gallo”, in input si considera la feature dipendente dal contesto “è aggressivo” (138).

Si vuole ora discutere l'interessante caso in cui siano presenti contemporaneamente due differenti contesti per lo stesso concetto di "gallo". Facendo ancora una volta riferimento alla Figura 15, si considera il primo contesto (evidenziato in verde) costituito dalle *features* "combatte" (136), "ha le zampe robuste" (137) ed "è aggressivo" (138) mentre al secondo (identificato dal colore viola) appartengono le proprietà "è maschio" (133), "ha lo sperone nelle zampe" (134) e "cerca i semi" (135).

A fronte di un singolo input *context-dependent* appartenente al primo o al secondo contesto, la rete evoca correttamente tutte le *features* salienti per il concetto di "gallo" e le altre due proprietà CD del contesto specifico di cui l'input fa parte, senza che i due differenti contesti interferiscano in alcun modo tra loro. Dunque, il contesto che non viene stimolato dal alcun input esterno rimane totalmente inattivo.

Si può inoltre osservare che, se in input vengono fornite contemporaneamente due proprietà *context-dependent*, una per ciascun contesto, la rete mostra un comportamento interessante, il quale viene mostrato nella figura seguente (Figura 22). La *feature* del primo contesto data in input (136) richiamerà altre due CD ad essa correlate (137, 138) mentre la proprietà in input appartenente al secondo contesto (133) richiamerà le caratteristiche CD da essa condizionate (134,135). In aggiunta a ciò, la rete riesce correttamente ad evocare tutte la *features* salienti del concetto di "gallo", riuscendo anche in questo caso a svolgere efficientemente il compito di riconoscimento del concetto. Si può infine notare come le *features* contestuali che vengono evocate arrivino a saturazione in intervalli temporali simili ma non del tutto identici (a seconda del contesto di appartenenza). Al medesimo istante di simulazione, ad esempio attorno al ventesimo *step*, si può osservare come le due proprietà 137 e 138 appartenenti al contesto in verde risultino avere una maggiore attività neurale rispetto alle proprietà 134 e 135 del contesto evidenziato in viola.

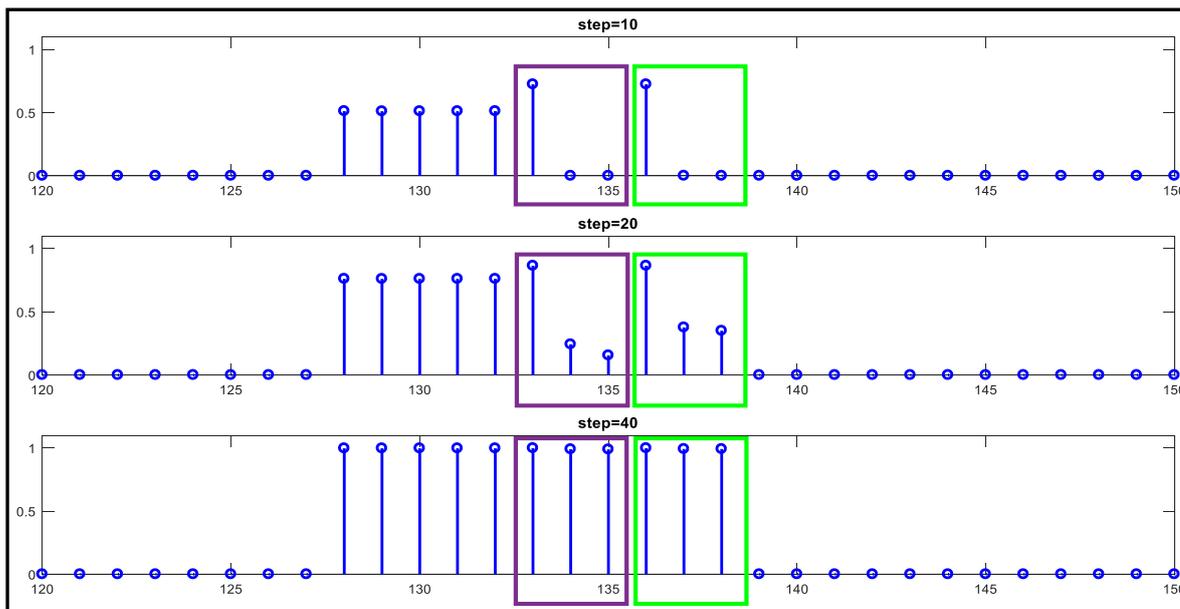


Figura 22 – Riconoscimento del concetto di “gallo”, in input si considerano due features appartenenti a contesti differenti, in particolare: “combatte” (136) per il primo contesto (riquadro verde) e “è maschio” (133) per il secondo contesto (riquadro in viola).

Per completezza della trattazione si riporta infine una simulazione del compito di riconoscimento relativa alla rete semantica con contesto basata sulla tassonomia di oggetti. Per la precisa corrispondenza delle *features* utilizzate in questo esempio si faccia riferimento allo schema riportato in Figura A2 della sezione Appendice. Nello specifico si è voluto simulare (come mostrato dalla Figura 23) il compito di riconoscimento del concetto di “ombrello” per cui sono state scelte quattro proprietà *context-dependent*.

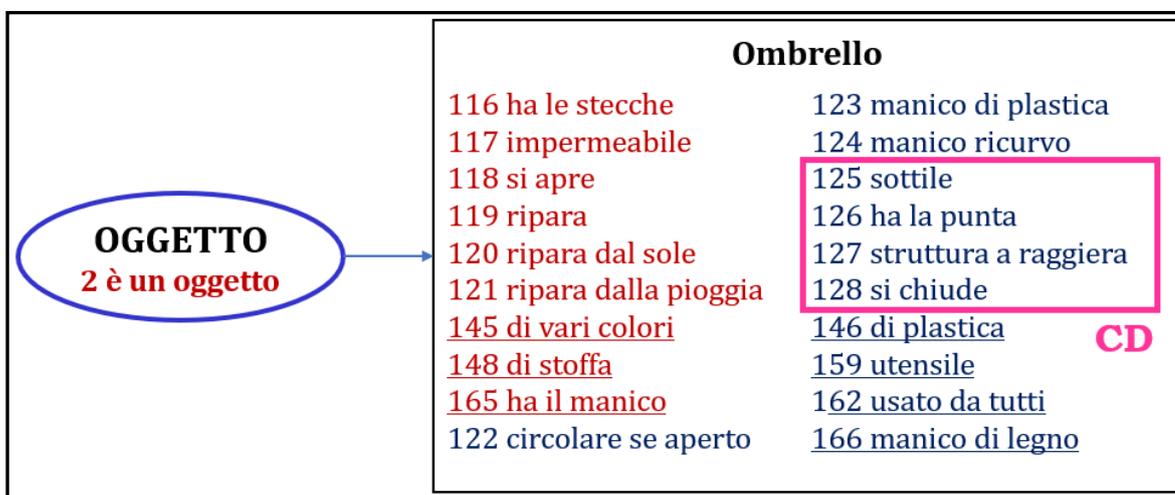


Figura 23– Schema relativo al concetto di “ombrello” e possibili features CD in evidenza.

Lo schema scelto per l'addestramento prevede che ognuna di queste caratteristiche CD condizioni la caratteristica successiva (in accordo con l'addestramento riportato in Figura 14). Le quattro *features* CD selezionate sono le seguenti: “sottile” (125), “ha la punta” (126), “ha struttura a raggiera” (127) e “si chiude” (128).

La simulazione riportata nella figura sottostante (Figura 24) evidenzia il corretto riconoscimento del concetto in quanto la rete è riuscita ad evocare spontaneamente tutte le *features* salienti di “ombrello” e allo stesso tempo anche le altre tre proprietà CD (125, 126, 127) a fronte di un input costituito dalla proprietà CD “si chiude” (128). In particolare, tra le caratteristiche salienti sono emerse, oltre alla caratteristica di categoria “è un oggetto” (2), tutte le proprietà salienti distintive (da 116 a 121) ed infine quelle condivise salienti (145, 148 e 165). Un comportamento del tutto equivalente si ottiene se in input alla rete vengono fornite le altre tre *features context-dependent*.

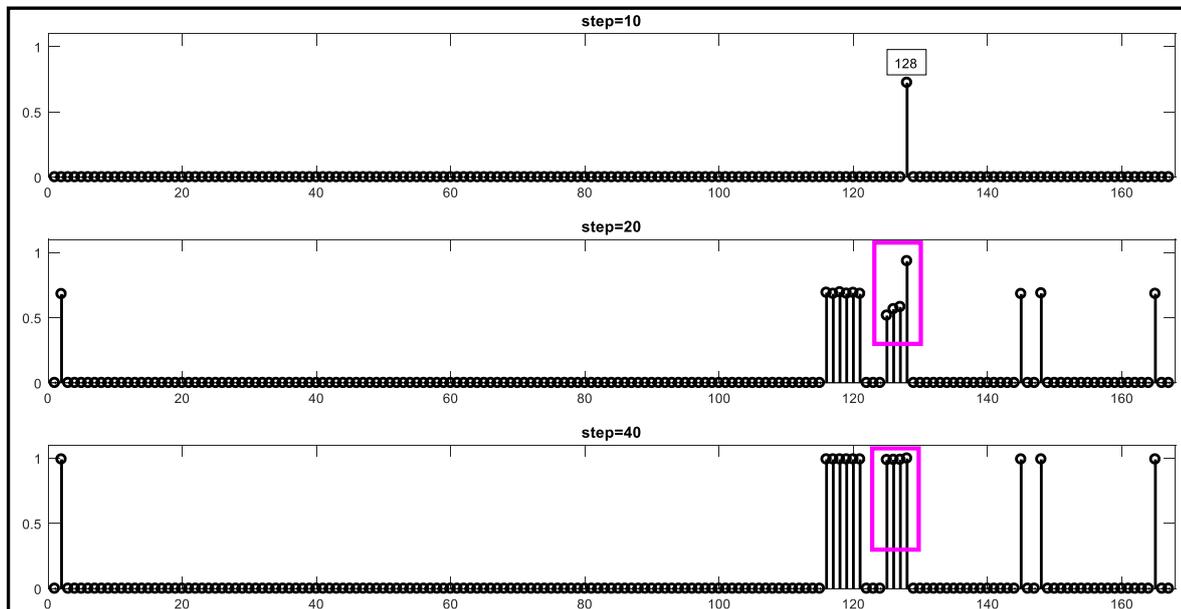


Figura 24 – Riconoscimento del concetto di “ombrello”, in input si considera la feature dipendente dal contesto “si chiude” (128).

Discussione

Le simulazioni per il compito di riconoscimento del concetto sono state ripetute per tutti i singoli animali e oggetti che realizzano le due tassonomie, a fronte di input sia indipendenti sia dipendenti dal contesto. Le prove sono state effettuate variando le *features* e il numero di attributi forniti in input e si è generalmente osservato un corretto funzionamento della rete.

Per quanto riguarda i dati relativi al contesto, si sono ottenuti risultati soddisfacenti in tutti e tre gli schemi implementati per l'addestramento (si faccia riferimento al paragrafo 3.3.1) che prevedono sia tre sia quattro caratteristiche *context-dependent*. Non si è mai verificato che una *features* CD abbia evocato erroneamente altre proprietà. Si noti che unico modo per eccitare significativamente le caratteristiche appartenenti allo stesso contesto è infonderne almeno una attraverso l'input esterno. Infine dalle simulazioni che prevedono il richiamo di più contesti contemporaneamente, è emerso come le *features* CD si comportino in maniera del tutto autonoma le une dalle altre, evidenziando come la presenza di un contesto non ne richiami un altro.

In accordo con le teorie di Barsalou, le informazioni contenute nella memoria a lungo termine vengono utilizzate dalla memoria di lavoro per costruire i concetti relativamente ai contesti e a situazioni specifiche [1,6]. Da questo punto di vista risulta chiaro come i concetti siano dei costrutti temporanei e sembrano avere una stretta dipendenza dai contesti. Una volta formati, i concetti vengono immagazzinati nella memoria a lungo termine: non costituiscono rappresentazioni rigidamente definite bensì presentano una connotazione flessibile e variabile. Anche se alcune informazioni sono completamente indipendenti dal contesto, cioè sempre incorporate in un concetto (*features* che vengono evocate spontaneamente sempre cioè salienti), altre potrebbero partecipare alla sua formazione solo nelle situazioni più pertinenti e dipendere dalle contingenze del momento. Si può dunque pensare che quelli appena citati potrebbero essere gli elementi distintivi in grado di fornire le caratteristiche di stabilità o dinamicità alle categorie.

Capitolo 4

MODELLO DI MEMORIA SEMANTICA CON OSCILLATORI

4.1 Sincronismo in banda gamma e oscillatori

La rappresentazione di concetti in varie funzioni cognitive avviene in modo altamente parallelo e distribuito: diverse caratteristiche dei singoli concetti vengono elaborate e codificate in aree corticali distinte e distanti. Una questione centrale nella neurofisiologia moderna è come questa attività neuronale distribuita sia funzionalmente collegata raggruppando le diverse *features* in una rappresentazione concettuale unitaria coerente e in che modo le caratteristiche di concetti diversi, simultaneamente presenti nella stessa scena, siano separate per evitare interferenze e false congiunzioni. Un'ipotesi influente, denominata “ipotesi di correlazione temporale”, postula che gruppi neuronali rappresentanti svariati aspetti dello stesso concetto siano legati, all'interno di un più ampio insieme di cellule, attraverso la sincronizzazione della loro attività nel *range* di frequenze gamma (30-100 Hz) [13,15]. Si ritiene che il cervello utilizzi prevalentemente il cosiddetto “*ensemble coding*”: le proprietà che codificano un medesimo oggetto devono essere simultaneamente attive, mentre le *features* di un oggetto diverso devono essere simultaneamente inibite; questo risultato può proprio essere raggiunto sfruttando il ritmo gamma. Neuron che codificano le proprietà di un medesimo oggetto percepito oscillano con un ritmo gamma in sincronismo fra loro (cioè si eccitano e si inibiscono tutti in fase) mentre l'attività dei neuroni che codificano proprietà di un altro oggetto risultano sfasate rispetto alle precedenti. Inoltre, la tendenza dei neuroni a sincronizzare la loro risposta aumenta se essi vengono ripetutamente impegnati in scariche oscillatorie sincrone in banda gamma, mentre la tendenza alla sincronizzazione diminuisce se i neuroni sono ripetutamente impegnati in scariche oscillatorie non sincronizzate. I concetti quindi (o, più in generale, i fenomeni percepiti) apparirebbero nel cervello in divisione temporale, uno alla volta, sfruttando il ritmo gamma.

Ovviamente il fenomeno avviene in modo così rapido da non risultare avvertito dalla coscienza, per cui l'impressione è quella di riconoscere più entità contemporaneamente [8]. La sincronizzazione dell'attività corticale oscillatoria in banda gamma è stata osservata in risposta a diverse classi di stimoli sensoriali e durante vari tipi di elaborazione, non solo in corteccia visiva ma anche nel sistema somato-sensoriale, uditivo e olfattivo. Forti argomentazioni supportano l'idea che la sincronizzazione neurale giochi un ruolo essenziale nell'elaborazione di compiti altamente cognitivi, oltre che nella percezione sensoriale di basso livello. Si pensa che le oscillazioni gamma svolgano una funzione primaria nella formazione della memoria dichiarativa così come nel recupero di informazioni; inoltre si ritiene che cambiamenti nella sincronia potrebbero essere determinanti per processi quali l'attenzione. Uno studio suggerisce infine che la sincronizzazione aumenti con la percezione cosciente rispetto ai processi inconsci [12, 13].

In molti compiti cognitivi di alto livello è necessario riconoscere più oggetti contemporaneamente a partire da stimoli esterni: a tal fine è necessario confrontare gli stimoli in input con una rappresentazione interna, recuperare le informazioni mancanti dall'esperienza precedente e conservare le informazioni in memoria evitando confusione. Inoltre, il riconoscimento degli oggetti deve essere indipendente dagli attributi spaziali (posizione, distanza, prospettiva), deve diffondersi attraverso varie modalità sensoriali e gli oggetti dovrebbero essere riconosciuti anche se mostrano moderati cambiamenti rispetto a una precedente rappresentazione prototipica. Quindi, il problema del riconoscimento multiplo è simile al classico problema di “*binding and segmentation*” della percezione sensoriale, problematica che negli ultimi decenni ha assunto grande importanza nell'ambito delle neuroscienze cognitive.

Nel caso di nostro interesse (memorizzazione e riconoscimento di concetti) i concetti sono considerati come una raccolta di *features* (consentendo l'invarianza spaziale) e le regole per il loro riconoscimento si basano su conoscenza pregressa e su un criterio di somiglianza (mentre le regole Gestalt di basso livello sono essenziali nell'elaborazione sensoriale precoce).

Ispirati dai precedenti incoraggianti risultati della modellazione di reti neurali e da recenti studi che supportano il ruolo della sincronizzazione in banda gamma nelle funzioni corticali superiori, numerosi modelli di reti basati su oscillatori sono stati proposti a partire dalla metà degli anni ottanta nel tentativo di chiarire il ruolo della sincronizzazione nel riconoscimento e recupero di oggetti. Questi modelli si differenziano per diversi aspetti: la tipologia di oscillatori utilizzati per implementare la rete, il livello di complessità ma anche l'affidabilità fisiologica [3,13]. In questo elaborato verrà discussa una rete di oscillatori di Wilson-Cowan che aspira a simulare la segmentazione ad alti livelli cognitivi, piuttosto che a bassi livelli sensoriali. La rete realizza la separazione di oggetti simultanei e il loro riconoscimento (cioè la ricostruzione dalla memoria) su un singolo livello di elaborazione, raggruppando un insieme di caratteristiche fondamentali.

4.2 Formulazione matematica degli oscillatori

Per fornire un meccanismo computazionale robusto che sia in accordo con la teoria del sincronismo oscillatorio in banda gamma, la rete di oscillatori deve rispettare i seguenti requisiti basilari [15]:

- i. Deve essere in grado di sincronizzare un gruppo di oscillatori localmente accoppiati;
- ii. Deve essere in grado di desincronizzare gruppi di oscillatori attivati da concetti diversi;
- iii. Sincronizzazione e desincronizzazione devono avvenire rapidamente in modo che il riconoscimento dei concetti avvenga in modo efficiente.

Al fine di ridurre la complessità matematica, la rete di oscillatori viene di seguito descritta in forma monodimensionale come una catena ordinata di oscillatori di Wilson-Cowan. Ovviamente questa implementazione rappresenta una schematizzazione semplificata della realtà fisiologica, ma i risultati sono ugualmente esemplari del comportamento del modello in una varietà di circostanze reali.

Il modello è composto da M (numero totale di *features*) oscillatori neurali di Wilson-Cowan ognuno dei quali codifica per una singola proprietà. Si suppone che ciascun attributo non sia immediatamente presente nell'input sensoriale ma che sia stato estratto da una precedente fase di elaborazione nella neocorteccia [13,14]. L'oscillatore di Wilson Cowan è stato originariamente proposto per studiare le proprietà dinamiche di due popolazioni di neuroni interagenti (una popolazione eccitatoria e una popolazione inibitoria) in cui ognuna manda una sinapsi sia a sé stessa sia all'altra [8]. Ogni gruppo neurale può tacere, se non riceve abbastanza eccitazione, o può oscillare nella banda di frequenza gamma, se eccitato da un input sufficiente. Nel nostro modello ogni oscillatore (Figura 25) rappresenta l'attività media in una popolazione di neuroni e consiste in una connessione in retroazione (*feedback*) tra un'unità eccitatoria X e un'unità inibitoria Y , mentre l'uscita della rete corrisponde all'attività di tutte le unità eccitatorie [12]. Le equazioni del modello con oscillatori di Wilson-Cowan (Equazioni 17a, 17b, 18) rappresentano quindi i due gruppi neuronali eccitatori e inibitori, costituenti le unità di elaborazione di base. Tali equazioni differenziali sono state ampiamente utilizzate nella modellizzazione di vari processi cerebrali e sono caratterizzate da una molteplicità di parametri, consentendo una estesa gamma di dinamiche:

$$(17a) \quad \frac{d}{dt} x_i(t) = -x_i(t) + H(x_i(t) - \beta y_i(t) + E_i(t) + I_i + v_i(t) - \phi_x)$$

$$(17b) \quad \frac{d}{dt} y_i(t) = -\gamma y_i(t) + H(\alpha x_i(t) - \phi_y) + J_i(t) + Inib(t)$$

dove $H(\psi)$ rappresenta una funzione di attivazione sigmoideale così definita:

$$(18) \quad H(\psi) = \frac{1}{1 - e^{-\frac{\psi}{p}}}$$

I parametri α e β (entrambi positivi) definiscono il collegamento rispettivamente dall'unità eccitatoria a quella inibitoria e da quella inibitoria a quella eccitatoria. Nello specifico α influenza in maniera significativa l'ampiezza delle oscillazioni [3,14].

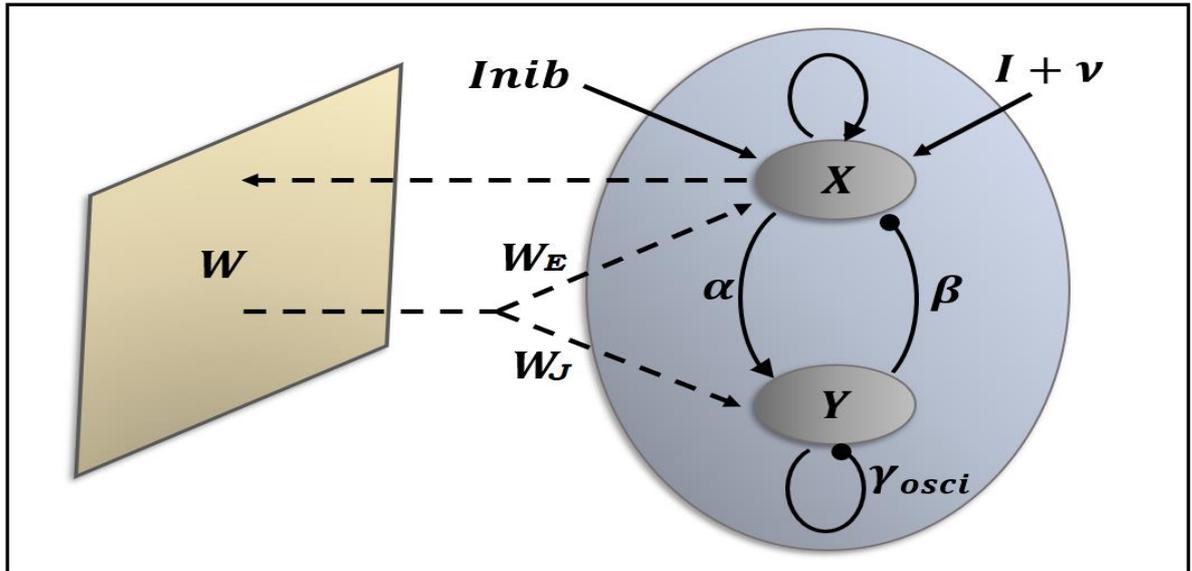


Figura 25 – Schema di un singolo oscillatore costituito da una unità eccitatoria X ed una unità inibitoria Y .

Il parametro γ_{osci} è invece il reciproco di una costante di tempo e influenza la frequenza di oscillazione. L'auto-eccitazione di X è settata al valore unitario per stabilire una scala per i pesi sinaptici. I due parametri ϕ_x e ϕ_y rappresentano due termini di offset per le funzioni sigmoidali nell'unità eccitatoria (H_x) ed inibitoria (H_y). I_i rappresenta uno stimolo esterno per l'oscillatore localizzato in posizione i -esima mentre v_i rappresenta un contributo di rumore non uniformemente distribuito. Infine, E_i e J_i identificano due termini di accoppiamento rispettivamente eccitatorio e inibitorio tra tutti gli oscillatori della rete, le cui espressioni in forma scalare sono (Equazioni 19a e 19b):

$$(19a) \quad E_i = \sum_j W_{Eij} x_j$$

$$(19b) \quad J_i = \sum_j W_{Jij} x_j$$

dove il pedice i denota la posizione dell'oscillatore post-sinaptico mentre il pedice j indica la posizione dell'oscillatore pre-sinaptico. La sommatoria si estende a tutti gli oscillatori pre-sinaptici, W_{Eij} e W_{Jij} rappresentano infatti le sinapsi eccitatorie e inibitorie tra gli M oscillatori della rete.

Per tutte le unità eccitatorie e inibitorie degli oscillatori sono state scelte condizioni iniziali casuali, nello specifico valori compresi tra zero e uno.

Una problematica dei modelli a oscillatori consiste nel fatto che i concetti tendono ad oscillare indipendentemente l'uno dall'altro e questa indipendenza può comportare la possibilità di una sincronia accidentale, che si verifica quando *features* appartenenti a concetti diversi presentano la stessa fase (risultando indistinguibili). Inoltre, all'aumentare del numero di concetti, aumenta significativamente la probabilità di correlarli erroneamente tra loro. Per questo motivo nel nostro modello si è deciso di implementare un l'inibitore globale (riportato in Figura 26) il quale calcola l'attività eccitatoria complessiva nella rete (z), ricevendo informazioni da tutti gli oscillatori, e restituisce un segnale inibitorio di entità $inib_val$ quando tale attività supera una certa soglia (th). L'equazione che caratterizza l'inibitore globale è riportata di seguito (Equazione 20) ed in generale il suo ruolo è quello di assicurare la separazione tra oggetti che sono rappresentati simultaneamente: il segnale inibitore evita che un oggetto appaia quando il precedente è ancora attivo. [3, 12,14].

$$(20) \quad Inib(t) = \begin{cases} inib_val & se \ z(t) > th \\ 0 & altrimenti \end{cases} \quad con \ z(t) = \sum_i x_i(t)$$

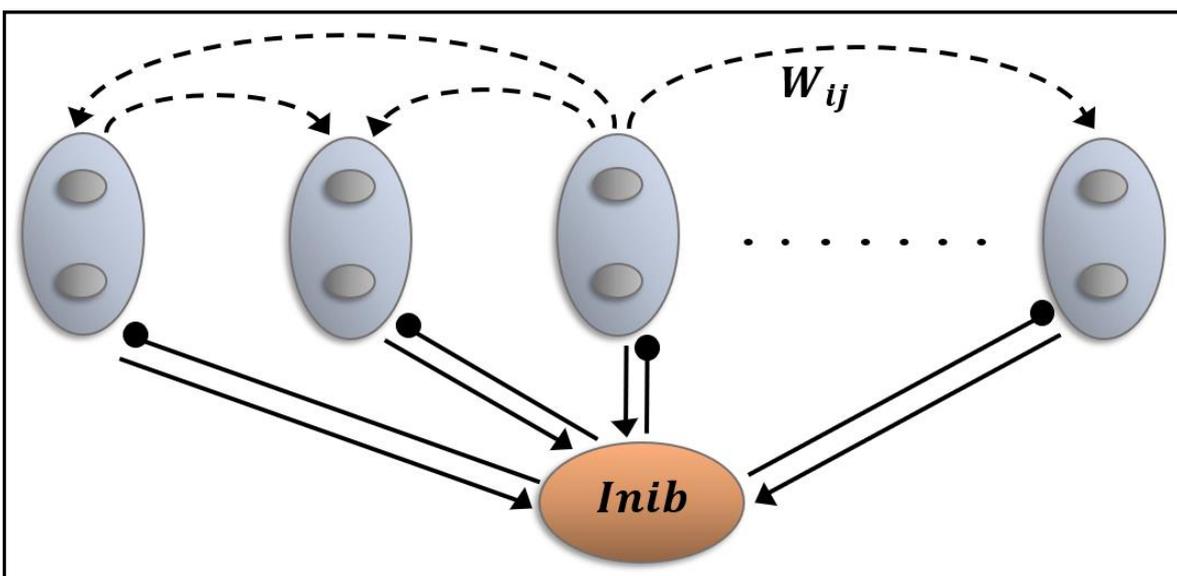


Figura 26 – Schema della rete di oscillatori con il contributo dell'inibitore globale.

4.3 Addestramento delle sinapsi

Viene eseguito un addestramento di mille epoche consecutive e le sinapsi eccitatorie W_{ij} all'interno della rete vengono addestrate con il meccanismo hebbiano a soglia variabile; tale fase ricalca esattamente l'addestramento implementato per il modello di base e quello con contesto. La differenza sostanziale rispetto all'addestramento discusso nei capitoli precedenti consiste in una successiva separazione, a partire dalla generica matrice dei pesi W_{ij} (dimensione $M \times M$), in due sottomatrici chiamate "matrice delle eccitazioni" W_E e "matrice delle inibizioni" W_J , (ancora $M \times M$). La prima (W_E) contiene i pesi sinaptici che vanno ad agire sulle unità eccitatorie degli oscillatori mentre la seconda (W_J) contiene i pesi sinaptici che vanno ad agire sulle unità inibitorie. In conseguenza a queste definizioni, le Equazioni 19a e 19b possono essere espresse in forma vettoriale (Equazioni 21a e 21b):

$$(21a) \quad E = W_E X$$

$$(21b) \quad J = W_J X$$

dove X è un vettore ($M \times 1$) contenente gli output degli oscillatori, mentre E e J sono i vettori ($M \times 1$) dei termini di accoppiamento. A partire da una soglia fissa S , le due sottomatrici appena descritte vengono implementate come segue. Si confronta il generico elemento di W_{ij} con la soglia S : se il valore dell'elemento di interesse supera il valore di soglia allora si assegna al corrispondente elemento ij -esimo della matrice W_E un valore pari a $W_{e_{val}}$ altrimenti si assegna un valore nullo. Un ragionamento analogo si potrebbe fare per la matrice delle inibizioni W_J ma, in questo lavoro di tesi, tale matrice è stata sempre considerata nulla per motivi di semplicità della trattazione e anche perché si sono ottenuti risultati incoraggianti senza bisogno di supporre un meccanismo inibitorio. Di conseguenza non risulta mai presente alcun input J sulle unità inibitorie di singoli oscillatori. Unico contributo inibitorio che giunge agli oscillatori è dunque quello fornito dall'inibitore globale.

4.4 Assegnazione dei parametri

I parametri utilizzati per l'addestramento sinaptico eseguito in condizioni non oscillanti sono elencati nella Tabella 8 mentre la successiva Tabella 9 incorpora i parametri specifici per l'implementazione degli oscillatori e dell'inibitore globale.

Significato	Simbolo	Valore
Costante di tempo	τ	3 ms
Pendenza della sigmoide	T	0.01
Posizione della sigmoide	φ	0.55
Soglia pre-sinaptica	ϑ_{pre}	0.05
Soglia post-sinaptica (valore basale)	ϑ_{post}	0.55
Saturazione della soglia post-sinaptica	$\vartheta_{post-sat}$	0.95
Velocità di incremento della soglia post-sinaptica	$\Delta\vartheta_{post}$	0.5
Rateo di apprendimento	γ_0	0.02
Massima forza sinaptica	W_{max}	0.8
Probabilità associata a <i>features</i> salienti	$P_{kn(s)}$	0.7
Probabilità associata a <i>features</i> marginali	$P_{kn(m)}$	0.35
Soglia per settare pesi sinaptici eccitatori di W_E	S	0.7
Pesi sinaptici eccitatori della matrice W_E	W_{eval}	--

Tabella 8 – Parametri utilizzati per l'addestramento sinaptico in condizioni non oscillanti.

Per primi si ripropongono i parametri relativi alla costante di tempo (τ), pendenza della sigmoide (T), ascissa centrale della sigmoide (φ) e i parametri dell'apprendimento hebbiano. Si noti che in questo caso il valore di soglia post-sinaptica (valore basale) si discosta da quello utilizzato nei capitoli precedenti. Nello specifico è stato considerato un valore più basso pari a $\vartheta_{post} = 0.55$, visto il numero minore di *features* utilizzate e quindi vista la minore probabilità che una proprietà marginale divenga casualmente saliente. Inoltre, si ritrovano invariati: il rateo di apprendimento hebbiano γ_0 , il massimo valore sinaptico W_{max} e le frequenze di occorrenza delle *features* salienti $P_{kn(s)}$ e di quelle marginali $P_{kn(m)}$.

Gli ultimi due parametri della Tabella 8 sono quelli corrispondenti alla soglia (S) per la costruzione della matrice delle eccitazioni e infine il valore assunto dai pesi sinaptici eccitatori della matrice W_E , identificato dal parametro $W_{e_{val}}$. Per questo ultimo parametro non viene al momento specificato il valore assunto poiché, osservando i dati numerici ottenuti durante le simulazioni, si è visto che esso deve essere ottimizzato a seconda del numero di caratteristiche che si considerano.

Significato	Simbolo	Valore
Forza sinaptica da unità eccitatoria a inibitoria	α	0.3
Forza sinaptica da unità inibitoria a eccitatoria	β	2,5
Pendenza delle funzioni sigmoidali H	p	0.025
Offset di H_x	ϕ_x	0.7
Offset di H_y	ϕ_y	0.15
Inverso della costante di tempo	γ_{osci}	--
Deviazione standard del rumore	σ	--
Valore dell'inibizione globale	$inib_val$	--
Soglia dell'inibitore globale	th	0.5

Tabella 9 – Parametri utilizzati per l'implementazione di oscillatori e inibitore globale.

Per quanto riguarda i parametri relativi alla rete di oscillatori (Tabella 9), si riportano le forze sinaptiche tra le due diverse unità costituenti l'oscillatore (α e β), la pendenza p e i rispettivi valori di offset (ϕ_x e ϕ_y) delle funzioni sigmoidali. Vengono inoltre elencati i due parametri che rappresentano l'inverso della costante di tempo (γ_{osci}) e la deviazione standard della componente di rumore (σ). Infine, gli ultimi due parametri presenti in tabella caratterizzano l'inibitore globale, in particolare l'entità della sua inibizione ($inib_val$) e la soglia oltre la quale si attiva il contributo inibitorio (th). I valori assunti rispettivamente da γ_{osci} , σ e $Inib$ verranno specificati di seguito, sempre in riferimento al numero di *features* che compongono la rete.

4.5 Risultati

In questa sezione vengono presentati i risultati di alcune simulazioni riguardanti attività di riconoscimento di concetti. A differenza di quanto ottenuto con le reti precedenti, basate sulle dinamiche di attrattori, il compito di riconoscimento di concetti si ritiene svolto correttamente se:

- Oscillatori che codificano *features* salienti relative ad un singolo concetto si mettono ad oscillare tra loro in modo sincrono;
- Oscillatori che codificano *features* salienti appartenenti a concetti diversi oscillano ma in modo tra loro desincronizzato per consentire la netta distinzione tra un concetto e l'altro.
- Oscillatori che codificano *features* marginali devono attivare l'oscillazione delle proprietà salienti ma non quella di altre proprietà marginali.

Per testare le capacità del modello si utilizzano due reti di oscillatori piuttosto semplici (Modello I e Modello II descritte in dettaglio nei paragrafi seguenti), le quali sono state realizzate sulla base di un numero limitato di *features* appartenenti alla tassonomia di oggetti descritta nel primo capitolo di questo elaborato. Risultati del tutto simili si ottengono per reti analoghe implementate a partire dalla tassonomia di animali. Le simulazioni sono riportate, nelle figure a seguire, in termini di andamento nel tempo della sola unità eccitatoria X dell'oscillatore. Equivalenti simulazioni e risultati si ottengono per le unità inibitorie Y ma, per brevità della trattazione, non verranno riportate. In ascissa si individua l'intervallo temporale mentre in ordinata si individuano i valori di attività neurale normalizzata dell'unità eccitatoria. In tutte le simulazioni effettuate, la rete è partita da una condizione iniziale casuale per cui i risultati che verranno di seguito presentati saranno accomunati dalla presenza di un breve transitorio iniziale prima della fase di assestamento del comportamento oscillatorio di interesse.

4.5.1 Risultati delle simulazioni del Modello I

Le prime simulazioni sono state eseguite sul Modello I (schema completo in Figura 27) costituito da un totale di 30 *features* che individuano tre diversi oggetti (“pentola”, “forchetta”, “tazzina”). Ogni oggetto è identificato da nove proprietà salienti ed una sola proprietà marginale e non è presente alcuna caratteristica condivisa. Le simulazioni eseguite sul Modello I hanno come obiettivo principale quello di discriminare correttamente i tre concetti, desincronizzandoli tra loro, ma anche il compito di mostrare il diverso ruolo tra *features* salienti e marginali. Per questo modello è stato necessario settare il valore della frequenza di oscillazione utilizzando un valore del relativo parametro $\gamma_{osci} = 0.5$ mentre la deviazione standard del contributo di rumore risulta $\sigma = 0.9$. Questa scelta relativa al contributo di rumore si deve al fatto che si è voluto evitare che, una volta eccitate, le *features* rimanessero in uno stato saturazione (privo di oscillazioni). Infine, sono stati assegnati opportuni valori sia per il parametro relativo ai pesi sinaptici eccitatori della matrice W_E sia per l'entità della inibizione globale. A seguito di varie prove empiriche, a tali parametri sono stati assegnati i seguenti valori ottimizzati: $We_{val} = 0.65$ e $inib_val = 0.23$.



Figura 27 – Schema completo del Modello I basato su *features* di tre diversi oggetti. Le prime nove *features* di ogni oggetto sono salienti mentre l'ultima di ognuno è marginale.

Una prima simulazione analizza il caso in cui in input siano fornite tre *features* distintive salienti, una sola per ogni oggetto. La rete deve essere in grado di evocare tutte le proprietà salienti dei tre oggetti (mettendole ad oscillare in sincronia) e al tempo stesso, grazie al contributo dell'inibizione globale, riuscire a desincronizzare tra loro (seppure non in maniera perfetta) caratteristiche appartenenti a concetti differenti. Di contro, nessuna proprietà marginale deve essere evocata. A titolo di esempio si considera la simulazione riportata nella Figura 28 in cui si può osservare una evidente sincronizzazione tra *features* salienti appartenenti allo stesso oggetto (nove proprietà di “pentola” in blu, nove di “forchetta” in rosso e altre nove di “tazzina” in verde) ed allo stesso tempo una buona desincronizzazione tra i tre oggetti, consentendo una immediata discriminazione tra essi. In basso, nella stessa figura, si intravedono le tre proprietà marginali che invece non vengono evocate spontaneamente: esse non presentano infatti alcun andamento oscillatorio, se non per un minimo contributo di rumore che crea una lievissima fluttuazione. In generale per tutte le *features* evocate si possono osservare oscillazioni caratterizzate da un transitorio iniziale abbastanza marcato, il quale è dovuto al contributo di rumore random.

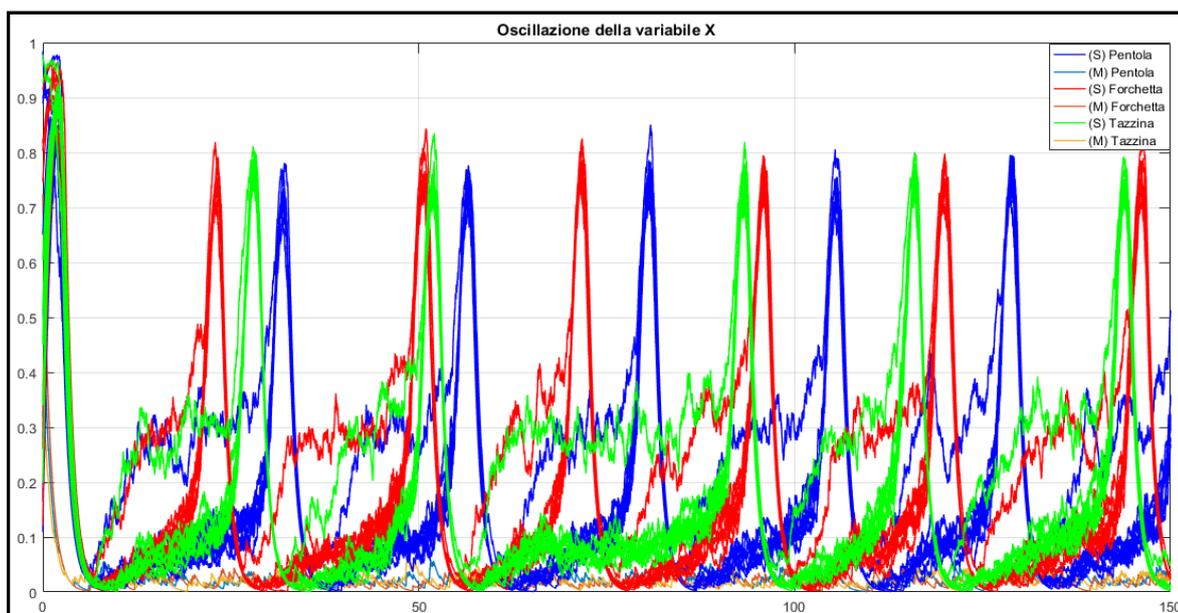


Figura 28 – Oscillazione delle unità eccitatorie del Modello I quando in input sono presenti *features* distintive salienti, una per ciascun oggetto.

Una variante interessante di questa prima simulazione si può osservare nella Figura 29, la quale mostra l'andamento oscillatorio delle unità eccitatorie quando in input vengono fornite due *features* salienti del primo oggetto e una sola proprietà saliente per gli altri due. Il modello riesce ancora ad evocare tutte le caratteristiche salienti facendo oscillare in ottimo sincronismo quelle del medesimo concetto, sfasandole rispetto alle proprietà degli altri due. Ora però l'oggetto che presenta un numero maggiore di input si presenta con maggiore frequenza. Infatti, vediamo che le oscillazioni delle *features* appartenenti al concetto di "pentola", in blu, dal momento che tale oggetto ha ricevuto input doppio rispetto a quello fornito agli altri due utensili da cucina, si presentano con più assiduità rispetto alle oscillazioni delle caratteristiche di "tazzina" e "forchetta". Inoltre, la desincronizzazione fra cucina e forchetta non è sempre ottimale.

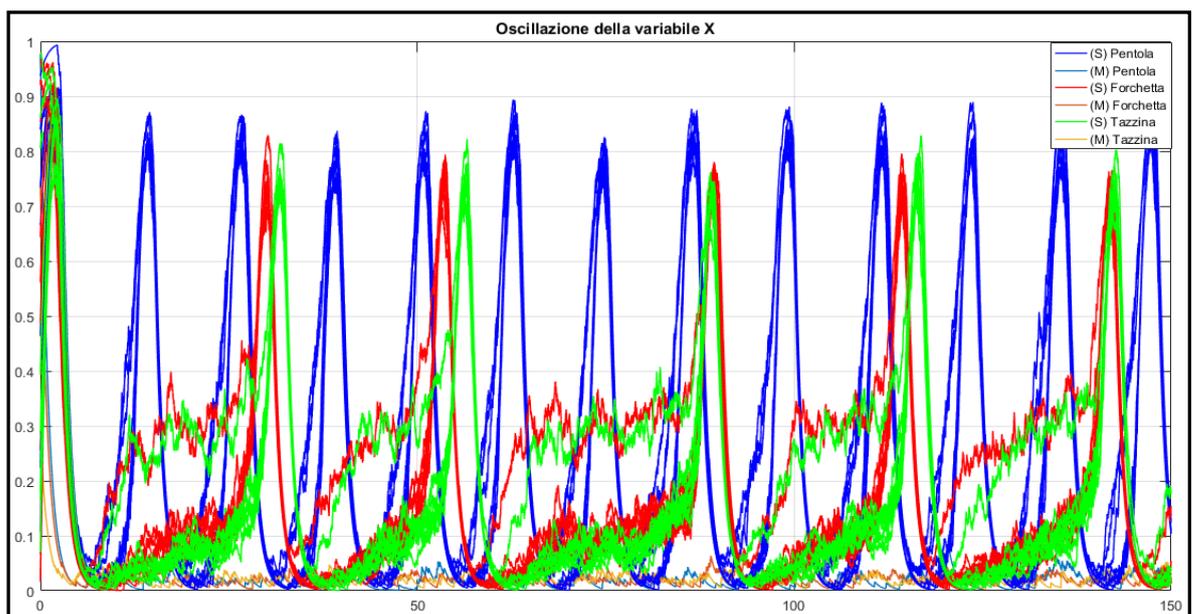


Figura 29 – Oscillazione delle unità eccitatorie del Modello I quando in input sono presenti quattro *features* distintive salienti: due per “pentola”, una per “forchetta” ed una per “tazzina”.

Si considera ora il comportamento della rete di oscillatori a fronte di input marginali (si veda la Figura 30): in particolare viene data in ingresso una proprietà marginale per ciascun concetto. Il modello evoca le caratteristiche salienti di tutti e tre gli utensili e le fa oscillare in maniera distinta ma tali oscillazioni, a parità di ampiezza, risultano meno frequenti rispetto ai casi precedenti. Inoltre, anche le tre *features* marginali date in input, evidenziate nel riquadro in nero in Figura 30, entrano in uno stato oscillatorio ma caratterizzato da una ampiezza molto minore (si presume dovuta al loro carattere marginale) rispetto a quella delle proprietà salienti evocate

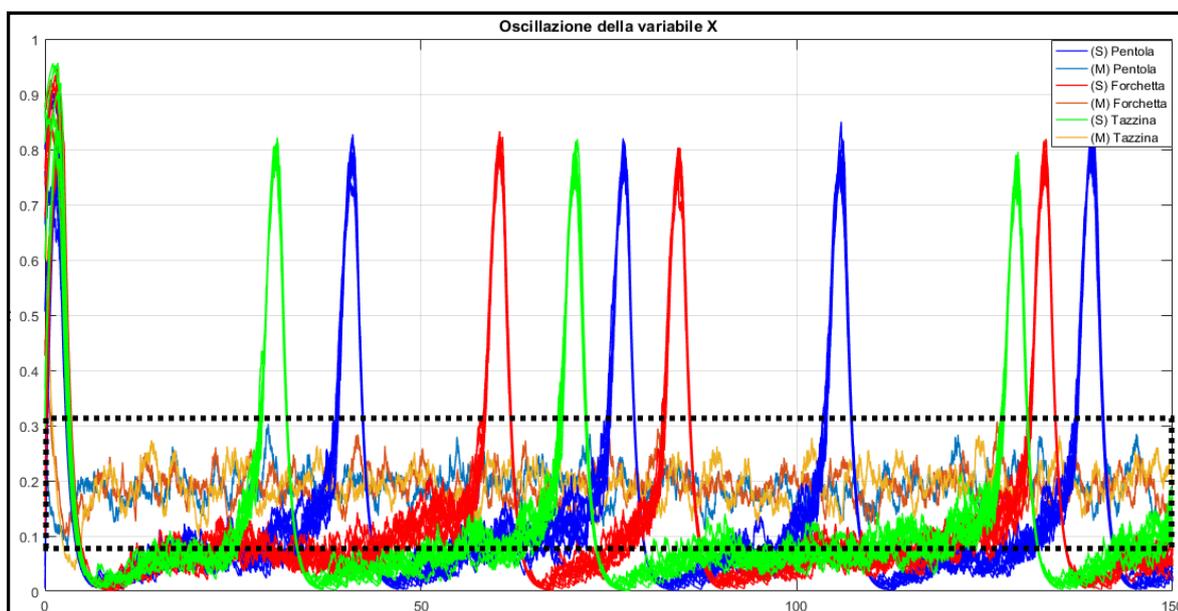


Figura 30 – Oscillazione delle unità eccitatorie del Modello I quando in input sono presenti *features* marginali, una per ciascun oggetto.

Infine, per completezza della trattazione, si riporta il caso semplificato in cui un utensile dei tre non viene in alcun modo stimolato. Questo equivale a fornire in input solo *features* appartenenti a due soli oggetti, ad esempio una proprietà saliente di “pentola” ed una saliente di “forchetta”. Quello che si ottiene è un comportamento quasi del tutto silente degli oscillatori che codificano le caratteristiche del concetto non stimolato mentre le proprietà salienti degli altri due risultano attivarsi in maniera oscillante. Un esempio di questo comportamento si può osservare nella Figura 31, in cui l’oggetto “tazzina” (in verde) non viene stimolato da nessun input esterno.

Quello che si ottiene dunque è la discriminazione, grazie a due fasi di oscillazione distinte delle caratteristiche, dei due soli oggetti stimolati: “pentola” (in blu) e “forchetta” (in rosso). Ragionamento analogo per la stimolazione di un singolo concetto per cui gli altri due resteranno silenti.

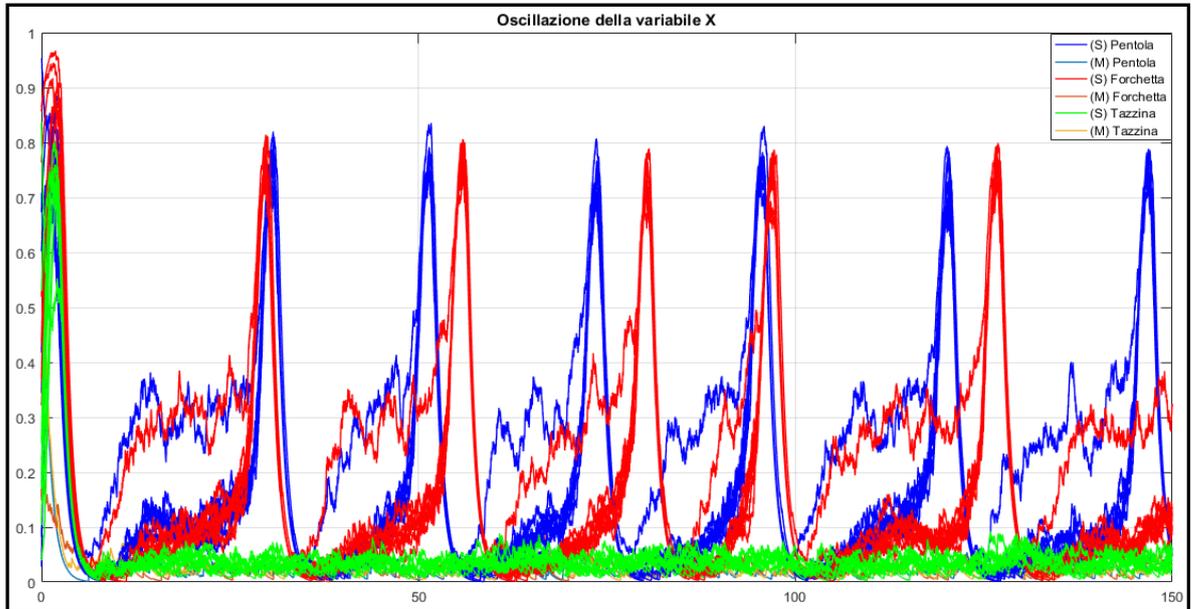


Figura 31 – Oscillazione delle unità eccitatorie del Modello I quando in input sono presenti due sole features salienti appartenenti ai concetti di “pentola” e “forchetta”.

Discussione

Le simulazioni sono state ripetute per tutti e tre gli oggetti presenti variando la scelta delle *features* e il numero di attributi forniti in input e si è generalmente osservato un corretto funzionamento del Modello I. Riassumendo il significato dei dati numerici ottenuti dalle questo primo set di simulazioni, si può concludere che la rete basata su oscillatori è in grado di svolgere correttamente il compito di riconoscimento del concetto; questo accade per qualsiasi proprietà distintiva fornita in input, sia essa marginale o saliente. Risulta inoltre sufficiente una singola *feature* distintiva per ogni concetto per evocare correttamente tutte le caratteristiche salienti. A differenza dei modelli discussi nei capitoli precedenti, attraverso gli oscillatori e il contributo di una inibizione globale, le caratteristiche salienti appartenenti a concetti differenti vengono evocate in maniera simultanea ma risultano differenziate grazie alla desincronizzazione.

Un limite riscontrato nelle simulazioni effettuate con il Modello I riguarda il sincronismo non così robusto nel tempo delle proprietà appartenenti ad un singolo concetto: all'aumentare del periodo temporale si assiste infatti ad uno sfasamento, seppure lieve, delle oscillazioni. Si presume che questo fenomeno sia dovuto al numero elevato di *features* (nove per ciascun concetto) che devono sincronizzarsi tra loro. Infine, un ulteriore limite di questa rete consiste nel fatto che le oscillazioni dei tre oggetti non presentino un *pattern* di alternanza regolare (in riferimento alla Figura 28).

4.5.2 Risultati delle simulazioni del Modello II

Una seconda serie di simulazioni è stata effettuata sul Modello II (schema completo in Figura 32), costituito da un totale di 17 *features* che individuano ancora i tre oggetti già presentati nel Modello I. Ogni utensile in questo caso è identificato da sole sette proprietà di cui due condivise e cinque distintive (nello specifico tre salienti e due marginali). Le simulazioni eseguite sul Modello II hanno come obiettivo ancora quello di separare i diversi concetti grazie alla desincronizzazione ma anche quello di mostrare il comportamento delle *features* condivise.

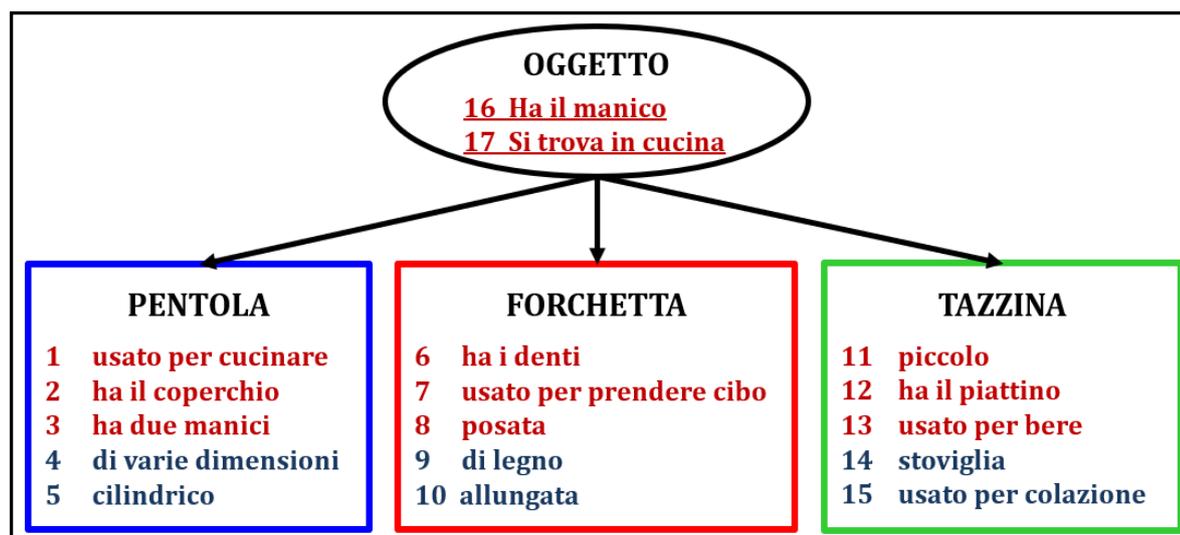


Figura 32 – Schema completo del Modello II basato su *features* di tre diversi oggetti. Per ogni oggetto si hanno due *features* condivise e cinque distintive (di cui le prime tre salienti e le ultime due marginali).

Per questo modello è stato necessario impostare il valore del parametro γ_{osci} ad un valore più basso rispetto al modello precedente ($\gamma_{osci} = 0.2$), che corrisponde ad una costante di tempo maggiore. Per quanto riguarda il contributo di rumore, si è invece scelta una deviazione standard inferiore ($\sigma = 0.1$) rispetto a quella utilizzata nel modello precedente, sufficiente per evitare la condizione di saturazione delle *features*. Infine, sono stati assegnati opportuni valori sia per il parametro relativo ai pesi sinaptici eccitatori della matrice W_E sia per l'entità dell'inibizione globale. Dopo varie prove si è scelto un valore ottimale per l'inibizione ($inib_val = 0.3$) e per i pesi sinaptici ($W_{eval} = 2.7$).

Si analizza dapprima il caso in cui in input siano fornite alla rete tre caratteristiche salienti, una per ciascun oggetto. Come mostrato nella Figura 33, la rete evoca tutte le proprietà salienti dei tre oggetti: questa volta prevedono sia proprietà distintive salienti (in blu le *features* di “pentola”, in rosso quelle di “forchetta” ed in verde quelle di “tazzina”) sia le uniche due caratteristiche condivise tra i tre oggetti (riportate in nero). Al tempo stesso, grazie al contributo dell'inibizione globale, si desincronizzano tra loro caratteristiche appartenenti ai tre differenti utensili. Non viene, al contrario, evocata nessuna proprietà marginale.

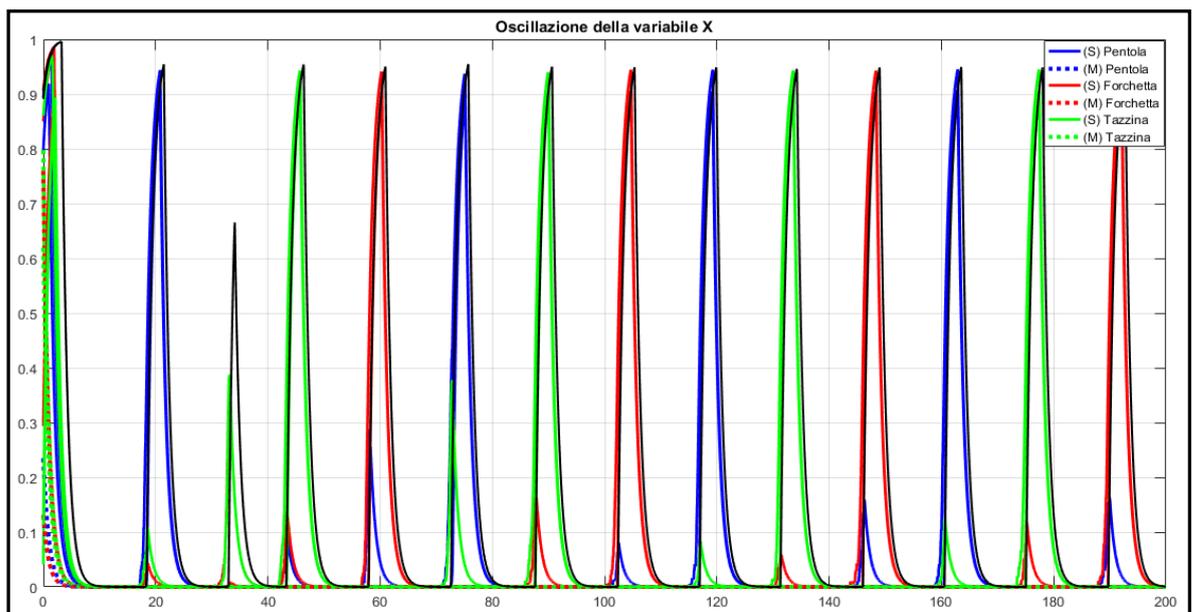


Figura 33 – Oscillazione delle unità eccitatorie del Modello II quando in input sono presenti *features* distintive salienti, una per ciascun oggetto.

Come ultimo caso si è voluta dare in pasto alla rete una delle due proprietà condivise (medesimo risultato si ottiene se in input viene fornita l'altra o entrambe). Il comportamento della rete è quello di evocare e far oscillare in sincrono solamente le uniche due proprietà condivise da tutti e tre gli oggetti, comportamento che si può osservare nella Figura 34. Nell'esempio di seguito riportato, a fronte dell'input condiviso "ha il manico" la rete evoca anche il secondo attributo condiviso dai tre utensili, ovvero "si trova in cucina" facendo oscillare con stessa fase e medesima ampiezza i due soli attributi condivisi, non evocando alcuna altra proprietà del modello.

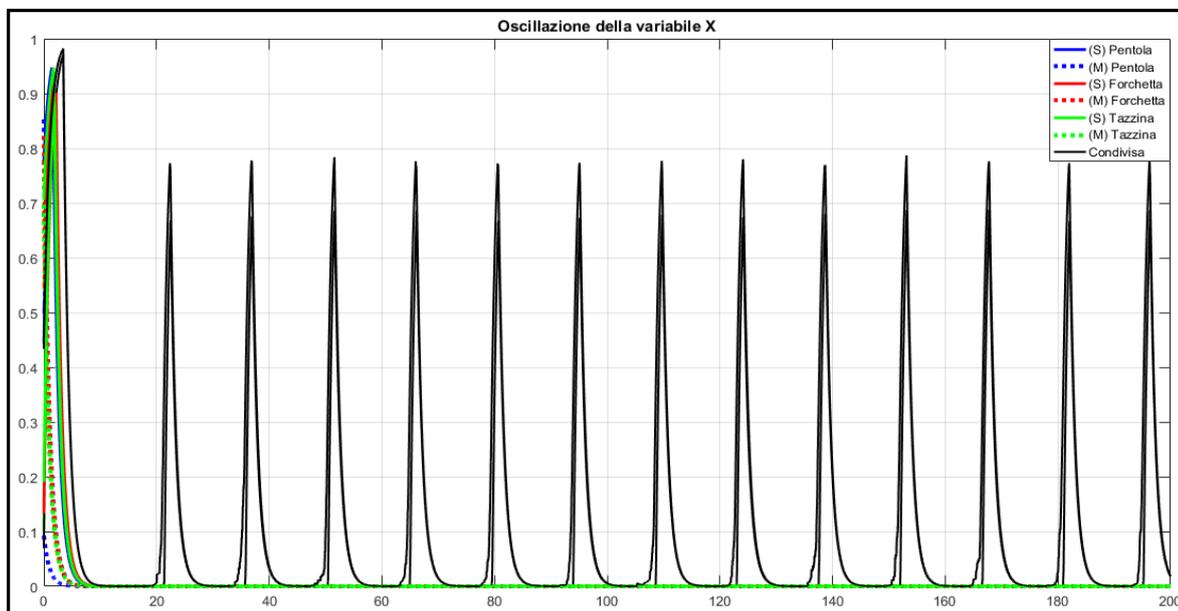


Figura 34 – Oscillazione delle unità eccitatorie del Modello II quando in input è presente una feature condivisa.

Discussione

Anche per questo secondo set di simulazioni, le prove sono state ripetute per tutti e tre gli oggetti presenti variando la scelta delle *features* e il numero di attributi forniti in input e si è generalmente osservato un corretto funzionamento del Modello II. A differenza del modello precedente, il sincronismo risulta più preciso e stabile nel tempo, probabilmente grazie alla presenza di un minor numero di proprietà che devono oscillare in sincrono.

Questa rete basata su oscillatori è in grado di svolgere correttamente il compito di riconoscimento del concetto se in input vengono fornite *features* distintive mentre si può osservare che il riconoscimento di una semplice categoria avviene solo se sono utilizzate come input le caratteristiche condivise. In questo caso la rete simula correttamente la categoria corrispondente, facendo oscillare le proprietà condivise in modo sincrono, ma non evoca mai le caratteristiche distintive dei singoli concetti (quindi non riesce a distinguere un singolo oggetto dall'altro).

Conclusioni

Nel presente elaborato, ho fatto uso di un modello neuro-computazionale per migliorare la comprensione dell'organizzazione della memoria semantica. L'obiettivo principale era quello di indagare il ruolo di diversi tipi di caratteristiche nella rappresentazione concettuale, questione centrale per lo studio della memoria semantica. Innanzitutto, in accordo con diverse idee in letteratura (*Barsalou, 1982*) il modello assume che, quando riconosciamo un concetto, recuperiamo le sue *features* dominanti, cioè ricostruiamo tutti i frammenti di informazione frequentemente presenti nella nostra esperienza passata. Quindi, non manipoliamo semplicemente simboli, bensì ricostruiamo l'esperienza.

Le teorie della rappresentazione concettuale basata su caratteristiche suggeriscono che le *features* semantiche possano spiegare la formazione di categorie attraverso proprietà che sono condivise da molti concetti, mentre le caratteristiche distintive consentano l'identificazione di concetti strettamente correlati.

Nel presente lavoro, sono state utilizzate proprietà generali per rappresentare la semantica di concetti reali (animali e oggetti) selezionati a partire da un ampio *set* di dati iniziali. Questa scelta, ovviamente, ha l'enorme vantaggio di consentire un confronto diretto tra risultati del modello e risultati di *test* psicologici. Le singole *features* utilizzate non sono percezioni elementari, bensì elementi complessi che consentono di ottenere una rappresentazione modale e organizzata della rete; ciò è in accordo con gli attuali modelli neurali di memoria semantica, i quali suggeriscono una distribuzione complessa e topologicamente organizzata su più aree del cervello. Inoltre, nel modello, le unità non rappresentano singoli neuroni, piuttosto una popolazione di neuroni che partecipa alla rappresentazione della medesima caratteristica. La loro attivazione è il prodotto di una fase di pre-elaborazione, che estrae le informazioni principali da dati esterni e decide sulla presenza/assenza di una proprietà.

Per semplicità, questi passaggi di pre-elaborazione non sono direttamente inclusi nelle equazioni ma sono stati assunti impliciti al modello di base.

Al contrario delle prime versioni, nel modello attuale è stata utilizzata una regola hebbiana di apprendimento sinaptico la quale prevede una soglia post-sinaptica variabile in funzione della frequenza presinaptica (riflettendo quanti concetti condividono una caratteristica). Questa scelta è stata fatta al fine di pervenire ad un concetto di salienza, diverso per le proprietà distintive e per quelle condivise. In particolare, è stato possibile evitare che una *feature* condivisa da molti membri della stessa categoria divenga dominante e sia erroneamente attribuita all'intera categoria. L'uso di una soglia post-sinaptica variabile ha quindi come obiettivo quello di controllare la salienza, in modo che la salienza per le caratteristiche condivise in una categoria segua uno schema diverso rispetto a quella di proprietà distintive. Vale la pena notare che la regola di apprendimento così implementata risulta biologicamente plausibile; in particolare è una regola locale, cioè sfrutta le informazioni già presenti (ovvero i valori delle attività pre-sinaptiche e post-sinaptiche e la forza della sinapsi).

Il modello suggerisce dunque che la regola hebbiana, con soglia post-sinaptica variabile, possa fornire una rappresentazione semantica affidabile di concetti che sfruttano le statistiche delle *features* di input dal momento che presuppone un adattamento automatico alla statistica di tali proprietà. Per quanto riguarda il ruolo delle caratteristiche distintive, esse risultano cruciali al fine di discriminare concetti strettamente correlati; tipicamente vengono utilizzate per distinguere concetti della stessa categoria semantica, dal momento che tali *features* si presentano solo in uno o in pochissimi concetti. Non tutte le caratteristiche distintive, tuttavia, hanno la stessa rilevanza per la rappresentazione di un concetto. Alcune proprietà distintive risultano marginali, in quanto sono riportate di rado nei compiti di elencazione e presumibilmente non svolgono un ruolo di primo piano nella rappresentazione concettuale. Le caratteristiche distintive possono quindi essere suddivise in caratteristiche salienti e marginali, dove le prime risultano più importanti nell'identificazione di un concetto.

Tutti i risultati ottenuti dalle simulazioni confermano che la rete può discriminare tra *features* salienti e marginali.

In questo lavoro, sono state utilizzate due frequenze di occorrenza costanti per ciascuna caratteristica all'interno di un concetto (35% per le proprietà marginali e 70% per quelle salienti), ottenendo così una rappresentazione semantica finale fissa. Questa scelta è stata motivata principalmente dalla semplicità, per facilitare l'analisi dei risultati. Il modello potrebbe in realtà essere compatibile con una rappresentazione della semantica flessibile, basata sull'esperienza individuale. Come possibile sviluppo futuro si potrebbe dunque addestrare il modello assumendo un'esperienza flessibile, in cui la frequenza di occorrenza della caratteristica cambi da persona a persona, o eventualmente nel tempo.

Un altro possibile sviluppo per ulteriori studi potrebbe essere quello di considerare più fattori che incidono sulla dominanza. Occorre tenere presente che nelle simulazioni attuali l'unico fattore che influenza la dominanza è la frequenza di occorrenza. Altri elementi, non considerati qui (come l'impatto emotivo di una caratteristica o il suo ruolo nel funzionamento dell'oggetto) possono influenzare la dominanza. Le caratteristiche ad alto impatto emotivo o funzionale, ad esempio, possono diventare dominanti anche se si verificano raramente. Tuttavia, tenere conto di questi aspetti richiederebbe una diversa strategia di apprendimento (probabilmente una diversa costante di apprendimento nella regola di Hebb), problematica che non è stata affrontata nell'analisi da me condotta.

Ritengo necessario riportare un limite del modello attuale, il quale è legato alle caratteristiche condivise tra due concetti ed ha provocato un comportamento anomalo ma occasionale della rete: se una *feature* condivisa è saliente per un concetto e marginale per un altro, a volte, il modello tende a considerarla dominante per entrambi. Tale anomalia si è però verificata con una frequenza assai poco significativa e sembrerebbe essere legata alla iniziale creazione casuale di sinapsi.

Inoltre, si vuole riconoscere una grande semplificazione del presente modello relativamente alla sua struttura. Una rete semantica più completa dovrebbe incorporare più di un singolo livello di elaborazione (ad esempio diversi strati modali, seguiti da strati multimodali), estraendo *features* elementari dalle regioni sensoriali, motorie o emotive.

Ulteriori aspetti della memoria semantica sono stati analizzati nell'ambito di questo elaborato ed alcuni di essi potrebbero rappresentare spunti futuri per l'implementazione di modelli neuro-computazionali più complessi.

Innanzitutto, si è voluta simulare la dipendenza della semantica di un concetto dal contesto, utilizzando un addestramento che ha fatto uso di probabilità condizionate tra vari concetti. Sebbene siano presenti alcune informazioni sempre dominanti per un dato concetto (indipendenti), si è potuto osservare come le proprietà *context-dependent* (CD), partecipino alla formazione concettuale solo in situazioni pertinenti e dipendenti dalle contingenze. Inoltre, le simulazioni relative al contesto hanno evidenziato come *features* CD si comportino in maniera totalmente autonoma fra un contesto e l'altro, mostrando che due o più contesti simultanei non si influenzano tra loro.

In secondo luogo, è stata considerata la sincronizzazione delle oscillazioni in banda gamma, implementando una rete di oscillatori di Wilson Cowan, attraverso la quale si ottiene la rappresentazione simultanea in memoria di concetti diversi, anche a partire da informazioni in ingresso parziali, favorendo la soluzione del problema di "*binding & segmentation*".

Le simulazioni ottenute da due differenti modelli (Modello I e Modello II) hanno mostrato come una rete di oscillatori sia in grado di svolgere correttamente il compito di riconoscimento del concetto e la formazione di semplici categorie. In aggiunta a questo, le *features* salienti appartenenti a concetti differenti vengono evocate in maniera simultanea ma risultano differenziate grazie alla desincronizzazione operata da un inibitore globale. Tuttavia, un limite di queste reti è che, per ottenere una desincronizzazione corretta, risulta necessaria una selezione ad hoc dei parametri degli oscillatori e dell'inibitore.

I due modelli con oscillatori hanno inoltre mostrato dei limiti relativi alla robustezza nel tempo della sincronizzazione per un elevato numero di caratteristiche.

Complessivamente per tutte le reti analizzate in questo lavoro di tesi occorrerà, in futuro, ottimizzare un *set* di parametri standard da poter utilizzare con qualsiasi tipo di tassonomia.

In conclusione, posso affermare che i modelli da me studiati hanno generalmente rispecchiato gli obiettivi prefissati a inizio lavoro e possono rappresentare uno strumento promettente sia per lo studio della memoria semantica, sia per possibili applicazioni in ambito neurologico.

Appendice

Per completezza della trattazione, all'interno di questa sezione ho voluto riportare una versione schematica delle due tassonomie semantiche rispettivamente di animali e oggetti. Entrambe le rappresentazioni schematiche sono accomunate dai seguenti requisiti:

- All'interno degli ovali sono riportate le *features* totalmente condivise, le quali costituiscono una categoria.
- All'interno dei riquadri rettangolari sono riportate le *features* distintive appartenenti ad un singolo animale/oggetto e le proprietà parzialmente condivise con altri animali/oggetti;
- Il colore rosso identifica le *features* salienti mentre il colore blu è utilizzato per identificare le caratteristiche marginali;
- Il testo sottolineato indica le *features* parzialmente condivise da più animali/oggetti mentre il testo non sottolineato identifica tutte le caratteristiche distintive.
- Ogni caratteristica è preceduta da un numero che individua la posizione della singola *feature* all'interno della rete così da rendere più facile la trattazione e immediata la visualizzazione durante le simulazioni. Per motivi implementativi la prima caratteristica è associata alla seconda posizione all'interno della rete (in modo da permettere una più facile visualizzazione grafica).
- Accanto ad ogni concetto è riportato un numero in nero, posto entro parentesi tonde, che corrisponde alla posizione assunta dal singolo concetto in modo da ottenere un elenco di proprietà raggruppate in modo contiguo.

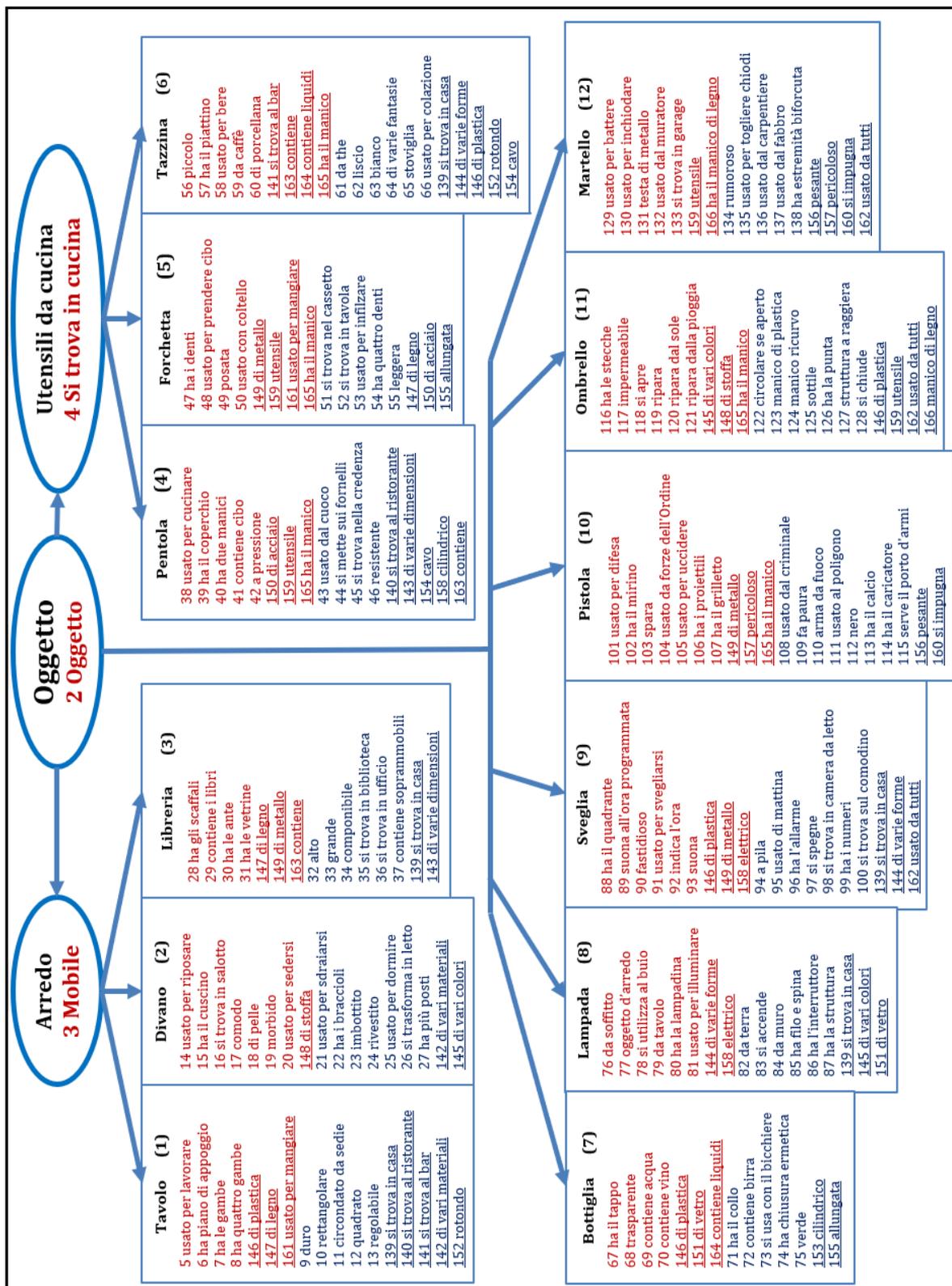


Figura A.2 – Schema relativo alla tassonomia semantica degli oggetti.

Bibliografia

- [1] Barsalou Lawrence W. (1982), “Context-independent and context-dependent information in concepts”, *Memory & Cognition*, Vol. 10 (1), pp. 82-93.
- [2] Binder J.R., Desai R.H. (2011), Department of Neurology, Medical College of Wisconsin, USA. “The neurobiology of semantic memory”, *Cognitive Sciences*, pp. 1-10.
- [3] Campbell S., Wang DL. (1996), “Synchronization and desynchronization in a network of locally coupled Wilson-Cowan oscillators”, *IEEE transactions on neural networks*, Vol. 7 (3), pp. 541-554.
- [4] Catricalà E, Ginex V, Dominici C, Cappa S (2015a), “A new comprehensive set of concept feature norms”, Special Issue in Honour of J. Frederico Marques, *Revista Portuguesa de Psicologia*, Vol. 44, pp. 111–120.
- [5] Eiling Yee, Sharon L. Thompson-Shill (2016), “Putting concepts into context”, *Psychonomic Bulletin & Review*, Vol. 23 (4), pp. 1015-1027.
- [6] Lebois Lauren A. M., Wilson-Mendenhall Christine D, Barsalou Lawrence W. (2015), “Are Automatic Conceptual Cores the Gold Standard of Semantic Processing? The Context-Dependence of Spatial Meaning in Grounded Congruency Effects”, *Cognitive Science*, Vol. 39, pp. 1764-1801.

- [7] Levine D.M., Krehbiel T.C., Berenson M.L., *Statistica*, Milano, APOGEO, 2006, (seconda edizione), pp. 117-150.
- [8] Ursino M. (2017), “*Dispense di Sistemi Neurali*”, Facoltà di Ingegneria, Corso di Laurea Magistrale in Ingegneria Biomedica, Università di Bologna.
- [9] Ursino M., Cuppini C., F. Cappa S., Catricalà E. (2018), “A feature-based neurocomputational model of semantic memory”, *Cognitive Neurodynamics*, Vol. 12 (6), pp. 525–547.
- [10] Ursino M., Cuppini C., Magosso E. (2015), “A neural network for learning the meaning of objects and words from a featural representation”, *Neural Networks*, Vol. 63, pp. 234–253.
- [11] Ursino M., Cuppini C., Magosso E. (2013), “The formation of categories and the representation of feature saliency: analysis with a computational model trained with an Hebbian paradigm”, *Journal of Integrative Neuroscience*, Vol. 12 (04), pp. 401–425.
- [12] Ursino M., Cuppini C., Magosso E. (2011), “An integrated neural model of semantic memory, lexical retrieval and category formation, based on a distributed feature representation”, *Cognitive Neurodynamics*, Vol. 5 (2), pp. 183–207.
- [13] Ursino M., Magosso E., Cuppini C. (2009), “Recognition of abstract objects via neural oscillators: interaction among topological organization, associative memory and gamma band synchronization”, *IEEE transactions on neural networks*, Vol. 20 (2), pp. 316-335.

- [14] Ursino M., Magosso E., La Cara G.E., Cuppini C. (2006), “Object segmentation and recovery via neural oscillators implementing the similarity and prior knowledge gestalt rules”, *BioSystems*, Vol. 85, pp. 201-218.

- [15] Wang DL. et al. (2005), “The time dimension for scene analysis”, *IEEE transactions on neural networks*, Vol. 16 (6), pp. 1401-1426.