

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:
STUDIO DEL SINCRONISMO IN BANDA GAMMA
E DIPENDENZA CONTESTUALE**

Tesi In Sistemi Neurali LM

Relatore:

Prof. Mauro Ursino

Presentato da:

Silvia Mastromarco

II Sessione

Anno Accademico 2018/2019

“In the brain, you have connections between the neurons called synapses, and they can change. All your knowledge is stored in those synapses”

G. Hinton

INDICE

INTRODUZIONE	1
-------------------------------	----------

CAPITOLO 1

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

1.1 La memoria semantica	5
------------------------------------	---

1.2 Organizzazione delle conoscenze	5
---	---

1.3 Modelli di memoria semantica	6
--	---

1.4 Database e tassonomie semantiche	8
--	---

1.4.1 Tassonomia Oggetti	10
------------------------------------	----

1.4.2 Tassonomia Animali.	13
-----------------------------------	----

CAPITOLO 2

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

2.1 Modello unidimensionale.	17
--------------------------------------	----

2.2 Equazioni del modello.	18
------------------------------------	----

2.3 Addestramento sinaptico.	19
--------------------------------------	----

2.3.1 Regola di Hebb.	19
-------------------------------	----

2.3.2 Fase di addestramento.	21
--------------------------------------	----

2.4 Assegnazione parametri	22
--------------------------------------	----

2.5 Risultati.	27
------------------------	----

2.5.1 Risultati addestramento sinaptico.	27
--	----

2.5.2 Risultati simulazioni di riconoscimento animali . .	31
---	----

CAPITOLO 3

MODELLO DI MEMORIA SEMANTICA CON OSCILLATORI .	37
---	-----------

3.1 “ <i>Binding and segmentation</i> ” e sincronismo in banda γ . . .	37
---	----

3.2 Equazioni associate all’oscillatore di Wilson – Cowan	38
---	----

3.3 Sincronismo tra 10 oscillatori: scelta parametri e risultati	40
--	----

3.4 Modelli con oscillatori	42
---------------------------------------	----

3.4.1	Rete a 1 animale e 10 proprietà	42
3.4.1.1	Assegnazione parametri	43
3.4.1.2	Risultati addestramento e simulazioni	44
3.4.2	Rete a più animali: riconoscimento simultaneo di concetti . . .	47
3.4.2.1	Rete con 3 animali: implementazione e risultati	48
3.4.2.2	Rete con 3 animali 10 proprietà	57
3.4.2.3	Rete con 3 oggetti 10 proprietà	66

CAPITOLO 4

DIPENDENZA DAL CONTESTO	75
4.1	Conoscenza concettuale: proprietà CD e proprietà CI.	75
4.2	Modello con contesto	76
4.2.1	Descrizione quantitativa e qualitativa	76
4.2.2	Addestramento	79
4.3	Risultati	80
4.3.1	Risultati addestramento	80
4.3.2	Risultati simulazioni	81
4.4	Modello contesto e oscillatori	83
4.4.1	Descrizione e addestramento	83
4.4.2	Risultati	85
CONCLUSIONI	89
APPENDICE	93
BIBLIOGRAFIA	99

INTRODUZIONE

L'organizzazione concettuale e i meccanismi neurali che avvengono nella memoria semantica sono un tema centrale per le neuroscienze cognitive. La maniera in cui i concetti vengono organizzati e archiviati nella memoria a lungo termine è cruciale per diversi processi cognitivi ed influenza direttamente il comportamento dell'individuo, governando il modo in cui il mondo esterno determina le conoscenze individuali e la maniera in cui queste ultime influenzano le azioni. Per queste ragioni, l'analisi della memoria semantica attraverso modelli teorici (qualitativi o matematici) è un argomento di studio di grande interesse delle moderne neuroscienze cognitive. Il modello neurocomputazionale utilizzato nel presente lavoro è una rappresentazione della memoria semantica attraverso una rete auto-associativa, che sfrutta la regola di Hebb con una diversa soglia pre-sinaptica e post-sinaptica, producendo *pattern* di sinapsi asimmetriche all'interno della rete stessa. In particolare, il modello qui adoperato è una versione semplificata monodimensionale di un modello bidimensionale sviluppato in un lavoro di ricerca dal Prof. Ursino et al. nel 2013. Il principio alla base è che ogni concetto sia descritto da diversi attributi caratterizzati da un diverso grado di salienza. La memoria semantica è descritta da una rete, in cui ogni attributo viene codificato da un neurone. Durante l'addestramento sinaptico ogni concetto riceve in ingresso le sue proprietà (una per volta) con una diversa probabilità di occorrenza. Proprietà salienti hanno un'alta probabilità di verificarsi, di conseguenza dopo l'addestramento vengono subito richiamate alla mente quando si pensa ad un dato concetto e giocano un ruolo centrale nella rappresentazione concettuale. Diversamente, proprietà marginali hanno una bassa probabilità di verificarsi, di conseguenza non riceveranno forti sinapsi e, dopo l'addestramento, non vengono richiamate alla mente quando si pensa ad un dato concetto, ma aiutano la ricostruzione del suddetto se date come input.

Un ulteriore aspetto riguardante le proprietà è la discriminazione tra proprietà distintive e condivise: le prime appartengono ad un solo concetto, le seconde appartengono a più concetti e permettono la formazione di categorie. La descrizione dei singoli concetti adoperati nel seguente lavoro è stata ricavata da un *database* fornito dalla Dott.ssa Eleonora Catricalà (IUSS Pavia). In particolare sono state costruite due differenti tassonomie contenenti più concetti rispetto a quelle usate in lavori precedenti: una composta da 16 animali e una composta da 12 oggetti. Ogni concetto è stato descritto da un numero diverso di proprietà, mantenendo, però, un certo equilibrio tra proprietà salienti e marginali.

Lo scopo del presente lavoro, in primo luogo, è stato quello di testare il modello su tassonomie contenenti un numero di concetti elevato, con l'obiettivo di analizzare il ruolo delle diverse proprietà nella ricostruzione di un concetto e nella formazione di categorie. A seguire, sfruttando un modello di oscillatori neuronali, è stata studiata la possibilità del riconoscimento simultaneo di più concetti, ipotizzando la separazione di fase (cioè le proprietà di uno stesso oggetto oscillano in sincronismo, e le proprietà di oggetti diversi oscillano in modo asincrono). A tale fine, il modello è stato testato su reti più semplici, contenenti meno concetti rispetto a quelle di base. In queste reti ad ogni neurone è stato sostituito un oscillatore di Wilson – Cowan ed è stato aggiunto un inibitore globale con lo scopo di ritardare l'attivazione complessiva delle proprietà di un concetto rispetto a quella delle proprietà degli altri concetti presentati al modello. Successivamente è stato introdotto il concetto di contesto nelle reti di base; ciascuna di esse è stata quindi caratterizzata da proprietà *context dependent* (CD) e *context independent* (CI). Le successive simulazioni, con le reti addestrate, dimostrano la dipendenza contestuale delle caratteristiche dipendenti dal contesto. Dopo questi risultati, il concetto di contesto è stato inserito anche in una delle reti con oscillatori.

Questo è stato fatto per verificare che quando più concetti vengono presentati simultaneamente alla rete, e ciascuno di essi è caratterizzato anche da proprietà *context dependent*, queste ultime si richiamano a vicenda, oscillano in fase e risultano sfasate rispetto a quelle degli altri concetti; quindi i diversi contesti non si influenzano a vicenda.

La tesi è articolata nel seguente modo. Dopo un breve quadro generale sul concetto di memoria semantica e sull'organizzazione delle conoscenze concettuali al suo interno, sono state descritte in dettaglio le due tassonomie dei concetti utilizzate nello studio (*cap.1*). A seguire è stato descritto e testato il modello unidimensionale mostrando alcune prove di simulazioni su alcuni casi esemplari presi a partire dalle tassonomie di base (*cap. 2*). Successivamente, dopo una panoramica sul problema del "*binding and segmentation*" e sul sincronismo neuronale, è stato verificato il sincronismo tra 10 oscillatori puri di Wilson – Cowan. Il modello con oscillatori è stato poi testato su diverse reti semantiche costruite a partire dalle reti di base, ma contenenti un minore numero di concetti e di proprietà (*cap. 3*). Infine, è stato introdotto il concetto di dipendenza contestuale. Il modello di base è stato caratterizzato da proprietà dipendenti dal contesto (CD) e indipendenti dal contesto (CI); sono state descritte le modifiche apportate al modello basate sulla probabilità condizionata e il modello con contesto è stato poi testato su reti costruite a partire dalle tassonomie di partenza. Per concludere, il concetto di dipendenza contestuale è stato analizzato su una delle reti contenente gli oscillatori (*cap. 4*).

CAPITOLO 1

LA MEMORIA SEMANTICA

1.1 La memoria semantica

La memoria è un archivio all'interno del quale l'individuo conserva tracce della propria esperienza passata e al quale attinge per riuscire ad affrontare situazioni di vita presente e futura. Un modo per classificare la memoria, nell'ambito delle neuroscienze cognitive, si basa sulla durata dell'intervallo di ritenzione, ovvero quell'intervallo di tempo durante il quale l'informazione viene mantenuta in memoria prima di essere recuperata ed usata. Basandosi su quanto detto, la memoria viene suddivisa in memoria a breve termine (MBT) e memoria a lungo termine (MLT). La memoria a breve termine mantiene l'informazione per brevi intervalli di tempo (qualche secondo), mentre la memoria a lungo termine riesce a conservare l'informazione per intervalli di tempo più lunghi (ore, giorni o anni). Questa ultima è ulteriormente suddivisa in memoria dichiarativa e memoria procedurale. La memoria dichiarativa gestisce le informazioni che possono essere evocate consapevolmente, mentre la memoria procedurale riguarda informazioni relative a comportamenti automatici. La memoria semantica è quella parte di memoria dichiarativa organizzata in modo tassonomico e associativo, costituita da un insieme di concetti, conoscenze e nozioni acquisiti grazie ad esperienze di vita (ad esempio, "il gatto è un animale"), indipendenti dal contesto dell'originale episodio d'apprendimento e in larga parte condivisa con altri individui.

1.2 Organizzazione delle conoscenze

Le conoscenze presenti nella memoria semantica sono organizzate su base categoriale: ogni concetto è legato semanticamente ad un determinato numero di concetti e non ad altri in modo da formare delle categorie semantiche. Il processo di categorizzazione delle informazioni è importante perché permette la creazione di classi più o meno ampie di elementi, che possono essere recuperati anche se non sono esplicitati verbalmente.

Inoltre il processo di categorizzazione rende possibili risposte comportamentali riferite a una classe di concetti cognitivamente equiparabili e permette di rilevare analogie e differenze fra concetti a diversi livelli di astrazione [12]. La visione classica delle categorie semantiche prevede la scomposizione di un concetto in attributi che riassumono la somma delle esperienze sensoriali e motorie con l'ambiente, o esperienze di tipo enciclopedico, e che includono più tipi di informazioni. Per esempio, il concetto di “cane” comprende le caratteristiche: “ha quattro zampe”, “ha il pelo”, “abbaia”. In generale tali proprietà sono considerate “*defining features*” perché rappresentano la definizione di un concetto. Inoltre si sostiene che non tutte le caratteristiche semantiche abbiano un'importanza simile per la rappresentazione concettuale e che differenti tipi di caratteristiche svolgano ruoli specifici nell'organizzazione della memoria semantica [3]. La maniera in cui le conoscenze sono organizzate ed archiviate nella memoria e i meccanismi neurali che vi stanno alla base sono cruciali per molti processi della cognizione umana, compresi il riconoscimento di oggetti, la pianificazione di azioni, il linguaggio e la formazione di idee astratte. L'organizzazione delle informazioni concettuali non ha un'importanza puramente teorica, essa infatti influenza direttamente il comportamento dell'individuo dal momento che governa la modalità attraverso la quale gli stimoli esterni determinano i concetti e la modalità con cui questi ultimi influenzano, a loro volta, le azioni. Per questo motivo, lo studio della memoria semantica attraverso modelli teorici svolge un ruolo importante nell'ambito delle neuroscienze cognitive [1,2].

1.3 Modelli di memoria semantica

Per quanto riguarda il formato della rappresentazione, si possono identificare tre classi di modelli: rappresentazione astratta, rappresentazione per esemplari, e approccio connessionista. Secondo i modelli di rappresentazione astratta, le informazioni vengono conservate in memoria in un formato slegato dalle informazioni sensoriali-motorie delle entità rappresentate.

Secondo i modelli per esemplari il sistema mentale è costituito dalle memorie degli esemplari che sono stati codificati nel tempo. In altre parole, la rappresentazione del concetto «gatto» è costituita dai prototipi che la mente forma per tale concetto. Infine, i modelli connessionisti della rappresentazione delle conoscenze implicano, nella maggior parte dei casi, una architettura distribuita in cui la rappresentazione di un concetto viene «spalmata» su diversi sottosistemi. Pertanto, secondo questo approccio, non esiste un nodo concettuale corrispondente a «gatto» e neppure l'insieme di esemplari di «gatto». Esistono invece insiemi di attributi di base, condivisi da un numero variabile di concetti, che si attivano in configurazioni appropriate in riferimento al concetto significativo. Semplificando in modo notevole, si può affermare che i concetti di «cane» e di «gatto» saranno utilizzabili quando verrà attivata una configurazione che comprenderà, fra le altre, le caratteristiche “abbaia”, “da guardia” e “fedele” per il primo, “caccia i topi”, “miagola” e “fa le fusa” per il secondo. Nei modelli connessionisti le categorie «emergono» dalla somiglianza dei *pattern* di attivazione di insiemi di attributi relativi ai diversi concetti. Ad esempio, l'attivazione delle caratteristiche “ha quattro zampe”, “ha il pelo” e “ha la coda” attiverà gli esemplari della categoria dei mammiferi mentre l'attivazione delle caratteristiche “ha due zampe”, “ha le piume” e “ha le ali” attiverà gli esemplari della categoria degli uccelli. Le tre classi di modelli descritte riflettono differenti concezioni della rappresentazione delle conoscenze nella mente umana. Non è possibile stabilire, sulla base delle conoscenze attuali, quale di questi modelli renda meglio conto del funzionamento della memoria semantica; tuttavia diversi modelli matematici, chiamati “reti di attrattori”, sfruttano l'approccio connessionista per indagare i processi della mente umana. In questi modelli matematici i processi cognitivi sono implementati in termini di complesse interazioni non-lineari tra un ampio numero di semplici unità di elaborazione.

In particolare in queste reti ogni concetto viene rappresentato come un insieme distribuito di nodi attivi (che simula la codifica di gruppi neurali per le caratteristiche) mentre le regolarità statistiche del verificarsi delle caratteristiche sono codificate nelle connessioni tra i nodi (cioè nelle sinapsi). Queste ultime vengono assegnate sulla base di alcune regole di apprendimento sinaptico attraverso il quale la rappresentazione del concetto viene memorizzata come punto di equilibrio della rete. In seguito, per ripristinare un concetto, si sfruttano le dinamiche attrattive del modello che, partendo da un punto circoscritto (cioè dall'attivazione di pochi nodi), converge verso un punto di equilibrio. Le reti di attrattori basate su caratteristiche forniscono intuizioni significative su diversi fenomeni comportamentali riguardanti la memoria semantica, sia in condizioni normali sia in condizioni patologiche. In particolare, le dinamiche attrattive spiegano chiaramente il diverso ruolo svolto da caratteristiche distintive e condivise nella rappresentazione semantica dei concetti, e la distinzione tra concetti subordinati e sovraordinati [2].

1.4 Database e tassonomie semantiche

Le tassonomie semantiche adoperate nel presente lavoro di tesi sono state costruite a partire da un *database* sviluppato da Catricalà et al. (2013, 2015a). L'intero *database* sul quale ci si è basati contiene 82 concetti, suddivisi in viventi ("uccelli", "animali terrestri", "ortaggi", "frutti") e non viventi ("mobili", "utensili", "articoli da cucina", "capi di abbigliamento", "veicoli"). Nel *database* sono contenuti diversi descrittori rilevanti per l'organizzazione della memoria semantica ricavati dalle *feature* generate da 20 partecipanti. Tali descrittori sono la distintività, la dominanza, la rilevanza semantica e la distanza semantica. La distintività serve per distinguere un concetto da altri. In particolare sono presenti due misure di distintività: la prima si riferisce al numero di concetti in cui appare la caratteristica semantica in esame, diviso il numero totale di concetti del *database*; la seconda si riferisce alla percentuale di concetti all'interno di una categoria per la quale la caratteristica in questione è stata elencata.

La dominanza valuta la frequenza con cui è prodotta una certa proprietà; indica cioè la percentuale di partecipanti che ha elencato una caratteristica specifica per un determinato concetto. La rilevanza semantica quantifica l'importanza di una caratteristica e viene calcolata come combinazione non lineare tra dominanza e distintività. Caratteristiche con alti valori di rilevanza semantica consentono l'identificazione del concetto e la sua discriminazione da altri concetti simili. Infine la distanza semantica è una misura della somiglianza tra rappresentazioni semantiche. Concetti vicini semanticamente risultano essere più difficili da identificare poiché condividono un numero sostanziale di caratteristiche e possono essere tra loro confusi (es. "divano" e "poltrona"). In particolare sono presenti due diverse misure di distanza semantica: una tra ogni coppia di concetti appartenente alla stessa categoria, e una tra ogni concetto e il centroide della relativa categoria semantica [2]. L'obiettivo del presente lavoro è stato quello di realizzare un modello neuro-computazionale ad attrattore che preveda l'uso di tassonomie semantiche più ampie rispetto a quelle adoperate in studi precedenti. In particolare sono stati scelti 16 animali e 12 oggetti dal *database* descritto in precedenza. Tutti i concetti adoperati contengono i tipi di caratteristiche di interesse, ovvero caratteristiche condivise (totalmente o parzialmente) e caratteristiche distintive (salienti o marginali) descritte in modo dettagliato di seguito:

- una caratteristica è condivisa se appartiene a diversi concetti. A sua volta, con riferimento a una specifica categoria, una caratteristica condivisa può esserlo totalmente o parzialmente. Nel primo caso la caratteristica appartiene a tutti i membri di una categoria; nel secondo caso invece appartiene solo ad alcuni membri di una categoria (più di un singolo membro).
- una caratteristica è distintiva se appartiene a un solo concetto, consentendo la differenziazione tra concetti strettamente correlati, in genere membri della stessa categoria semantica.

Non tutte le caratteristiche distintive hanno una salienza identica per lo stesso concetto, di fatto una caratteristica distintiva può essere saliente o marginale. Nel primo caso è evocata spontaneamente quando si pensa all'oggetto/animale, poiché frequentemente usata per definire il concetto (es. "tuba" per il concetto di "piccione"); al contrario una caratteristica risulta marginale se non è frequentemente evocata quando si pensa al concetto (es. "usato per mandare messaggi" per il concetto di "piccione"). Le *feature* condivise contribuiscono alla formazione di categorie (es. "ha il pelo" per il concetto "mammifero"), le caratteristiche distintive consentono la differenziazione tra concetti correlati appartenenti alla stessa categoria semantica (es. "miagola" per il concetto di "gatto"). A questo punto si hanno due dimensioni principali per quantificare il ruolo delle caratteristiche in una struttura semantica: una dimensione globale, distintività contro condivisione, e una dimensione locale, dominanza contro marginalità. A differenza di elaborati precedenti, l'assegnazione della salienza alle varie proprietà è stata eseguita nel modo seguente: le caratteristiche totalmente condivise, ovvero quelle che formano categorie come ANIMALE, OGGETTO, MAMMIFERO, ARREDO, etc.. sono state ritenute salienti in quanto spontaneamente evocate quando si pensa al concetto in questione (es. "mangia" o "si riproduce" per la categoria ANIMALE). Per le caratteristiche dei singoli concetti (distintive e parzialmente condivise) invece, dopo averle ordinate in base al valore di dominanza crescente, si è scelto di considerare salienti un numero di proprietà pari alla metà del numero totale di caratteristiche di ogni concetto. Le restanti proprietà sono state etichettate come marginali. Questa distinzione è stata fatto in modo da mantenere un certo equilibrio tra proprietà salienti e marginali.

1.4.1 Tassonomia Oggetti

Per formare la tassonomia di oggetti utilizzata nel seguente elaborato si sono scelti dal *database* originale 12 oggetti in base alla vicinanza semantica ovvero la misura di somiglianza tra rappresentazioni semantiche.

È opportuno notare che, dal momento che gli esseri viventi sono semanticamente più vicini tra loro rispetto a quelli non viventi, i concetti presenti nella tassonomia degli oggetti risultano avere una vicinanza semantica minore rispetto ai concetti presenti nella tassonomia degli animali. Questo comporta la possibilità di formare un numero minore di categorie nella rete degli oggetti. In particolare in questa tassonomia sono stati scelti 3 oggetti per la categoria ARREDO, 3 oggetti per la categoria UTENSILI DA CUCINA e 6 oggetti che fanno parte della generica categoria OGGETTO. Ogni concetto ha un numero di proprietà compreso tra 15 e 22 (non è stato preso lo stesso numero di proprietà per ciascun oggetto dato che in una vera memoria semantica alcuni concetti possono avere più proprietà rispetto ad altri), per un totale di 165 proprietà. Alcune di queste caratteristiche sono totalmente condivise e permettono la creazione di categorie, altre sono parzialmente condivise ed altre ancora sono invece proprietà distintive che permettono di differenziare un oggetto rispetto agli altri (tabella 1, tabella 2, tabella 3 nella pagina seguente). In particolare tutti gli oggetti hanno in comune la proprietà “oggetto” che forma la macro categoria OGGETTO. Da questa derivano due sottocategorie: ARREDO ed UTENSILI DA CUCINA. Gli oggetti che appartengono alla prima sottocategoria hanno in comune (oltre ad “oggetto”) la proprietà “mobile”, mentre quelli che appartengono alla seconda sottocategoria hanno in comune (oltre ad “oggetto”) la proprietà “si trova in cucina”. Infine ogni oggetto ha proprietà distintive (salienti e marginali) e parzialmente condivise (salienti e marginali). L'intera tassonomia degli oggetti è riportata nello schema in *fig. A1* in appendice.

Posizione	Oggetto	Proprietà totalmente condivise
1	Tavolo	2,3
2	Divano	2,3
3	Libreria	2,3
4	Pentola	2,4
5	Forchetta	2,4
6	Tazzina	2,4
7	Bottiglia	2
8	Lampada	2
9	Sveglia	2
10	Pistola	2
11	Ombrello	2
12	Martello	2

Tabella 1. Nella prima colonna sono riportate le posizioni che gli oggetti assumono nella rete in modo da avere le proprietà raggruppate in modo contiguo, nella seconda colonna i nomi di tutti gli oggetti e nella terza colonna le posizioni che le proprietà totalmente condivise associate ad ogni oggetto assumono all'interno della rete.

Posizione	Oggetto	Proprietà parzialmente condivise
1	Tavolo	139,140,141,142,146,147,152,161
2	Divano	142,145,148
3	Libreria	139,143,147,149,163
4	Pentola	140,143,150,154,158,159,163,165
5	Forchetta	147,149,150,155,159,161,165
6	Tazzina	139,141,144,146,152,154,163,164,165
7	Bottiglia	146,151,153,155,164
8	Lampada	139,144,145,151,158
9	Sveglia	139,144,146,149,158,162
10	Pistola	149,156,157,160,165
11	Ombrello	145,146,148,159,162,165,166
12	Martello	156,157,159,160,162,166

Tabella 2. Nella prima colonna sono riportate le posizioni che i oggetti assumono nella rete in modo da avere le proprietà raggruppate in modo contiguo, nella seconda colonna i nomi di tutti gli oggetti e nella terza colonna le posizioni che le proprietà parzialmente condivise associate ad ogni oggetto assumono all'interno della rete.

Posizione	Oggetto	Proprietà 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 3. Nella prima colonna sono riportate le posizioni che i oggetti assumono nella rete in modo da avere le proprietà raggruppate in modo contiguo, nella seconda colonna i nomi di tutti gli oggetti e nella terza colonna le posizioni che le proprietà distintive associate ad ogni oggetto assumono all'interno della rete.

1.4.2 Tassonomia Animali

Per formare la tassonomia di animali utilizzata nel seguente elaborato, si sono scelti dal *database* originale 16 animali in base alla vicinanza semantica. A differenza della tassonomia degli oggetti, i concetti presenti nella tassonomia degli animali risultano essere semanticamente più vicini. Questo comporta la possibilità di formare un numero maggiore di categorie rispetto alla tassonomia degli oggetti descritta nel sotto paragrafo precedente. In particolare sono stati scelti 10 animali per la categoria MAMMIFERO e 6 animali per la categoria UCCELLO. Ogni animale ha un numero di proprietà compreso tra 25 e 34 (non è stato preso lo stesso numero di proprietà per ciascun oggetto dato che in una vera memoria semantica alcuni concetti possono avere più proprietà rispetto ad altri), per un totale di 231 proprietà. Alcune di queste caratteristiche sono totalmente condivise e permettono la creazione di categorie, altre sono parzialmente condivise ed altre ancora sono invece proprietà distintive che consentono di differenziare un animale rispetto agli altri (tabella 4, tabella 5, tabella 6). In particolare tutti gli animali hanno in comune le proprietà “animale”, “mangia”, “ha le zampe”, “si riproduce” che formano la macro categoria ANIMALE. Da questa derivano due sottocategorie: MAMMIFERO ed UCCELLO. Gli animali che appartengono alla prima sottocategoria hanno in comune (oltre alle proprietà della categoria ANIMALE) la proprietà “è un mammifero”, “ha quattro zampe”, “ha il pelo”, “ha la coda”. Quelli che appartengono alla seconda sottocategoria hanno in comune (oltre alle proprietà della categoria ANIMALE) le proprietà “è un uccello”, “ha due zampe”, “ha le piume” ed “ha le ali”. Dalle due sottocategorie principali (MAMMIFERO e UCCELLO) si formano ulteriori sottocategorie. In particolare la categoria UCCELLO si divide in VOLATILI e NON VOLATILI. Gli animali appartenenti ai VOLATILI hanno in comune (oltre alle caratteristiche della categoria ANIMALE e UCCELLO) le proprietà “vola” e “fa il nido”.

Gli animali appartenenti alla categoria NON VOLATILI hanno in comune (oltre alle caratteristiche della categoria ANIMALE e UCCELLO) le proprietà “svolazza” e “ha la cresta”. La categoria MAMMIFERO si suddivide a sua volta in DOMESTICO e SELVATICO. Gli animali che fanno parte della categoria DOMESTICO hanno in comune (oltre alle caratteristiche delle categorie ANIMALE e MAMMIFERO) le proprietà “domestico” e “da compagnia”. La categoria SELVATICO contiene la proprietà “è selvatico” che risulta comune a tutti i concetti subordinati e si suddivide in CARNIVORO e ERBIVORO. Gli animali che appartengono alla categoria CARNIVORO hanno in comune gli attributi “ha gli artigli” e “mangia carne” (oltre alle caratteristiche della categoria ANIMALE, MAMMIFERO e SELVATICO). La categoria ERBIVORO ha la proprietà “mangia l’erba” che risulta comune a tutti i concetti subordinati e si suddivide in AFRICANI e DA ALLEVAMENTO. Gli animali appartenenti alla prima classe condividono “vive in Africa” e “vive nella savana” mentre quelli della seconda classe “si ricava la carne” e “vive nella fattoria” (oltre alle proprietà delle categoria gerarchicamente superiori). Infine ogni concetto ha proprietà distintive (salienti e marginali) e parzialmente condivise (salienti e marginali). L’intera tassonomia degli animali è riportata nello schema in *fig. A2* appendice.

Posizione	Animale	Proprietà totalmente condivise
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 4. Nella prima colonna sono riportate le posizioni che gli animali assumono nella rete in modo da avere le proprietà raggruppate in modo contiguo, nella seconda colonna sono riportati i nomi di tutti gli animali e nella terza colonna sono riportate le posizioni che le proprietà parzialmente totalmente condivise associate ad ogni animale assumono nella rete.

Posizione	Animale	Proprietà 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	189,190,193,195,199,200,202,208,209,212,229,231
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	201,214,218,219,223,229,232
7	Elefante	192,193,195,199,208,212,212,229,230
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 5. Nella prima colonna sono riportate le posizioni che gli animali assumono nella rete in modo da avere le proprietà raggruppate in modo contiguo; nella seconda colonna sono riportati i nomi di tutti gli animali; nella terza colonna sono riportate le posizioni che le proprietà parzialmente condivise associate ad ogni animale assumono nella rete.

Posizione	Animale	Proprietà 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	154,155,156,157,158,169,160,161,162,163,164,165
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 6. Nella prima colonna sono riportate le posizioni che gli animali assumono nella rete in modo da avere le proprietà raggruppate in modo contiguo; nella seconda colonna sono riportati i nomi di tutti gli animali; nella terza colonna sono riportate le posizioni che le proprietà distintive associate ad ogni animale assumono nella rete.

CAPITOLO 2

MODELLO DI MEMORIA SEMANTICA

2.1 Modello unidimensionale

Il modello utilizzato in questo lavoro di tesi si basa su una versione semplificata monodimensionale della versione più recente del modello bidimensionale sviluppato nel lavoro di ricerca dal Prof. Ursino nel 2013. Il modello incorpora una rete di neuroni che implementa la rete semantica dedicata alla descrizione dei concetti rappresentati come una raccolta di caratteristiche (*fig. 1*).

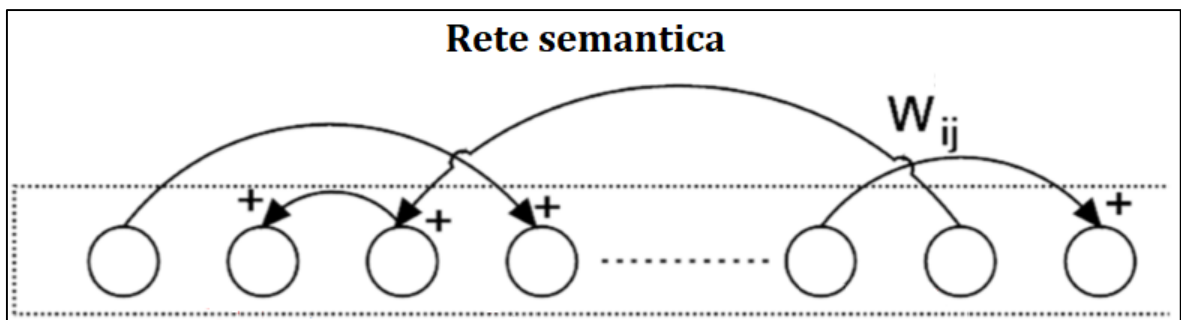


Figura 1 - Diagramma schematico del modello. Si noti che vengono riportati solo alcune sinapsi a titolo esplicativo per evitare di appesantire la notazione grafica.

Tali caratteristiche sono raccolte in un vettore di lunghezza M pari al numero di *feature* utilizzate. Ogni unità neurale codifica per una caratteristica ed ognuna di esse può ricevere sinapsi da tutte le altre (tranne che da sé stessa) formando una matrice delle sinapsi di dimensione $M \times M$. Ogni unità neurale viene indicata con un solo pedice "i", mentre una sinapsi tra due unità neurali ha due pedici "ij", il primo rappresenta la posizione del neurone post-sinaptico e il secondo la posizione del neurone pre-sinaptico. È bene ricordare che nel suddetto lavoro sono presenti solo sinapsi eccitatorie non simmetriche W_{ij} tra i neuroni della rete semantica. Tali sinapsi vengono create sulla base dell'esperienza passata della rappresentazione del concetto, con un paradigma Hebbiano. Il modello è stato addestrato usando due tassonomie di concetti ampliate rispetto alle precedenti (ANIMALI e OGGETTI).

I concetti vengono descritti in modo schematico attraverso un vettore di *feature*. Alcune proprietà sono condivise e altre distintive; inoltre esse presentano una diversa frequenza di occorrenza durante l'addestramento e questo si riflette in una differente salienza. La rete addestrata è in grado di risolvere compiti di riconoscimento di concetti, mantenendo una distinzione tra le categorie e i membri individuali all'interno della categoria, e dando un diverso ruolo alle proprietà salienti rispetto a quelle marginali [1,12].

2.2 Equazioni del modello

Ogni unità nella rete semantica riceve un input globale (indicato da u) e produce un output (indicato da x). Questa relazione è descritta attraverso un filtro passa-basso del primo ordine con costante di tempo τ (Eq.(1)) e una funzione sigmoideale (Eq.(2)) [2]. La prima espressione riproduce la risposta nel tempo dei neuroni, mentre la seconda rappresenta una soglia inferiore e una saturazione superiore dell'attività neuronale. Quindi:

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

dove τ è la costante di tempo che determina la velocità della risposta allo stimolo, $H(u_i(t))$ è una funzione di attivazione sigmoideale e il pedice “ i ” indica la posizione dell'unità nella rete. La relazione sigmoideale è descritta dalla seguente equazione:

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

dove T è un parametro che imposta la pendenza della sigmoide e φ stabilisce la sua posizione. Quindi, quando l'input è troppo basso, l'attività del neurone è trascurabile, mentre un input elevato conduce il neurone ad una attività vicina alla saturazione, che come si deduce dalla Eq. (2) è a 1. L'input per i neuroni viene calcolato come la somma di due contributi:

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

dove I_i rappresenta l'input esterno per l'unità neurale in posizione i , proveniente dalla catena di elaborazione sensoriale-motoria ed emotiva che estrae le caratteristiche. I_i può assumere il valore 0 (assenza di *feature*) o 1 (presenza di *feature*). E_i rappresenta un termine di accoppiamento eccitatorio proveniente da altre unità della rete semantica (cioè, dai neuroni che codificano per una caratteristica diversa). Il termine di accoppiamento all'interno della rete semantica ha la seguente espressione:

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

dove i indica la posizione del neurone post-sinaptico (bersaglio), mentre j indica la posizione del neurone pre-sinaptico; inoltre le somme si estendono a tutti i neuroni pre-sinaptici nella rete semantica. Il simbolo W_{ij} rappresenta le sinapsi eccitatorie, che formano una rete auto-associativa all'interno della semantica. Queste sinapsi sono soggette all'apprendimento descritto di seguito. Si trascurano inoltre le sinapsi inibitorie all'interno della rete semantica in quanto si assume che caratteristiche che non si verificano insieme semplicemente non siano collegate. Inoltre si trascurano anche le sinapsi auto-eccitatorie di ciascuna proprietà (la proprietà "ha le zampe" non richiama sé stessa).

2.3 Addestramento sinaptico

2.3.1 Regola di Hebb

Le sinapsi eccitatorie di entrambe le reti (ANIMALI e OGGETTI) vengono addestrate mediante la regola di Hebb, che modifica il peso sinaptico sulla base della correlazione tra l'attività pre e post sinaptica. Per tenere conto non solo del potenziamento a lungo termine ma anche del depotenziamento a lungo termine e rendere quindi la regola di addestramento biologicamente più plausibile, queste attività vengono confrontate con una soglia. In tal modo, 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.

Viceversa, quando entrambi i neuroni hanno un'attività elevata, la sinapsi si rafforza. Quindi, si può scrivere:

$$\Delta W_{ij} = \gamma_{ij}(x_i - \vartheta_{post})(x_j - \vartheta_{pre}) \quad (5)$$

dove ΔW_{ij} rappresenta il cambiamento nella forza della sinapsi, a causa delle attività pre e post-sinaptiche, x_i è l'attività dell'unità neurale alla posizione i nell'area post-sinaptica, e x_j è l'attività dell'unità neurale nella posizione j nell'area pre-sinaptica. ϑ_{post} e ϑ_{pre} sono le soglie per le attività post e pre-sinaptiche, mentre γ_{ij} denota un coefficiente di addestramento che stabilisce la velocità di apprendimento (dipendente dalla storia precedente della sinapsi). Affinché l'Eq. (5) sia fisiologicamente realistica occorrono alcune precisazioni. Innanzitutto, quando entrambe le attività pre-sinaptiche e post-sinaptiche sono basse, non dovrebbero verificarsi cambiamenti di peso, cioè la sinapsi dovrebbe rimanere invariata. Quindi:

$$se \ x_i < \vartheta_{post} \quad e \quad x_j < \vartheta_{pre} \quad \rightarrow \quad \Delta W_{ij} = 0$$

Secondo, le sinapsi non possono aumentare indefinitamente, ma devono raggiungere un livello massimo di saturazione. Ciò è ottenuto nel modello riducendo progressivamente il coefficiente di addestramento quando le sinapsi si avvicinano al loro massimo (indicato da W_{max}). Quindi:

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

dove γ_0 è il coefficiente di addestramento quando la sinapsi è zero. Infine, una sinapsi non può diventare negativa (altrimenti l'eccitazione verrebbe convertita in inibizione, che non è fisiologicamente accettabile). Quindi, quando calcoliamo il nuovo valore della sinapsi si ha:

$$W_{ij} \leftarrow \text{Max}\{W_{ij} + \Delta W_{ij}, 0\}$$

dove il simbolo \leftarrow significa che il valore calcolato sul membro di destra è assegnato al membro di sinistra [2].

2.3.2 Fase di addestramento

L'addestramento parte da una condizione iniziale, in cui tutte le sinapsi eccitatorie nella rete vengono impostate a zero. La procedura di apprendimento consiste in 1000 epoche consecutive; durante ciascuna epoca tutti i concetti vengono presentati separatamente e permutati in ordine casuale (cioè ogni concetto è presentato una sola volta in ogni epoca) e la regola di Hebb viene applicata in condizioni stazionarie. In questo modo vengono apprese le sinapsi che collegano le diverse caratteristiche. Questo corrisponde a una fase in cui il soggetto sperimenta gli oggetti e apprende la loro semantica. Tuttavia, non tutte le caratteristiche del concetto vengono presentate in input simultaneamente. Ogni caratteristica ha una probabilità che viene adoperata come input per un dato concetto: maggiore è questa probabilità, maggiore è la salienza della caratteristica. Consideriamo la caratteristica j -esima nella rete e denotiamo con P_{kj} la probabilità che tale caratteristica venga presentata come input per il concetto k -esimo. Di conseguenza, durante l'allenamento, quando viene presentato il concetto k -esimo, la caratteristica j -esima ha una probabilità di ricevere un input elevato (ovvero un valore sufficiente per portare il neurone vicino alla saturazione, ciò significa che la caratteristica è percepita) pari a $P_{kj(s)}$, se è saliente, oppure pari a $P_{kj(m)}$, se è marginale. La probabilità di ricevere un input basso (ovvero un input che non è in grado di eccitare il neurone corrispondente, ciò significa che la caratteristica non è percepita) sarà invece pari a $1 - P_{kj}$. Per semplicità e per consentire un'analisi più semplice dei risultati, nel presente lavoro a tutte le caratteristiche salienti è stata data un'alta probabilità, vale a dire il 70% (cioè $P_{kj(s)} = 0.7$) mentre a tutte le caratteristiche marginali è stata data una bassa probabilità, vale a dire il 35% (cioè $P_{kj(m)} = 0.35$) a differenza dei lavori precedenti in cui questa ultima probabilità era stata impostata pari al 40%. Questa modifica è stata introdotta in modo da discriminare meglio fra proprietà salienti e marginali.

2.4 Assegnazione parametri

Un punto fondamentale per ottenere un corretto addestramento sinaptico riguarda la scelta di valori appropriati per le soglie pre e post-sinaptiche nella regola di Hebb (Eq. (5)). Tale scelta deve tenere conto dei requisiti che il modello deve rispettare, ovvero:

Saliente vs. marginale. Le caratteristiche salienti devono essere richiamate da tutte le altre caratteristiche del concetto (salienti e non salienti), quindi dovrebbero ricevere forti sinapsi di input e dovrebbero inviare forti sinapsi di output solo alle altre caratteristiche salienti. Le caratteristiche marginali invece non dovrebbero essere richiamate dalle altre caratteristiche del concetto, quindi dovrebbero ricevere deboli sinapsi di input. Al contrario, dovrebbero inviare forti sinapsi di output verso le *feature* salienti per favorire la ricostruzione del concetto.

Condivisa vs distintiva. Le caratteristiche totalmente condivise da concetti diversi (cioè le *feature* che formano una categoria) dovrebbero richiamare tutte le altre caratteristiche totalmente condivise nella categoria, ma non dovrebbero richiamare le caratteristiche distintive dei singoli membri (ad es. la caratteristica “mangia” dovrebbe richiamare le caratteristiche comuni della categoria ANIMALE come "si riproduce" ma non dovrebbe richiamare le caratteristiche dei singoli membri “orso”, “leone”, ecc.. , come "è polare", "ruggisce" ecc.).

Una caratteristica distintiva invece dovrebbe richiamare non solo le altre caratteristiche distintive salienti dello stesso concetto ma anche le caratteristiche totalmente condivise ("ruggisce" dovrebbe richiamare tutte le proprietà salienti di “leone”, incluso "mangia", “si riproduce” etc.. condivise con altri animali). Le condizioni descritte in precedenza vengono riassunte nello schema in *fig.2* nella pagina seguente.

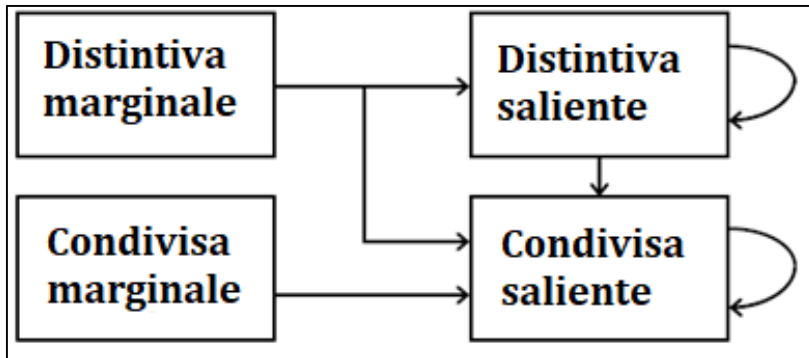


Figura 2 - Esempio qualitativo di connessioni sinaptiche tra caratteristiche condivise e distintive e tra caratteristiche salienti e marginali.

Questo particolare comportamento delle sinapsi semantiche può essere ottenuto assumendo che la soglia per l'attività post-sinaptica e la soglia per l'attività pre-sinaptica siano diverse; in questo modo si ottengono infatti sinapsi non simmetriche che garantiscono una buona distinzione tra proprietà condivise e proprietà distintive, e tra proprietà salienti e marginali. In particolare la soglia post-sinaptica di base è stata impostata ad un valore piuttosto alto ($\vartheta_{post-base} = 0.67$) mentre la soglia per l'attività pre-sinaptica è stata lasciata allo stesso valore dei modelli precedenti ($\vartheta_{pre} = 0.05$, vicino all'inibizione: qui è stato scelto un valore appena sopra lo zero per evitare che un'attività neuronale residua, indotta dalle connessioni nella rete semantica, possa causare un rinforzo indesiderato delle sinapsi). In generale il valore della soglia post sinaptica è strettamente correlato al concetto di salienza poiché stabilisce in maniera approssimativa la frequenza che discrimina tra salienza e marginalità. Avere un valore di soglia post sinaptica elevato aumenta il livello di frequenza richiesto per avere salienza. Nel nostro elaborato, dopo diverse prove, il valore di $\vartheta_{post-base}$ è stato settato a 0.67 poiché con tale valore si è ottenuto un comportamento ottimale del modello, comportamento non raggiungibile con valori più bassi. Il fatto di avere impostato una soglia pre sinaptica ad un valore basso e una soglia post sinaptica ad un valore più alto, comporta quanto spiegato in seguito [11]:

- se entrambi i neuroni pre-sinaptico e post-sinaptico sono attivi (attività prossima a 1), le sinapsi si rafforzano.
- se il neurone post-sinaptico è inibito (attività prossima a 0) mentre il neurone pre-sinaptico è eccitato (attività prossima a 1), la sinapsi si indebolisce. Questa situazione si verifica in quelle sinapsi che vanno da una caratteristica condivisa verso una caratteristica distintiva, quando la caratteristica condivisa appare in un concetto che non contiene quella particolare caratteristica distintiva (es. dalla caratteristica “ha il pelo” verso la caratteristica “miagola” quando si pensa al concetto di cane). La stessa situazione si ha anche per quelle sinapsi che vanno da una caratteristica saliente verso una marginale. Quindi, dopo un addestramento sufficiente, grazie all’occorrenza statistica delle proprietà, le caratteristiche condivise dovrebbero inviare sinapsi deboli verso quelle distintive e le caratteristiche salienti dovrebbero inviare sinapsi deboli verso quelle marginali.
- se il neurone post-sinaptico è eccitato (attività prossima a 1) e il neurone pre-sinaptico è inibito (attività prossima a 0), la sinapsi mostra solo un moderato indebolimento. Questa è una situazione che si verifica nelle sinapsi che vanno da una caratteristica marginale verso una caratteristica saliente (l'attività pre-sinaptica è spesso vicina allo zero, poiché le caratteristiche marginali sono spesso assenti). Di conseguenza, una caratteristica marginale distintiva continua a inviare forti sinapsi verso tutte le caratteristiche salienti, con un indebolimento solo limitato nei casi in cui non sia presente in input. La stessa condizione si verifica anche se consideriamo una sinapsi che va da qualsiasi caratteristica distintiva verso una caratteristica condivisa.

Come in elaborati precedenti, si è notato che il comportamento del modello con soglia post-sinaptica fissa risulta inadeguato per le caratteristiche salienti parzialmente condivise. Per esempio, si considerino le caratteristiche “vola” e “ha le ali” che, osservando la tassonomia, sono condivise da diversi uccelli. La caratteristica “ha le ali” è saliente per tutti e sei gli uccelli (si verifica nel 70% dei casi) e, dopo l’addestramento, viene evocato nella categoria uccello. La caratteristica “vola” invece, secondo la tassonomia, non si verifica per “gallina” e “gallo”, ma si verifica con un’elevata frequenza (70%) negli uccelli che volano (“oca”, “gufo”, “pappagallo” e “piccione”) e dopo l’addestramento risulta saliente per questi quattro uccelli. Considerando la connessione tra “vola” e “ha le ali”, dopo l’addestramento, la caratteristica “ha le ali” si verifica nel 100% degli uccelli, in quanto saliente. La caratteristica “vola” si verifica nel 66% degli uccelli (ovvero è saliente nei 4/6 di essi). Con la soglia post-sinaptica fissa pari a $\vartheta_{post} = \vartheta_{post-base} = 0.67$, è possibile che si crei una sinapsi che vada da “vola” verso “ha le ali”, ma anche una sinapsi che vada da “ha le ali” verso “vola” e “vola” diventi così saliente per tutti gli uccelli. Per tale motivo si assume, come nei lavori precedenti, che la soglia post-sinaptica dipenda dal numero di volte che la caratteristica pre-sinaptica si verifichi, questa diventa quindi funzione di j . In particolare, la soglia post-sinaptica aumenta per tutte le sinapsi che si originano da proprietà condivise da più membri di una categoria. La legge di Hebb effettivamente utilizzata per addestrare le sinapsi nel presente lavoro diventa quindi:

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

dove il pedice j nella soglia indica che essa dipende dalla caratteristica pre-sinaptica. In particolare vale:

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

dove N_j è il numero medio di occorrenze della caratteristica j (pre-sinaptica) durante un'epoca. $\vartheta_{post-base}$ rappresenta il valore di frequenza richiesta per avere salienza per caratteristiche distintive (cioè quando $N_j = 1$), mentre $\Delta\vartheta_{post}$ è la velocità di incremento della soglia post-sinaptica. $\vartheta_{post-base}$ nel presente lavoro è stato scelto pari a 0.67, mentre $\Delta\vartheta_{post}$ è stato lasciato pari al valore usato in elaborati precedenti (0.5). Abbiamo osservato infatti che 0.5 è un valore adatto per ottenere un apprendimento delle sinapsi corretto con le due tassonomie. Vale la pena notare che una caratteristica distintiva non può avere $N_j > 1$ e quindi la soglia post-sinaptica è fissata ad un livello basale. Viceversa, le caratteristiche condivise hanno $N_j > 1$, e quindi $\vartheta_{post,j} > \vartheta_{post-base}$. Ciò significa che le caratteristiche distintive possono facilmente creare sinapsi verso caratteristiche condivise, ma il contrario. Infine, si suppone che la soglia post sinaptica non possa mai superare un livello di saturazione massimo (chiamato $\vartheta_{post-sat}$):

$$\vartheta_{post,j} < \vartheta_{post-sat} \quad \forall j$$

Come in lavori precedenti il valore di $\vartheta_{post-sat}$ deve essere solo un po' più piccolo di 1 in modo da consentire che solo le caratteristiche condivise da quasi tutti i membri di una categoria (quindi con probabilità che diviene prossima a 1) siano attribuite alla categoria, pertanto il suo valore non è stato variato rispetto a lavori passati. Tutti i dettagli sui valori numerici dei parametri (soglie, saturazione superiore, pendenza della sigmoide, ecc.) sufficienti a ricostruire l'intero modello di lavoro, sono forniti nella tabella seguente:

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)	$\vartheta_{post-base}$	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 proprietà saliente	$P_{kj(s)}$	0.7
Probabilità proprietà marginale	$P_{kj(m)}$	0.35

Tabella 7. Valori numerici dei parametri utilizzati nel modello di lavoro.

2.5 Risultati

Di seguito sono presentati i risultati. Si noti che, per brevità, la trattazione che segue riguarda la sola tassonomia degli animali, ma risultati analoghi sono stati ottenuti anche per la tassonomia degli oggetti. Inizialmente vengono mostrati i *pattern* di sinapsi ottenuti in seguito alla fase di addestramento, che mettono in evidenza il diverso ruolo delle proprietà condivise rispetto alle proprietà distintive e delle proprietà salienti rispetto a quelle marginali. Successivamente vengono presentate alcune simulazioni riguardo compiti cognitivi di riconoscimento degli animali.

2.5.1 Risultati addestramento sinaptico

I valori delle sinapsi ottenuti dopo l'addestramento incorporano le differenze tra caratteristiche condivise e distintive e quelle tra caratteristiche salienti e marginali. Ogni figura rappresenta la forza delle sinapsi che entrano in una proprietà dagli altri neuroni dell'area semantica dopo un addestramento di 1000 epoche. Nei grafici sottostanti, in ascissa sono indicate le posizioni relative a ciascuna proprietà (in totale 231) mentre in ordinata è indicato il peso sinaptico il cui valore massimo è 0.8. Per non appesantire la trattazione vengono considerate solo alcune caratteristiche scelte dall'animale 10 (cavallo). In particolare vengono considerate:

- una caratteristica saliente totalmente condivisa che appartiene alla categoria principale ANIMALE (es. “mangia”);
- una caratteristica saliente totalmente condivisa nella sotto categoria MAMMIFERO (es. “ha quattro zampe”);

- una caratteristica saliente totalmente condivisa nella sotto categoria gerarchicamente inferiore rispetto alla precedente, DA ALLEVAMENTO (es. “vive in fattoria”)
- una caratteristica saliente distintiva (es. “nitrisce”);
- una caratteristica marginale distintiva (es. “veloce”);

Come si evince dal *pattern* in figura 3, una proprietà totalmente condivisa appartenente alla categoria principale ANIMALE riceve sinapsi da tutte le proprietà della rete tranne che da sé stessa. La forza di ciascuna sinapsi dipende dalla frequenza di occorrenza delle caratteristiche: le sinapsi che arrivano da caratteristiche salienti hanno una forza maggiore, mentre quelle che arrivano da caratteristiche marginali sono più deboli poiché queste proprietà sono raramente evocate.

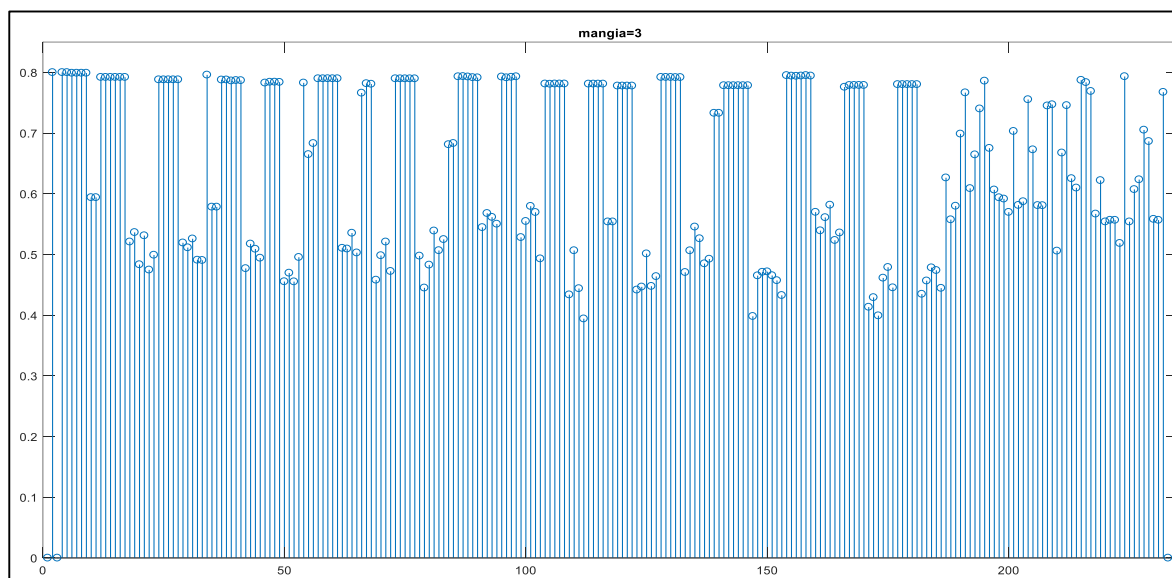


Figura 3 - Forza delle sinapsi entranti nella proprietà "mangia" dagli altri neuroni dell'area semantica.

Una proprietà totalmente condivisa appartenente alla sotto categoria MAMMIFERO (es. “ha quattro zampe” che occupa la posizione 7 nella rete) mostra un *pattern* simile al precedente ma ora questa proprietà riceve sinapsi da tutte le altre proprietà nella stessa categoria e dalle proprietà dei dieci animali che vi appartengono, mentre non riceve sinapsi dalle proprietà di una categoria a un livello gerarchico superiore, né dai membri in altre categorie (*fig. 4*).

Anche in questo caso la forza sinaptica dipende dalla frequenza di occorrenza delle caratteristiche.

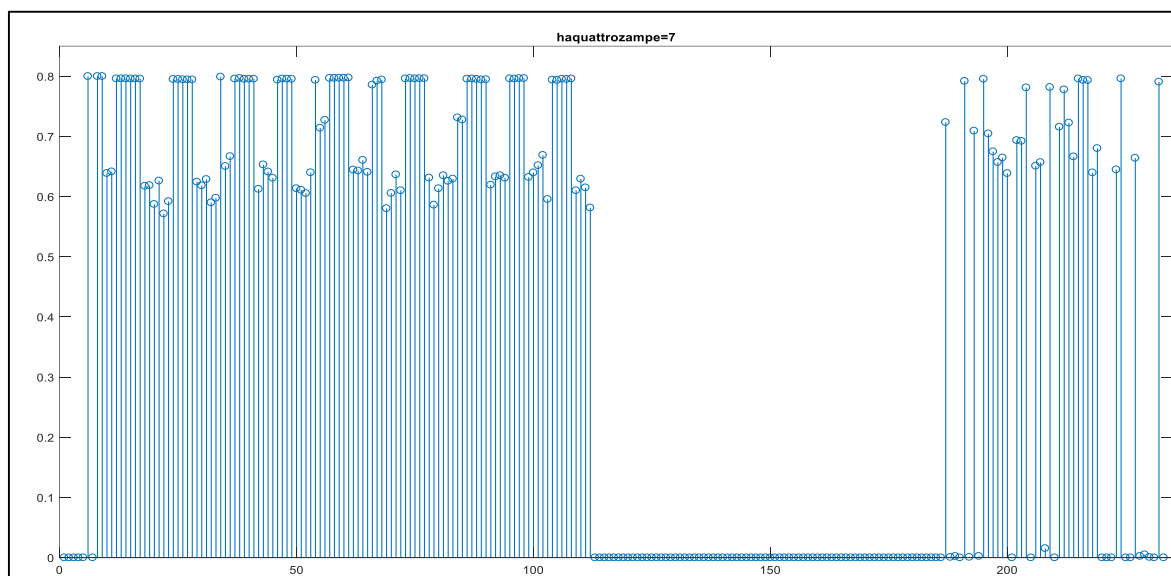


Figura 4 - Forza delle sinapsi entranti nella proprietà "ha quattro zampe" dagli altri neuroni dell'area semantica.

Un comportamento simile si ha per le altre proprietà che formano sotto categorie come "mangia erba" per ERBIVORO o "vive nella fattoria" per DA ALLEVAMENTO. In particolare dal *pattern* della proprietà "vive nella fattoria" (posizione 85) si evince che una proprietà di sotto categoria riceve sinapsi (sempre pesate in base alla salienza) dall'altra proprietà della stessa categoria (indicata con una freccia nel grafico) e dalle proprietà degli altri animali appartenenti alla categoria DA ALLEVAMENTO (pecora indicata con colore blu, mucca indicata con colore verde e cavallo indicata con colore rosso), mentre non riceve sinapsi dalle proprietà di una categoria a un livello gerarchico superiore, né dai membri in altre categorie (*fig. 5*). Una proprietà saliente distintiva dell'animale ("nitrisce" che si trova in posizione 108 nella rete) riceve sinapsi da tutte le altre proprietà distintive dell'animale tranne che da sé stessa e sinapsi trascurabili (praticamente nulle) da proprietà condivise (*fig. 6*).

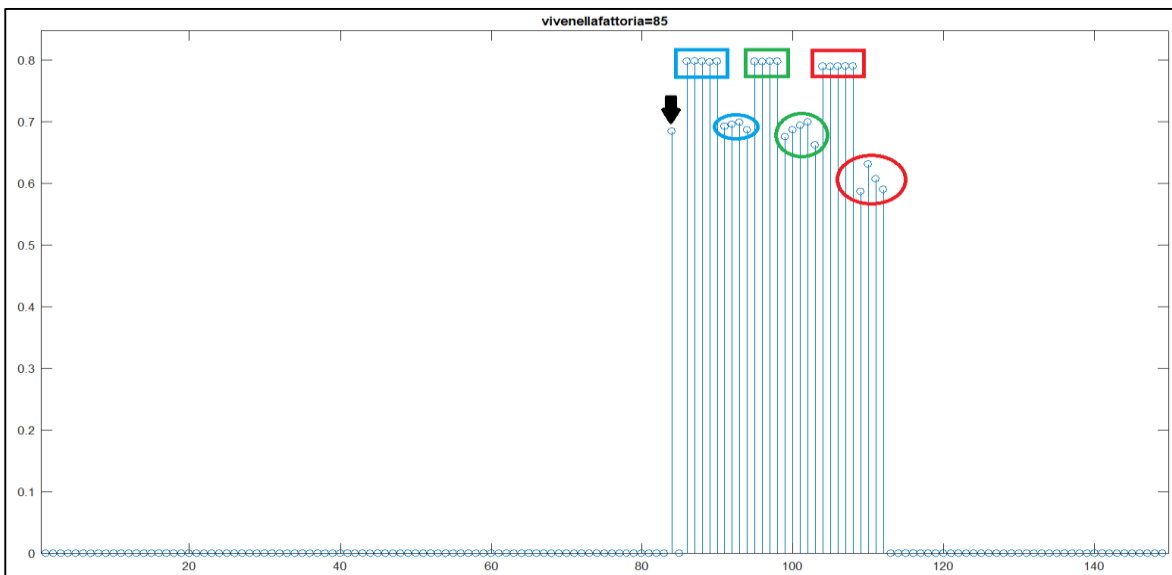


Figura 5 - Forza delle sinapsi entranti nella proprietà "vive nella fattoria" dagli altri neuroni dell'area semantica.

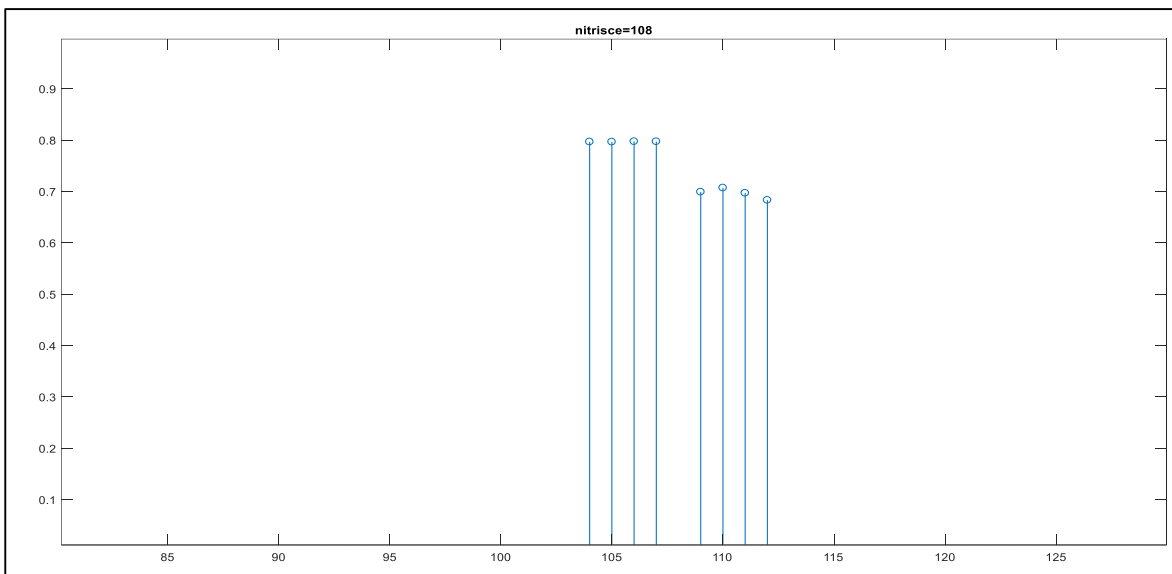


Figura 6 - Forza delle sinapsi entranti nella proprietà "nitrisce" dagli altri neuroni dell'area semantica.

Infine, una caratteristica marginale distintiva (es. "veloce" che occupa la posizione 110 nella rete) riceve sinapsi trascurabili (praticamente nulle) (fig. 7). La conseguenza è che le caratteristiche marginali non sono evocate spontaneamente dalle altre caratteristiche dello stesso animale, cioè non vengono portate alla mente spontaneamente quando si pensa a quest'ultimo.

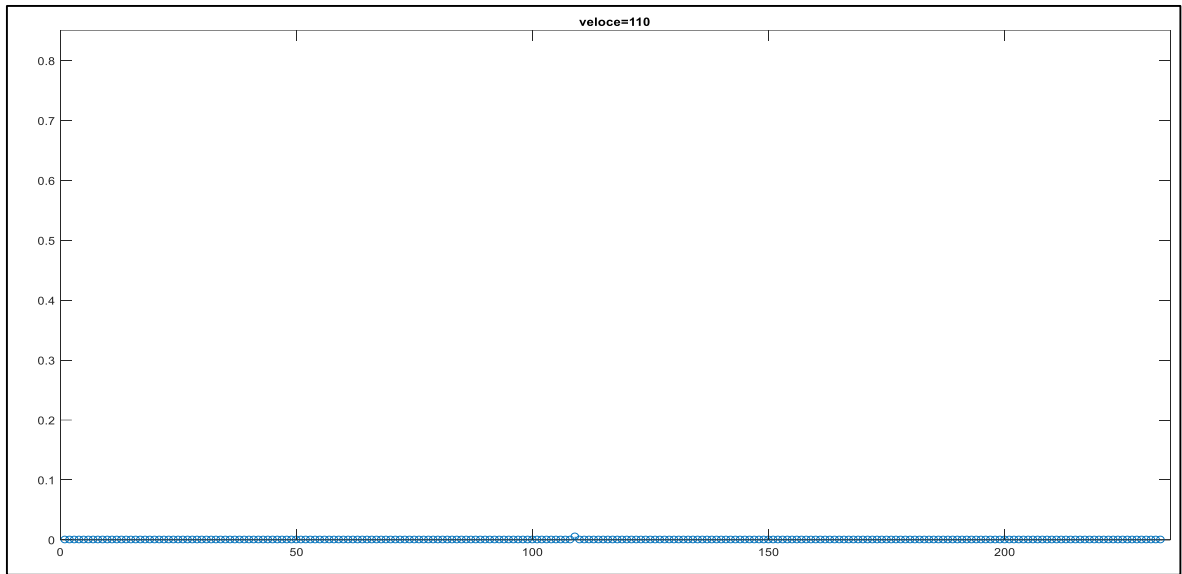


Figura 7 - Forza delle sinapsi entranti nella proprietà "veloce" dagli altri neuroni dell'area semantica.

2.5.2 Risultati simulazioni di riconoscimento animali

Il modello addestrato è stato successivamente testato attraverso delle simulazioni di compiti di riconoscimento di animali. Durante tali prove, sono state fornite in input alla rete alcune (da 1 fino ad un massimo di 2) proprietà di un animale e i risultati mostrano quali proprietà vengono spontaneamente evocate. In particolare il modello correttamente addestrato, in seguito all'input, dovrebbe evocare tutte le proprietà salienti del concetto in questione (distintive e condivise), mentre non dovrebbe richiamare le proprietà marginali (distintive e condivise). Inoltre, date in input proprietà condivise (totalmente o parzialmente), il modello dovrebbe essere in grado di formare delle categorie. Di seguito vengono riportati alcuni esempi e mostrati alcune istantanee dei *pattern* prese a 10, 20 e 50 *step* per chiarire la spiegazione. In ciascun pannello in ascissa vengono indicate le posizioni relative alle proprietà totali (231) mentre in ordinata è indicata l'attività neuronale normalizzata a 1. Per semplicità la trattazione viene eseguita solo su alcune caratteristiche dell'animale "cavallo", ma considerazioni analoghe possono essere estese a tutti i concetti presenti nella tassonomia di animali.

Dando in ingresso al modello una caratteristica distintiva saliente (es. “nitrisce” posizione 108) si nota che vengono richiamate tutte le caratteristiche salienti associate al concetto di “cavallo”. In particolare vengono richiamate le altre proprietà salienti distintive dell’animale (dalla 104 alla 108) e le proprietà salienti parzialmente condivise (203,214,218,219,223,228). Infine vengono richiamate le proprietà salienti totalmente condivise che vanno a formare le categorie gerarchicamente superiori al concetto: per la classe DA ALLEVAMENTO vengono evocate le proprietà 84,85, per la classe ERBIVORO la proprietà 54, per la classe SELVATICO la 34, per la categoria MAMMIFERO le proprietà dalla 6 alla 9 ed infine per la generica categoria ANIMALE le proprietà dalla 2 alla 5 (*fig. 8*). Si noti che la sinapsi corrispondente alla proprietà che viene data come input alla rete (108) raggiunge la saturazione più velocemente rispetto alle altre.

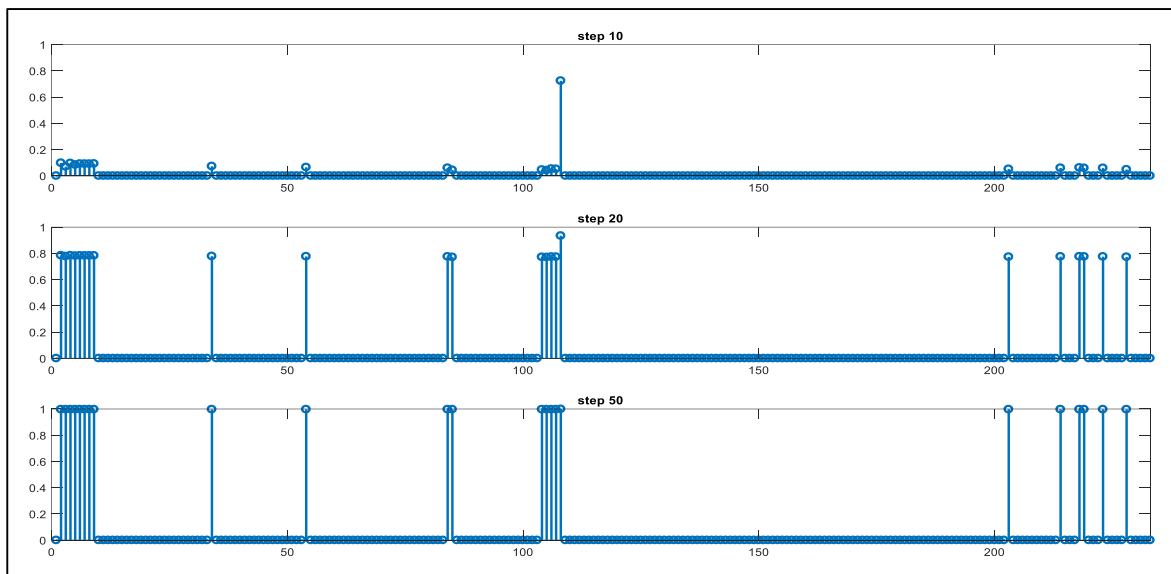


Figura 8 - Pattern relativi al riconoscimento del concetto "cavallo" data in input la proprietà saliente distintiva "nitrisce".

Dando in ingresso una caratteristica distintiva marginale (es. “veloce” posizione 110) il comportamento del modello è analogo al precedente. L’unica differenza è che nel *pattern* è presente anche la caratteristica distintiva marginale (indicata in rosso in *fig.9*) che spunta non perché evocata dalla rete ma perché usata come input esterno.

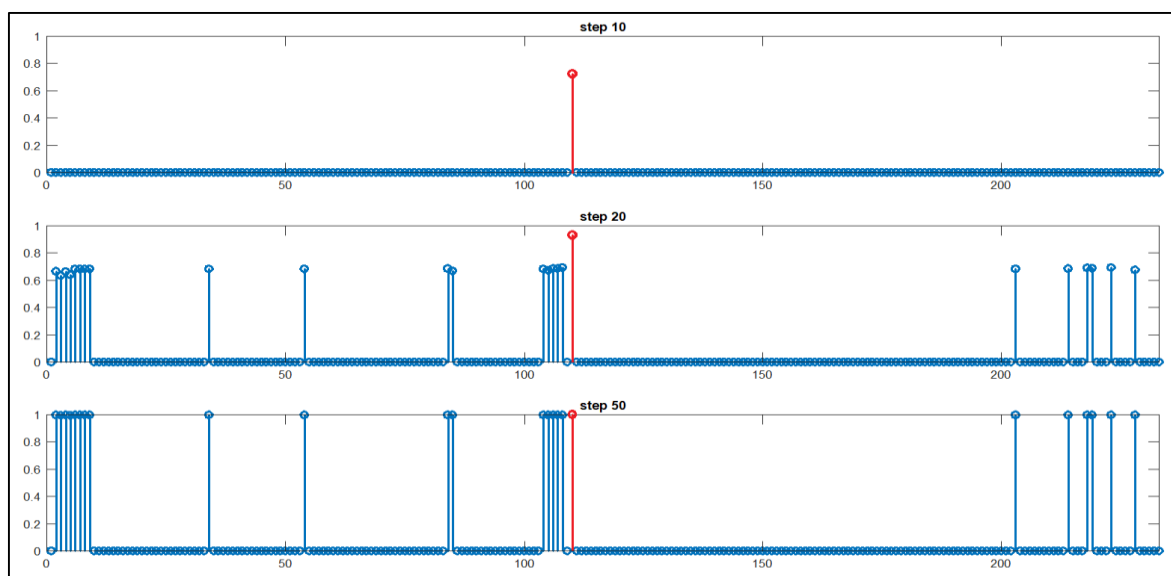


Figura 9 - Pattern relativi al riconoscimento del concetto "cavallo" data in input la proprietà marginale distintiva "veloce" (rossa).

Dando in ingresso al modello una caratteristica totalmente condivisa, cioè una caratteristica che forma una categoria, il modello evoca spontaneamente le altre caratteristiche appartenenti a quella categoria e le proprietà appartenenti a categorie gerarchicamente superiori. Le proprietà salienti distintive di un concetto subordinato e quelle salienti parzialmente condivise non vengono invece richiamate. Ad esempio, dando come ingresso la proprietà "vive nella fattoria" (85) appartenente alla categoria DA ALLEVAMENTO, il modello evoca spontaneamente l'altra proprietà della categoria (84) e le proprietà delle categorie gerarchicamente superiori: la 54 per ERBIVORO, la 34 per SELVATICO, dalla 6 alla 9 per MAMMIFERO e infine dalla 2 alla 5 per ANIMALE. Non vengono invece evocate le proprietà salienti dei concetti subordinati (*fig.10*). Infine si analizza il comportamento del modello quando come input viene data una caratteristica parzialmente condivisa di un dato concetto. La proprietà parzialmente condivisa scelta come input può essere marginale o saliente, il risultato è analogo: il modello evoca tutte le proprietà totalmente condivise delle categorie che accomunano i concetti che condividono la proprietà data in input. Per chiarire, si considera la proprietà "corre" condivisa da "cavallo", "zebra" e "leone" e non appartenente a "orso", "giraffa", "elefante", "pecora" e "mucca".

Come risulta dai *pattern* in *fig. 11* nella pagina successiva, il modello riesce a richiamare le proprietà salienti totalmente condivise della categoria ANIMALE (dalla 2 alla 5), della categoria MAMMIFERO (dalla 6 alla 9) e della categoria SELVATICO (34). Le proprietà salienti proprie di un concetto subordinato (es. “cavallo”) non sono invece richiamate perché il modello non ha sufficienti informazioni per discriminare il concetto dagli altri concetti che condividono la proprietà data come input. Al contrario se in input viene data una caratteristica distintiva di un concetto (es. “bruno” per il concetto di “orso”), la rete richiama correttamente le caratteristiche salienti del concetto (oltre a quelle di categoria), ma non evoca la proprietà parzialmente condivisa “corre” (223), come mostrato in *fig. 12* nella pagina seguente.

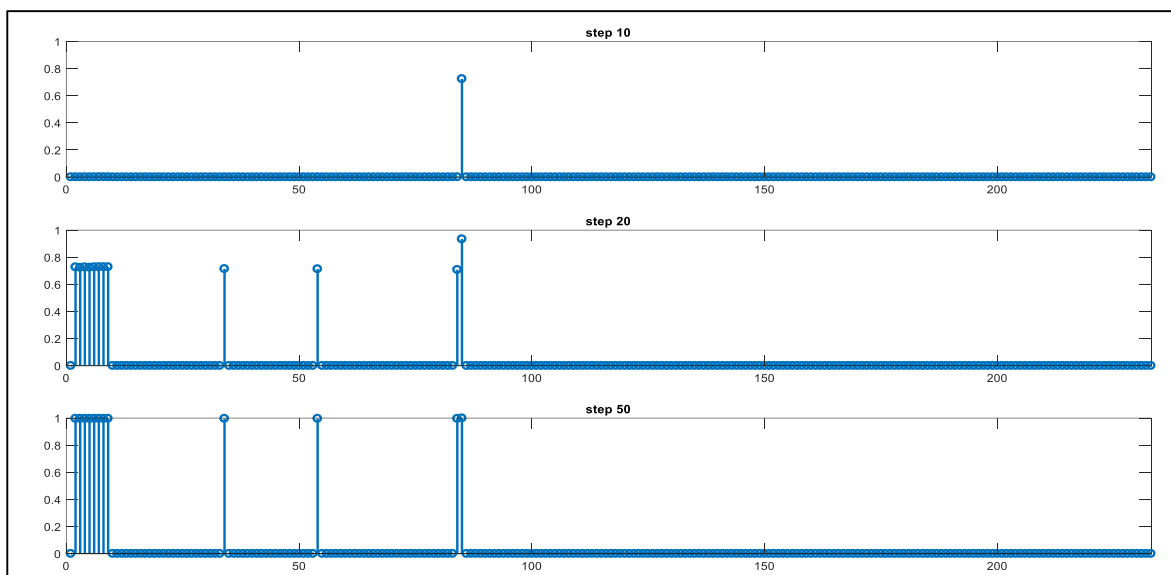


Figura 10 - Pattern relativi alla formazione di categorie data in input la proprietà totalmente condivisa “vive nella fattoria”.

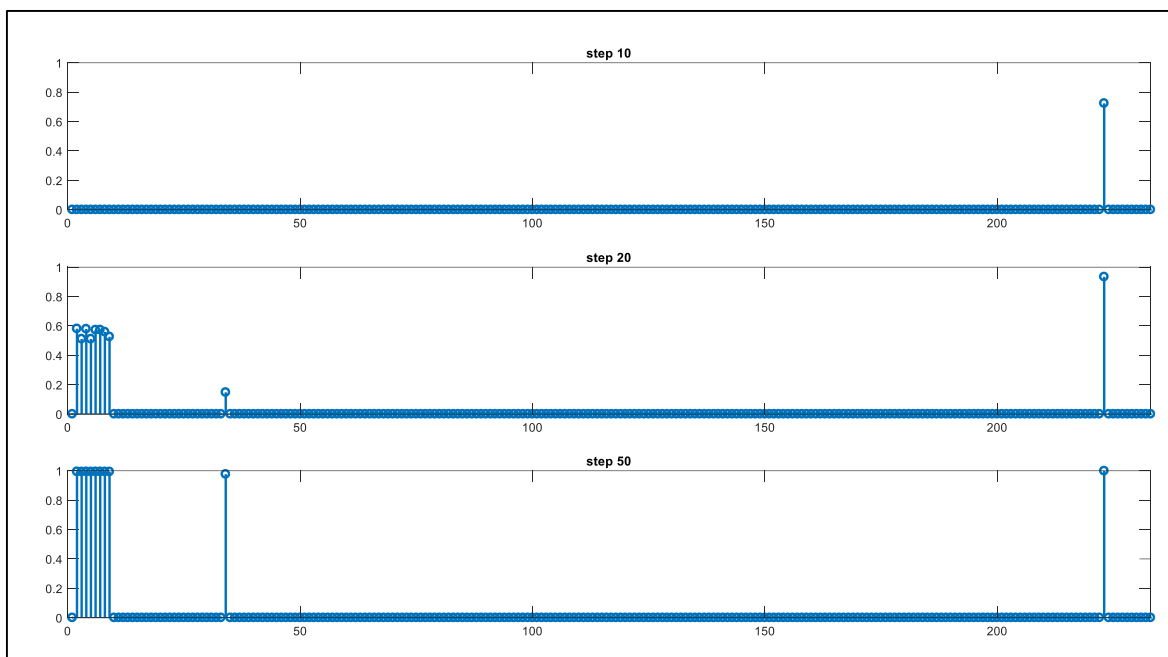


Figura 11 - Pattern relativi alla formazione di categorie data in input la proprietà parzialmente condivisa “corre”.

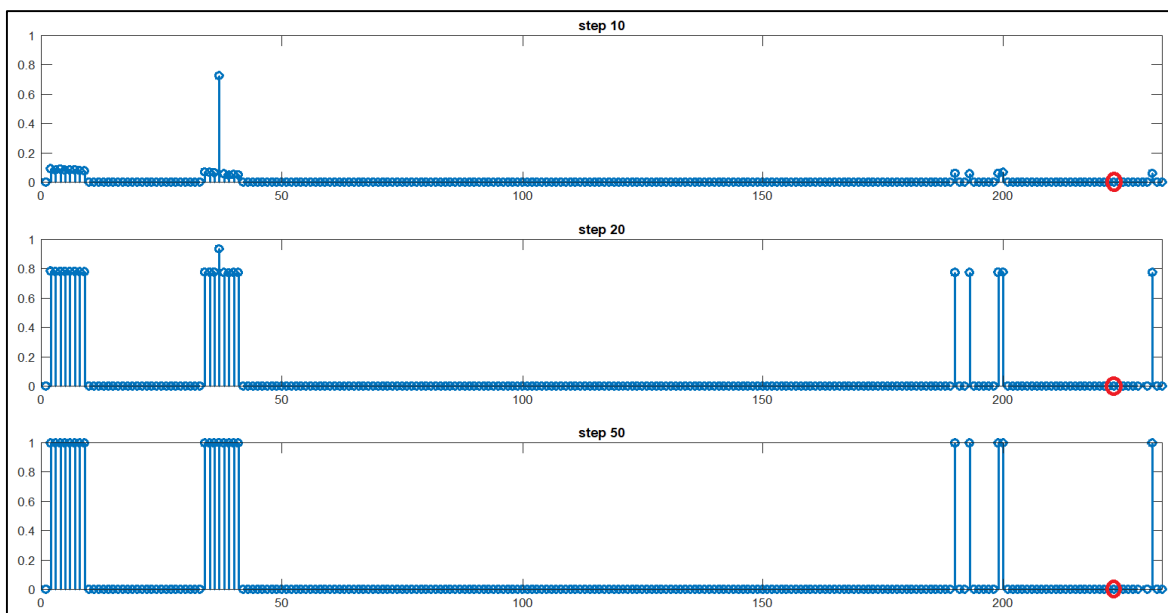


Figura 12 - Proprietà richiamate dalla rete se in input viene data la feature “bruno” (distintiva per il concetto di “orso”). Il cerchio rosso mette in evidenza che in questo caso la proprietà parzialmente condivisa “corre” (223) non è evocata.

Le prestazioni complessive del modello possono essere riassunte nel modo seguente:

Riconoscimento dei singoli concetti: se vengono stimulate le caratteristiche distintive (saliente o marginale), il modello riconosce il concetto corrispondente evocando solo le sue caratteristiche salienti (sia distintive che condivise) senza evocare quelle marginali (distintive e condivise) né quelle di altri concetti. Solo in alcuni casi si è verificato che una caratteristica marginale divenisse saliente e venisse quindi evocata dalla rete. Tale fenomeno è da attribuire esclusivamente alla randomizzazione delle sinapsi. In generale infatti l'unico modo per eccitare una caratteristica marginale è darla come input esterno. In questo caso essa contribuisce all'eccitazione delle caratteristiche salienti e aiuta il riconoscimento del concetto. In particolare i risultati mostrano che risulta sufficiente una caratteristica distintiva (marginale o saliente) affinché il modello funzioni correttamente per tutti e 16 gli animali della tassonomia.

Formazione di una categoria: se vengono stimulate le caratteristiche condivise (totalmente o parzialmente), la rete riconosce correttamente la categoria e non evoca le caratteristiche distintive dei singoli animali, né le caratteristiche parzialmente condivise se non quando utilizzate come input esterno. Anche in questo caso i risultati mostrano che è sufficiente una sola caratteristica totalmente condivisa affinché vengano formate sia la categoria della quale fa parte la proprietà in questione sia le categorie gerarchicamente superiori ad essa. Si noti che se da questa prestazione si vuole ottenere quella descritta al punto precedente, basta dare in input alla rete, oltre che una caratteristica condivisa, anche una caratteristica distintiva propria del concetto che si vuole discriminare rispetto agli altri.

CAPITOLO 3

MODELLO DI MEMORIA SEMANTICA CON OSCILLATORI

3.1 “*Binding and segmentation*” e sincronismo in banda γ

La percezione simultanea di differenti concetti è alla base delle attività che ciascun individuo deve svolgere ogni giorno. Affinché i vari concetti presentati simultaneamente in una scena riescano ad essere riconosciuti adeguatamente, il cervello deve risolvere nello stesso tempo due problemi complementari. Il primo consiste nel raggruppare insieme attributi codificati in diverse aree corticali e appartenenti allo stesso concetto (*binding*), il secondo consiste nel tenere tali attributi separati da quelli di altri concetti memorizzati simultaneamente (*segmentation*). Inizialmente si riteneva che il cervello risolvesse questi problemi attraverso processi *feed-forward* strutturati gerarchicamente, in modo tale che i singoli attributi di un medesimo concetto confluissero ad eccitare neuroni posti a un livello gerarchico superiore, in grado di codificare il concetto medesimo. Secondo tale ipotesi, l'eccitazione dei singoli neuroni, gerarchicamente avanzati, segnalerebbe la presenza del concetto. Questa idea tuttavia presenta diversi limiti. Tra gli altri, la necessità di avere un numero di neuroni gerarchicamente superiori troppo elevato, considerando tutte le possibili combinazioni di attributi. Ad oggi si ritiene che il cervello riesca a fare fronte al problema del “*binding and segmentation*” grazie al meccanismo chiamato “*assembly coding*”: le proprietà che codificano un medesimo concetto devono essere simultaneamente attive, mentre proprietà di un concetto diverso simultaneamente inibite. Ciò può essere ottenuto attraverso il sincronismo neuronale secondo cui neuroni che codificano gli attributi di un medesimo concetto percepito oscillano con una frequenza propria della banda gamma (intorno ai 30 Hz) in sincronismo fra loro (cioè si eccitano e si inibiscono tutti in fase) mentre l'attività dei neuroni che codificano attributi di un concetto diverso risultano sfasate rispetto alle precedenti.

La sincronizzazione neuronale può essere distribuita sia nella stessa area corticale, sia tra aree corticali distanti e dipende interamente dalle connessioni interne tra i neuroni. I concetti quindi 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 oggetti contemporaneamente [5,6,7].

3.2 Equazioni associate all'oscillatore di Wilson – Cowan

Gli oscillatori utilizzati nei modelli di seguito descritti sono oscillatori di Wilson – Cowan. Se opportunamente stimolato, un oscillatore di Wilson-Cowan è in grado di produrre un'oscillazione grazie alla presenza di connessioni in *feedback* tra una unità eccitatoria x_i ed una unità inibitoria y_i , come descritto in *fig. 13*.

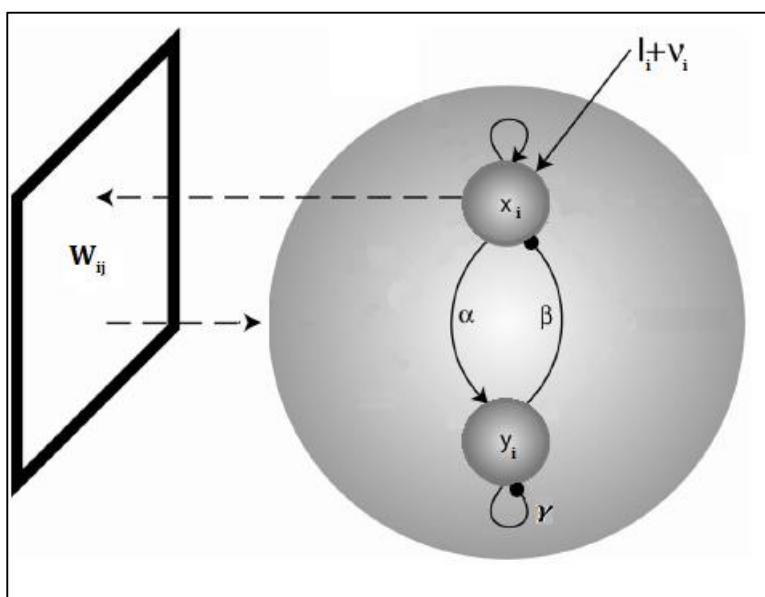


Figure 13 - Oscillatore di Wilson – Cowan. W_{ij} rappresenta la matrice delle sinapsi eccitatorie tra i vari oscillatori. I è l'input esterno, corrotto da rumore. α , β e γ sono parametri il cui significato viene spiegato nel testo.

Le equazioni che descrivono come varia nel tempo l'oscillazione del generico oscillatore i-esimo sono:

$$\frac{d}{dt}x_i(t) = -x_i + H(x_i(t) - \beta y_i(t) + E_i(t) + I_i(t) + v_i(t) - \varphi_x) \quad (6)$$

$$\frac{d}{dt}y_i(t) = -\gamma y_i(t) + H(\alpha x_i(t) - \varphi_y) + J_i(t) \quad (7)$$

dove $H(\psi)$ rappresenta la funzione sigmoideale di attivazione definita come

$$H(\psi) = \frac{1}{1+e^{-\frac{\psi}{T}}} \quad (8)$$

I parametri presenti nelle espressioni (6) e (7) vengono spiegati di seguito. α e β sono due parametri positivi che definiscono rispettivamente l'accoppiamento tra unità eccitatoria ed inibitoria e tra unità inibitoria ed eccitatoria dell'oscillatore. In particolare α influenza significativamente l'ampiezza dell'oscillazione. Il parametro γ è l'inverso di una costante di tempo ed influenza la frequenza di oscillazione. L'auto-eccitazione di x è posta a 1 per stabilire una scala per i pesi sinaptici. Anche la costante di tempo di x è posta a 1 e rappresenta una scala per il tempo. φ_x e φ_y sono termini di offset per le funzioni sigmoideali delle unità rispettivamente eccitatoria ed inibitoria. I_i esprime l'input esterno dell'i-esimo oscillatore, mentre v rappresenta un rumore gaussiano non uniformemente distribuito [4,9]. Infine E_i e J_i simboleggiano termini di accoppiamento tra tutti gli oscillatori della rete:

$$E_i = \sum_{j=1}^N W_{E_{ij}} x_j \quad (9)$$

$$J_i = \sum_{j=1}^N W_{J_{ij}} x_j \quad (10)$$

dove N è il numero totale di oscillatori, $W_{E_{ij}}$ e $W_{J_{ij}}$ sono, rispettivamente, le sinapsi eccitatorie e inibitorie che collegano l'oscillatore j -esimo (pre) con l'oscillatore i -esimo (post). In forma vettoriale i due termini di accoppiamento risultano

$$E = (W_E)X \text{ e } J = (W_J)X \quad (11 \text{ e } 12)$$

dove X è un vettore di dimensione $N \times 1$ che contiene gli output dell'oscillatore, E e J sono i vettori dei termini di accoppiamento entrambi di dimensione $N \times 1$, W_E e W_J sono le due matrici, entrambe di dimensioni $N \times N$ (fig. 14) che sono state ricavate ricalcando la matrice delle sinapsi W ottenuta dall'addestramento sinaptico. La prima contiene i pesi sinaptici (da noi impostati) che vanno ad agire sulla unità eccitatoria, mentre la seconda contiene i pesi sinaptici (da noi impostati) che vanno ad agire sulla unità inibitoria.

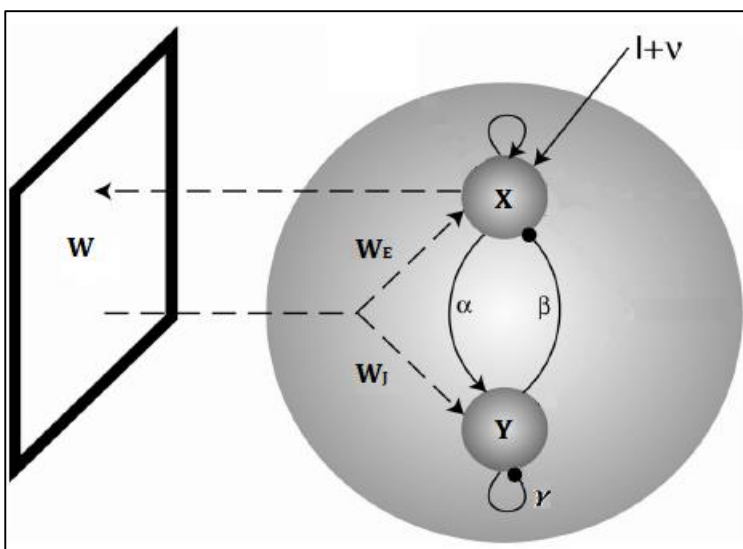


Figura 14 - Oscillatore Wilson-Cowan con in dettaglio le matrici W_E e W_J .

3.3 Sincronismo tra 10 oscillatori: scelta dei parametri e risultati

Partendo dalle equazioni (6) e (7), sono stati implementati 10 oscillatori di Wilson- Cowan. I valori dei parametri adoperati per l'implementazione di ciascun oscillatore, sono stati settati pari a quelli adoperati in uno studio precedente [9]. Si sono dovuti invece trovare sperimentalmente i valori ottimali da attribuire alle componenti delle matrici W_E e W_J per ottenere un buon sincronismo tra tutti e 10 gli oscillatori. Dopo diverse prove, effettuate a partire da uno stato iniziale pari a un valore casuale preso nell'intervallo $[0,1]$ e pre-moltiplicato per una costante pari a 0.6 per tutti e dieci gli oscillatori, è stato verificato che settando le sinapsi eccitatorie (contenute in W_E) ad un valore pari a 2 e le sinapsi inibitorie (contenute in W_J) ad un valore pari a 0.5, tutti gli oscillatori sono tra loro sincronizzati.

Con tali valori sinaptici inoltre tutti gli oscillatori mostrano un buon sincronismo per ogni combinazione data in ingresso: partendo dalla stimolazione di un solo oscillatore (un ingresso impostato a 1 e restanti impostati a 0) fino alla stimolazione di tutti e 10 gli oscillatori (dieci ingressi impostati a 1). A titolo di esempio viene riportato in *fig.15* il risultato ottenuto nel caso di 10 oscillatori tutti stimolati con un input unitario.

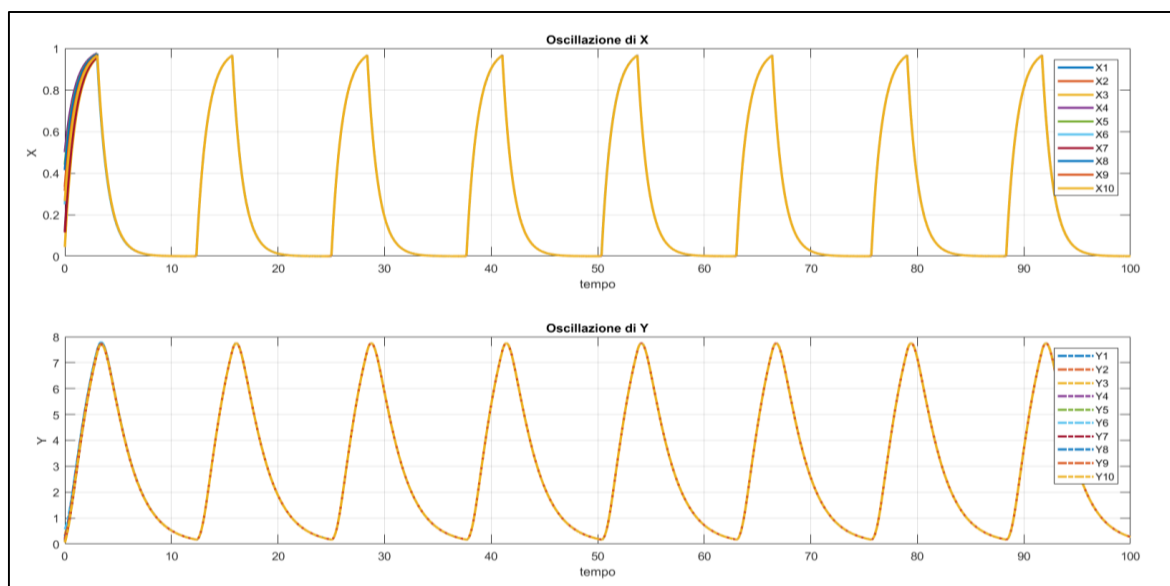


Figure 15 - Oscillazione dell'unità eccitatoria (X) ed inibitoria (Y) di 10 oscillatori di Wilson Cowan sincronizzati.

Nella tabella seguente sono riportati i valori numerici dei parametri utilizzati:

Significato	Simbolo	Valore
Parametri oscillatore Wilson-Cowan		
Forza sinaptica da unità eccitatoria a unità inibitoria	α	0.3
Forza sinaptica da unità inibitoria a unità eccitatoria	β	2.5
Parametro che influenza frequenza di oscillazione	γ	0.5
Pendenza delle sigmoidi	T	0.025
Offset di H_x	φ_x	0.7
Offset di H_y	φ_y	0.15
Deviazione standard del rumore	σ	0.02
Parametri sinapsi		
Pesi sinaptici eccitatori	W_E	2
Pesi sinaptici inibitori	W_J	0.5

Tabella 8. Parametri utilizzati per la sincronizzazione di 10 oscillatori di Wilson – Cowan.

3.4 Modelli con oscillatori

Di seguito vengono analizzate reti semantiche implementate a partire dalla rete semantica di base (16 animali) e contenenti oscillatori di Wilson- Cowan al posto di ogni singolo neurone. Nel presente lavoro sono stati adoperati solo pochi concetti (da 1 a 3 animali con un massimo di 30 proprietà totali) per verificare se si riuscissero a fare oscillare in maniera sincrona tutte le proprietà salienti di un concetto e a desincronizzare i diversi concetti dati in ingresso simultaneamente. Studi con reti semantiche più ampie possono essere oggetto di lavori futuri.

3.4.1 Rete a 1 animale e 10 proprietà

Partendo dalla rete semantica degli animali di base (*cap. 2*) è stato scelto un solo animale con 10 proprietà distintive salienti come mostra lo schema seguente:



In questa nuova rete ogni neurone è stato sostituito con un oscillatore di Wilson – Cowan precedentemente descritto, ottenendo così un modello che si basa complessivamente sulle equazioni (1), (2), (6) e (7). Tale rete è stata successivamente addestrata attraverso la regola di Hebb. La fase di addestramento, che consiste in 1000 epoche consecutive, è del tutto analoga a quella effettuata per la rete di base: il concetto viene presentato una volta durante ogni epoca e in questo caso le caratteristiche hanno tutte la stessa probabilità adoperata come input per il dato concetto, ovvero $P_{kj(s)} = 0.7$. Inoltre, durante l'addestramento non si è fatto uso di neuroni oscillanti.

3.4.1.1 Assegnazione parametri

Durante l'addestramento sinaptico sono stati utilizzati gli stessi parametri, e i rispettivi valori, adoperati per l'addestramento della rete di base, ad eccezione della soglia post sinaptica basale il cui valore è stato abbassato a 0.55 visto il numero ridotto di proprietà presenti nella rete (10 vs 231) e quindi la minore possibilità che una proprietà marginale divenga casualmente saliente. A fine addestramento sono state implementate le due matrici W_E e W_J in corrispondenza dei collegamenti eccitatori ed inibitori ricalcando la matrice delle sinapsi W come descritto di seguito. Si è fissata una soglia pari a 0.7, si è confrontato ogni valore della matrice W con tale soglia. Se il generico elemento W_{ij} risultava maggiore di 0.7, si associava alla posizione corrispondente in W_E un valore pari a 0.6. Se invece il generico elemento W_{ij} risultava minore di 0.7, si associava alla posizione corrispondente in W_J un valore pari a 0.0. Si noti che si sono dovuti trovare sperimentalmente nuovi valori da dare alle sinapsi eccitatorie contenute nella matrice W_E e a quelle inibitorie contenute nella matrice W_J e si è dovuta aumentare la deviazione standard del rumore (σ) a 0.9 per evitare che, una volta eccitate, le proprietà entrassero in saturazione smettendo di oscillare. Inoltre si è verificato che la rete presente, così come quelle descritte in seguito in linea con il modello di base, lavora correttamente se le sinapsi inibitorie vengono trascurate ($W_{J_{ij}} = 0 \forall i, j$). Per questo motivo da questo punto della trattazione in avanti la matrice W_J non verrà più presa in considerazione, questo corrisponde all'omissione del termine $J_i(t)$ nell'equazione (7). In generale tutti i valori numerici utilizzati nella rete sono riportati nella tabella nella pagina seguente, dove sono evidenziati i termini che hanno subito variazioni.

Significato	Simbolo	Valore
Parametri oscillatore Wilson-Cowan		
Forza sinaptica da unità eccitatoria a unità inibitoria	α	0.3
Forza sinaptica da unità inibitoria a unità eccitatoria	β	2.5
Parametro che influenza frequenza di oscillazione	γ	0.5
Pendenza delle sigmoidi	T	0.025
Offset di H_x	φ_x	0.7
Offset di H_y	φ_y	0.15
Deviazione standard del rumore	σ	0.9
Parametri sinapsi		
Pesi sinaptici eccitatori	W_E	0.6
Soglia	Soglia	0.7
Parametri modello		
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)	$\vartheta_{post-base}$	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 proprietà saliente	$P_{kj(s)}$	0.7

Tabella 9. Parametri utilizzati per rete semantica a 1 animale e 10 proprietà salienti

3.4.1.2 Risultati addestramento e simulazioni

Il modello correttamente addestrato, in seguito all'input, dovrebbe evocare tutte le proprietà salienti del concetto in questione, mentre non dovrebbe richiamare le proprietà marginali. Inoltre, una volta richiamate, le proprietà salienti dovrebbero oscillare in modo sincrono. Tutte le prove sono state effettuate a partire da uno stato iniziale pari ad un valore random preso tra 0 e 1 per ogni oscillatore. Durante le simulazioni sono stati eccitati alcuni degli oscillatori corrispondenti alle proprietà del concetto e i risultati mostrano che il modello riesce a richiamare tutte le altre proprietà salienti del concetto e a farle oscillare in modo sincrono.

Di seguito viene riportato un esempio al fine di chiarire quanto appena detto. Viene considerato il concetto “mucca”. Nei *pattern* in fig. 16 e fig.17 è raffigurato il comportamento della rete se viene dato un input $I = 1$ a un oscillatore (proprietà “muggisce”) e $I = 0$ agli altri (proprietà “bovino”, “mammelle”, “ruminante”, “commestibile”, “ha coda con ciuffo”, “ha le macchie”, “ha le corna”, “pascola” e “si ricava il latte”): l’unità eccitata oscilla e fa oscillare in fase tutte le proprietà salienti del medesimo animale. Il risultato ottenuto può essere esteso a tutte le combinazioni di ingresso possibili: partendo dalla stimolazione di un solo oscillatore (un ingresso impostato a 1 e restanti impostati a 0) fino alla stimolazione di tutti e 10 gli oscillatori (dieci ingressi impostati a 1).

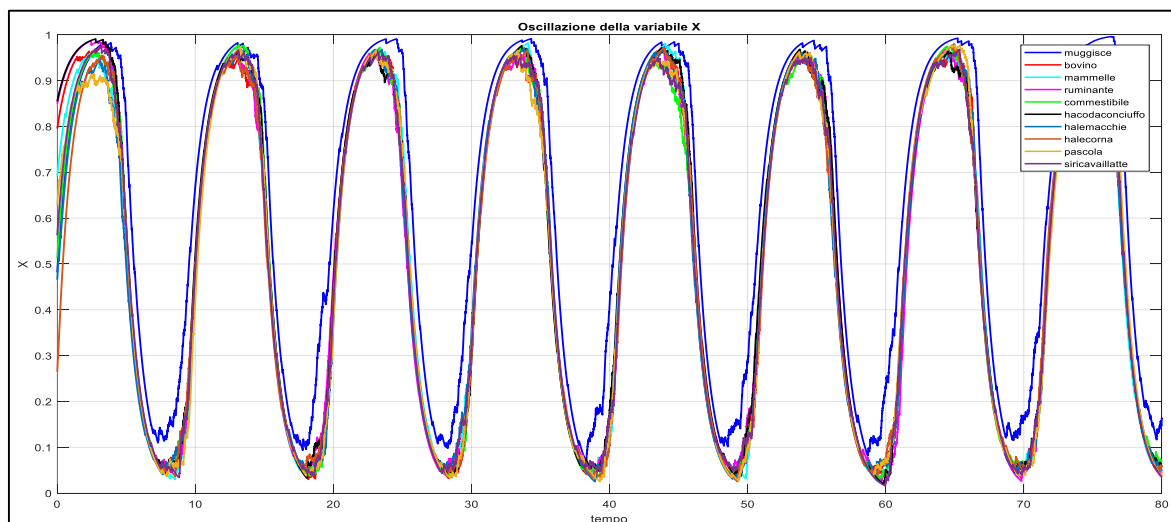


Figura 16 - Oscillazione delle unità eccitatorie (X) delle proprietà salienti del concetto “mucca”.

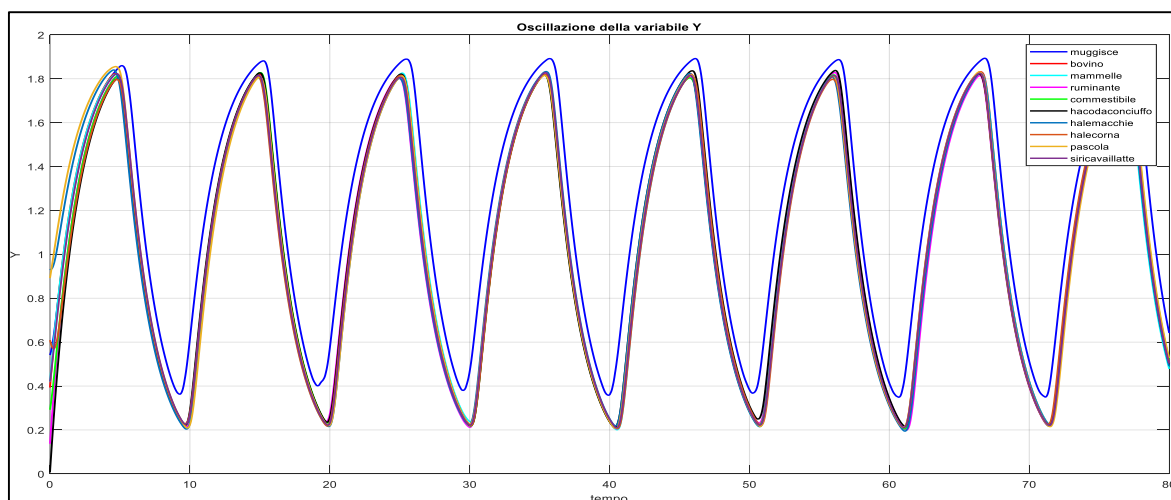


Figura 17 - Oscillazione delle unità inibitorie (Y) delle proprietà salienti del concetto “mucca”.

Per completezza si fa notare che nel caso in cui alcune delle dieci proprietà (in particolare da 1 fino ad un massimo di 3) fossero state etichettate come marginali ($P_{kj(m)} = 0.35$), il risultato sarebbe stato il seguente: la rete, qualunque fosse stato l'ingresso, avrebbe richiamato tutte le proprietà salienti del concetto e le avrebbe fatte oscillare in maniera sincrona. Le proprietà marginali invece non sarebbero state evocate, se non nel caso in cui usate come input esterno (fig. 18 e 19).

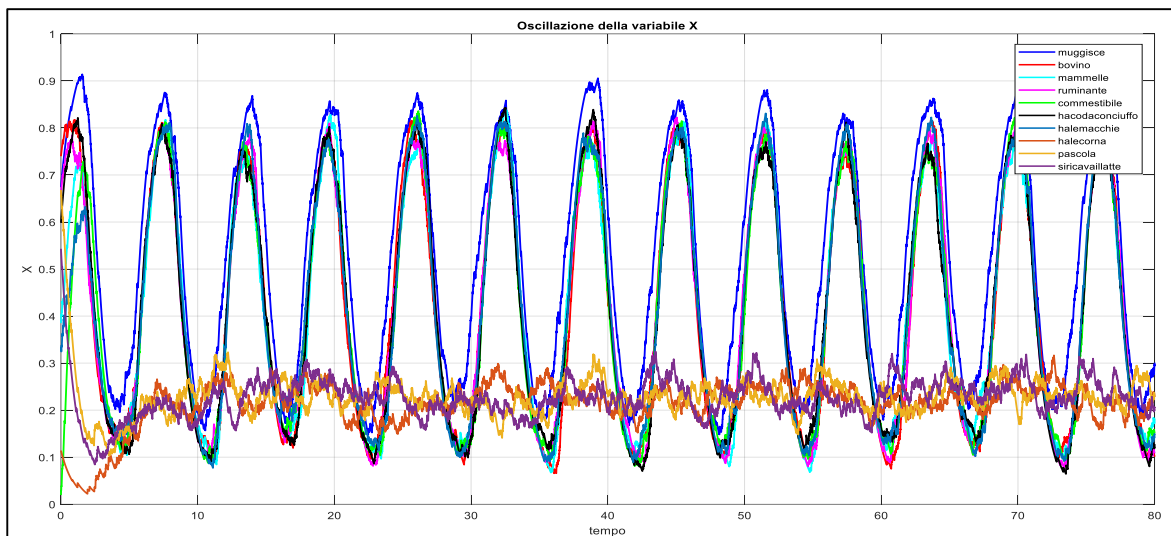


Figura 18 - Oscillazione delle unità eccitatorie (X) delle proprietà salienti del concetto "mucca". Si noti che le proprietà marginali ("ha le corna", "pascola", "si ricava il latte") non vengono richiamate dal modello e non oscillano in modo sincrono.

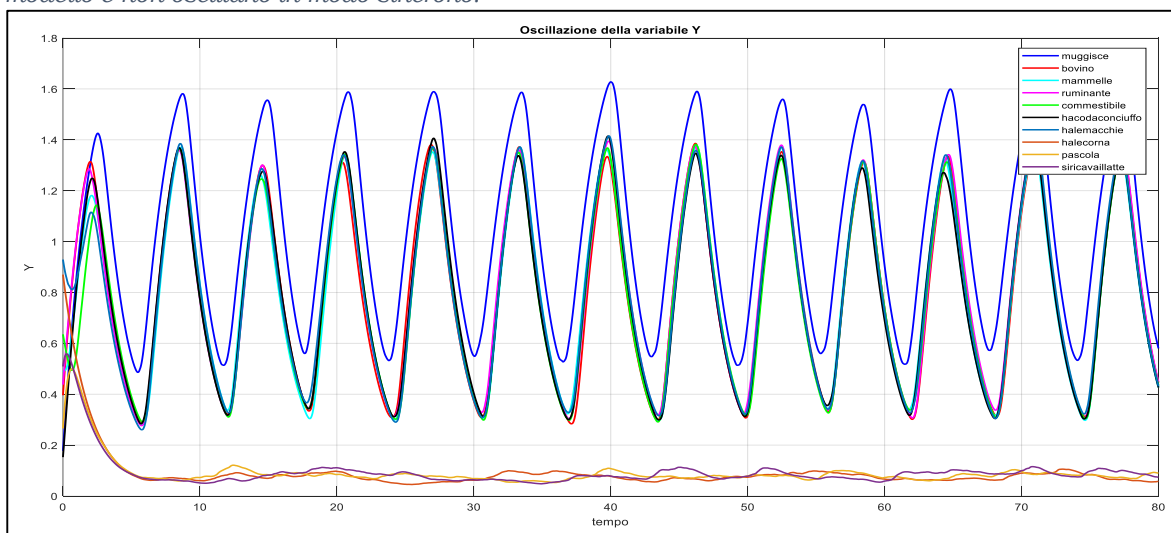


Figura 19 - Oscillazione delle unità inibitoria (Y) delle proprietà salienti del concetto "mucca". Si noti che le proprietà marginali ("ha le corna", "pascola", "si ricava il latte") non vengono richiamate dal modello e non oscillano in modo sincrono.

3.4.2 Rete a più animali: riconoscimento simultaneo di concetti

Partendo dai risultati sopra ottenuti, si è implementata una rete contenete più di un singolo concetto per verificare che, oltre ad attivare tutte le proprietà salienti dello stesso concetto in maniera sincrona, la rete riesca anche a riconoscere più concetti presentati simultaneamente. Questo ultimo obiettivo si ottiene desincronizzando i vari concetti dati in input, cioè facendo in modo che le proprietà di un concetto diverso risultino sfasate rispetto alle precedenti. A tal proposito, si è introdotto nel modello un inibitore globale (fig.20). L'inibitore globale presenta una unità inibitoria che riceve la somma di tutte le attività eccitatorie della rete semantica (z) ed invia un segnale inibitorio ($Inib$) se questa somma supera una certa soglia (th). In questo modo quando un concetto viene evocato dalla rete, tutti gli altri concetti sono momentaneamente inibiti. Le equazioni associate ad ogni oscillatore diventano:

$$\frac{d}{dt}x_i(t) = -x_i(t) + H(x_i(t) - \beta y_i(t) + E_i(t) + I_i(t) + v_i(t) - \varphi_x) \quad (6)$$

$$\frac{d}{dt}y_i(t) = -\gamma y_i(t) + H(\alpha x_i(t) - \varphi_y) + Inib(t) \quad (7')$$

$$\text{dove } Inib(t) = \begin{cases} k \text{ se } z(t) > th & \text{con } z(t) = \sum_i x_i(t) \\ 0 \text{ altrimenti} \end{cases}$$

il termine k indica un valore settato di caso in caso e riportato nelle successive tabelle dei parametri.

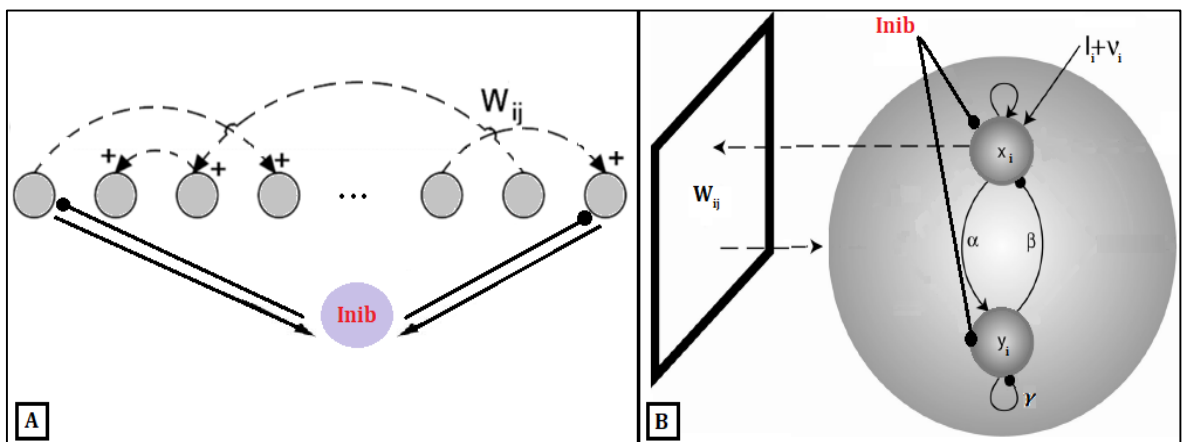


Figura 20 - Diagramma schematico che descrive il modello. (Panel A) Catena di oscillatori di Wilson-Cowan. Ogni oscillatore riceve sinapsi eccitatorie dagli altri oscillatori e un segnale inibitorio dall'inibitore globale ($Inib$). (Panel B) Dettaglio di ogni singolo oscillatore.

3.4.2.1 Rete con 3 animali: implementazione e risultati

Partendo dalla rete degli animali di base (*cap. 2*), sono stati scelti tre animali, limitando l'analisi a cinque proprietà ciascuno. Di queste, tre distintive salienti (rosse) e due distintive marginali (blu) come mostra lo schema seguente:



In questa rete ogni neurone è stato sostituito con un oscillatore di Wilson – Cowan precedentemente descritto, ottenendo così un modello che si basa complessivamente sulle equazioni (1), (2), (6) e (7'). Tale rete è stata successivamente addestrata attraverso la regola di Hebb, usando unità non oscillanti durante l'addestramento. La fase di addestramento, che consiste in 1000 epoche, è del tutto analoga a quella effettuata per la rete di base: i concetti vengono presentati una volta durante ogni epoca. Non tutte le proprietà di un singolo concetto hanno la stessa probabilità usata come input: le caratteristiche salienti hanno un'alta probabilità di verificarsi ($P_{kj(s)} = 0.7$), mentre a quelle marginali viene data una probabilità più bassa ($P_{kj(m)} = 0.4$). Durante l'addestramento per le proprietà marginali è stata scelta una probabilità più alta rispetto al modello di base (0.4 vs 0.35) e il valore basale della soglia post sinaptica è stato abbassato ad un valore pari a 0.55; visto il numero minore di proprietà totali (15 vs 231) è più difficile che una proprietà marginale divenga casualmente dominante, e quindi si può abbassare la soglia post sinaptica. Il valore degli altri parametri usati durante l'addestramento è rimasto invece immutato rispetto alla versione di base.

A fine addestramento è stata costruita la matrice W_E in corrispondenza dei collegamenti eccitatori ricalcando la matrice delle sinapsi W come descritto in seguito. Si è fissata una soglia pari a 0.7, si è confrontato ogni valore della matrice W con tale soglia. Se il generico elemento W_{ij} risultava maggiore di 0.7, si associava alla posizione corrispondente in W_E un valore pari a 2.7. Si noti che in questa rete si sono dovuti trovare nuovi valori da dare alle sinapsi eccitatorie contenute nella matrice W_E e si è dovuto inoltre abbassare il valore della deviazione standard del rumore (σ) a 0.1 e il valore di γ a 0.2. In particolare, abbassare γ equivale ad aumentare la costante di tempo, favorendo così la desincronizzazione degli oscillatori. Si sono inoltre settati sperimentalmente i valori adoperati per l'inibizione globale: è stato verificato che per valori di soglia (th) pari a 0.5 e di inibizione (k) pari a 0.4 il modello presenta un comportamento ottimale. Per completezza i valori numerici utilizzati nel modello sono riportati nella tabella sottostante, dove vengono evidenziati i parametri cambiati rispetto alla rete con un solo animale.

Significato	Simbolo	Valore
Parametri oscillatore Wilson-Cowan		
Forza sinaptica da unità eccitatoria a unità inibitoria	α	0.3
Forza sinaptica da unità inibitoria a unità eccitatoria	β	2.5
Parametro che influenza frequenza di oscillazione	γ	0.2
Pendenza delle sigmoidi	T	0.025
Offset di H_x	φ_x	0.7
Offset di H_y	φ_y	0.15
Deviazione standard del rumore	σ	0.1
Parametri costruzione sinapsi		
Pesi sinaptici eccitatori	W_E	2.7
Soglia	Soglia	0.7
Parametri inibitore globale		
Soglia	Th	0.5
Inibizione globale	k	0.4
Parametri modello		
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)	$\vartheta_{post-base}$	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 proprietà saliente	$P_{kj(s)}$	0.7
Probabilità proprietà marginale	$P_{kj(m)}$	0.4

Tabella 10. Parametri utilizzati per la rete semantica a 3 animale e 5 proprietà ciascuno (3 salienti e 2 marginali).

Il modello addestrato è stato successivamente testato. Il modello correttamente addestrato, in risposta all'input, dovrebbe essere in grado di evocare tutte le proprietà salienti dei concetti simultaneamente presentati. Le proprietà marginali non dovrebbero invece essere richiamate. Una volta richiamate, le proprietà salienti di un animale dovrebbero oscillare in modo sincrono tra loro e, vista la presenza dell'inibitore globale, dovrebbero essere desincronizzate rispetto a quelle degli altri due animali. Tutte le prove sono state effettuate a partire da uno stato iniziale pari ad un valore random preso tra 0 e 1 per ogni oscillatore. Durante le simulazioni sono stati eccitati simultaneamente degli oscillatori corrispondenti alle proprietà dei tre concetti e i risultati mostrano che il modello ha il comportamento atteso. Di seguito vengono riportati degli esempi per chiarire quanto appena detto. Vengono considerati i concetti "mucca", "pecora" e "cavallo". Nel *pattern* in *fig. 21* è raffigurato il comportamento della rete se in input vengono date simultaneamente tre proprietà salienti, una per ciascun animale (es. "muggisce" per il concetto di "mucca", "si tosa" per "pecora" e "si cavalca" per "cavallo"). Il modello riesce ad evocare tutte le proprietà salienti di "mucca" (linea blu continua), tutte le proprietà salienti di "pecora" (linea rossa continua) e tutte le proprietà salienti di "cavallo" (linea verde continua), a farle oscillare in fase e a desincronizzare i tre concetti.

Le proprietà marginali dei tre concetti (linee blu, rosse e verdi tratteggiate) non vengono invece richiamate.

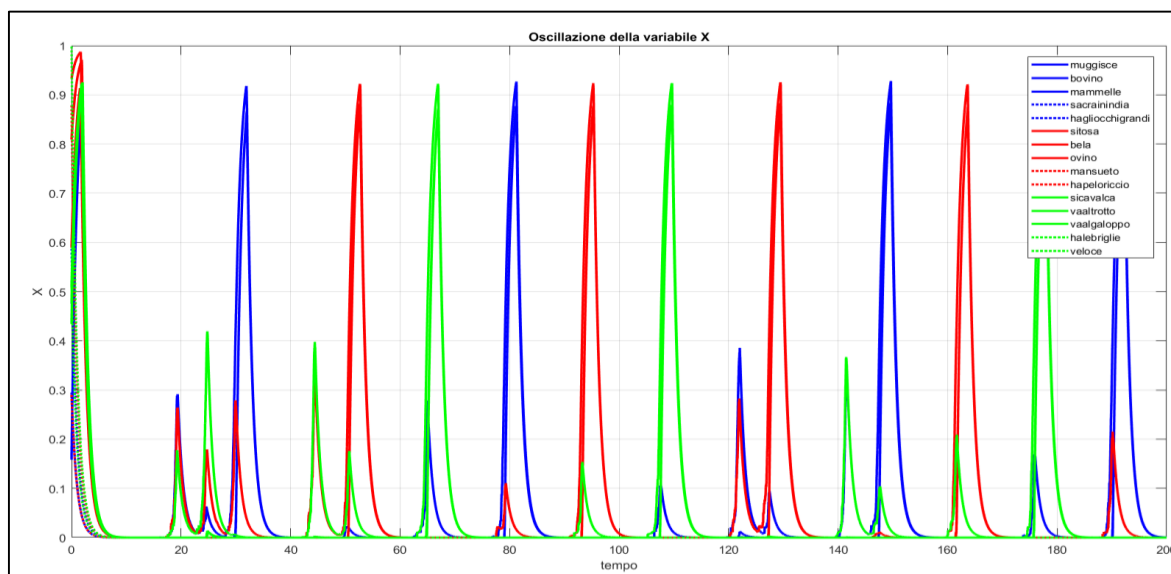


Figura 21 - Oscillazione dell'unità eccitatoria (X) delle proprietà salienti dei tre concetti desincronizzati. Un andamento simile si ottiene anche per l'unità inibitoria Y, qui omissa per brevità.

Un risultato interessante si è ottenuto se ad uno dei tre concetti vengono dati due input salienti, mentre gli altri due un solo input saliente. In particolare il concetto che riceve più input appare con una frequenza maggiore. A titolo esemplificativo si riporta il *pattern* in *fig. 22* nella pagina seguente dove il concetto di “mucca”, che riceve due input salienti (le proprietà “muggisce” e “bovino”), appare con una frequenza maggiore (7 picchi blu) rispetto agli altri due concetti (3 picchi rossi e 3 picchi verdi) che ricevono un solo input saliente (“si tosa” per il concetto di “pecora” e “si cavalca” per il concetto di “cavallo”).

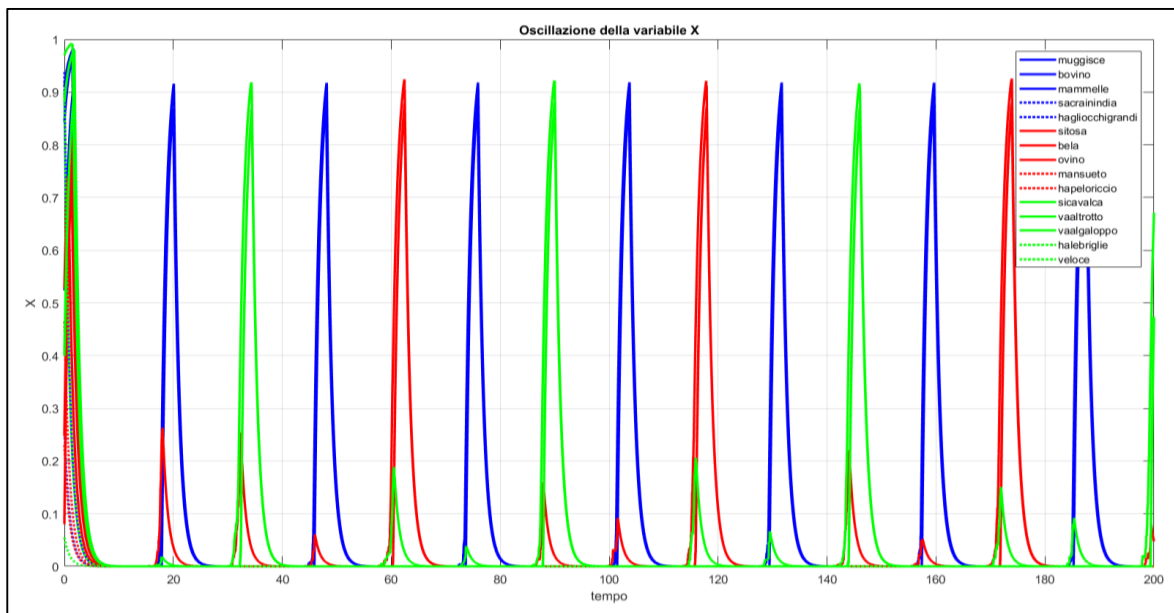


Figura 22 - Oscillazione dell'unità eccitatoria delle feature salienti di mucca (blu), pecora (rosso) e cavallo (verde). Si noti che l'animale che riceve più input salienti (mucca) viene richiamato dalla rete più frequentemente.

In *fig. 23* viene rappresentato cosa succede se viene dato come input una proprietà marginale a tutti e tre i concetti (es. “sacra in India” per il concetto di “mucca”, “mansueto” per il concetto di “pecora” e “ha le briglie” per il concetto di “cavallo”). In particolare la rete evoca le proprietà salienti di ogni concetto e le fa oscillare in fase, mentre le marginali sono presenti solo perché utilizzate come input della rete. Si noti inoltre che i picchi corrispondenti alle proprietà marginali risultano avere un'ampiezza minore rispetto a quelli relativi alle proprietà salienti.

Questo può trovare una corrispondenza fisiologica: le proprietà ritenute marginali per un concetto, se utilizzate come input, riescono ad evocare le proprietà salienti e quindi a fare riconoscere il concetto in questione, dopodiché vengono dimenticate quindi rimangono più basse. La rete riesce inoltre a desincronizzare i tre concetti presentati simultaneamente.

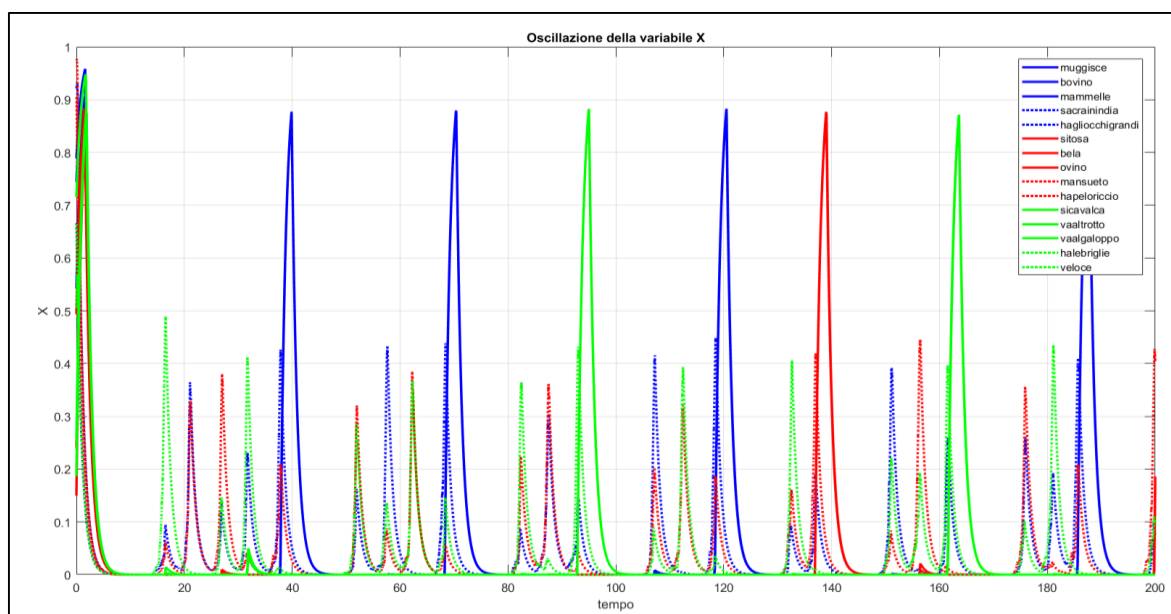


Figura 23 - Oscillazione unità eccitatoria delle proprietà salienti (linea continua) e marginali (linea tratteggiata) di mucca (blu), pecora (rosso) e cavallo (verde).

Infine la rete ha un buon comportamento sia nel caso in cui le venga dato come input una proprietà saliente per un concetto (es. “bovino” per il concetto di “mucca”) e una proprietà marginale per ciascuno degli altri due concetti (es. “mansueto” per il concetto di “pecora” e “ha le briglie” per il concetto di “cavallo”), oppure nel caso in cui le venga dato come input una proprietà marginale per un concetto e una proprietà saliente per ciascuno degli altri due concetti. Tutte le proprietà salienti di ogni concetto vengono richiamate, oscillano in fase e i tre concetti presentati contemporaneamente sono desincronizzati. Le proprietà marginali non vengono invece evocate, se non quando usate come input (in questo ultimo caso inoltre la loro ampiezza rimane più bassa rispetto a quella delle proprietà salienti). Un esempio è riportato in *fig. 24* nella pagina successiva, dove si nota inoltre che, nonostante i tre concetti risultino essere desincronizzati tra loro, le oscillazioni non sono regolari nel tempo e presentano un ordine diverso.

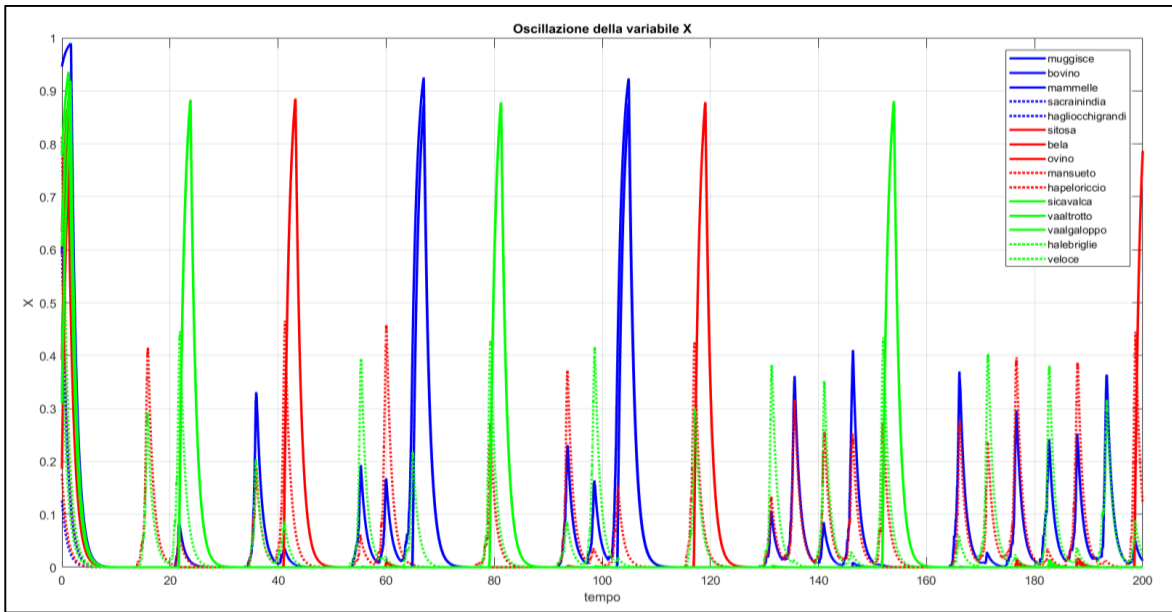
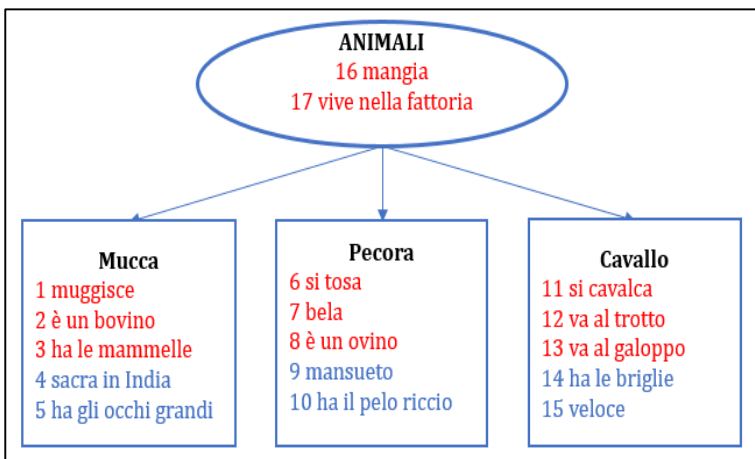


Figura 24 - Comportamento della rete nel caso in cui come input si dia: 1 proprietà saliente al concetto "mucca" (blu), 1 proprietà marginale al concetto "pecora" (rosso) e 1 proprietà marginale al concetto "cavallo" (verde). Si noti che il concetto con in input 1 proprietà saliente (mucca, picchi blu) presenta picchi leggermente più alti rispetto agli altri due concetti.

Fino a questo punto si sono prese in considerazione solo proprietà (salienti o marginali) distintive dei tre concetti. Quanto detto fino ad ora può essere esteso anche ad una rete in cui sono presenti delle proprietà totalmente condivise (che formano la categoria ANIMALE). Nelle simulazioni seguenti, ciascuno dei tre animali della rete non dispone solo di 5 proprietà, ma 7. Di queste, 2 totalmente condivise, 3 salienti distintive (rosse) e 2 marginali distintive (blu):



Si è verificato che in questo caso la rete, correttamente addestrata, riesce anche a formare la categoria ANIMALE. In particolare fornendo come input una proprietà di categoria (es. “mangia”) la rete evoca l’altra proprietà appartenente alla categoria (“vive nella fattoria”) e la fa oscillare in fase con quella data in input (*fig. 25*). Nessuna proprietà saliente di un concetto specifico viene invece evocata visto che il modello non ha sufficienti informazioni per discriminare un concetto dagli altri due. Se oltre alla proprietà di categoria viene fornita in input anche una proprietà distintiva (saliente o marginale) dei tre concetti, la rete riesce a riconoscere i tre concetti simultaneamente (i.e richiama in fase tutte le proprietà salienti per ciascun concetto oltre a quelle di categoria) e a desincronizzarli.

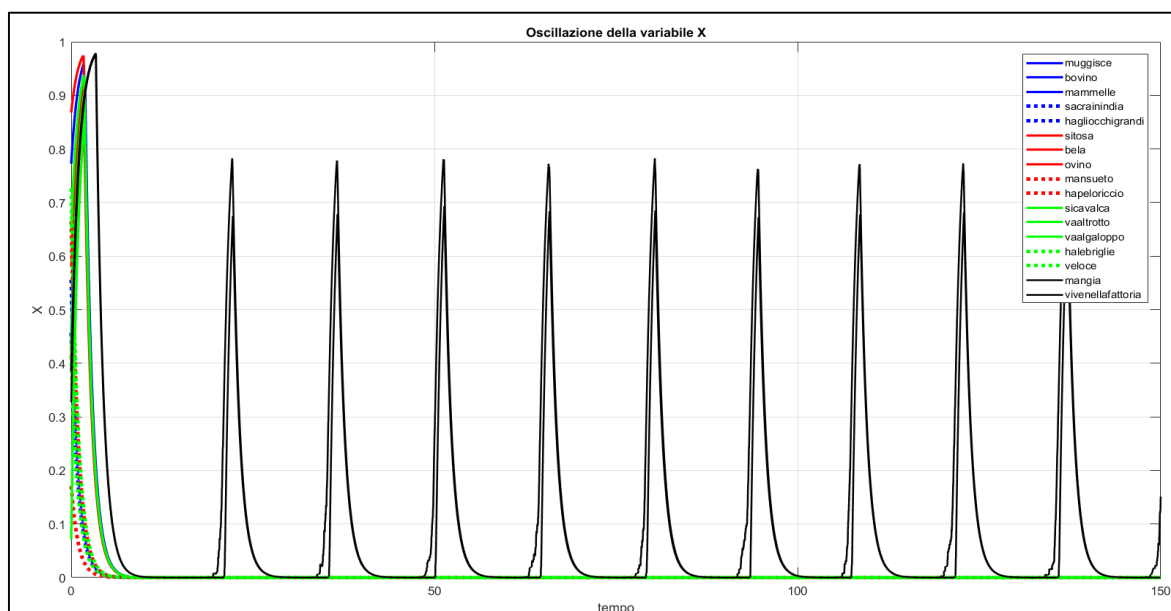


Figura 25 - Oscillazione delle unità eccitatorie delle proprietà appartenenti alla categoria ANIMALE.

Si fa notare che i parametri utilizzati nella rete a 3 animali e 7 proprietà per ciascun animale sono del tutto analoghi a quelli utilizzati nella rete a 3 animali e 5 proprietà ciascuno. L’unica differenza si ha nei parametri dell’inibitore globale: prove numeriche hanno evidenziato che per ottenere un buon comportamento della rete il parametro k deve essere ulteriormente abbassato e portato ad un valore pari a 0.3.

Si riporta di seguito solo la sezione della tabella 10 relativa ai parametri dell'inibitore globale, evidenziando il parametro variato rispetto al caso precedente:

Parametri inibitore globale		
Soglia	Th	0.5
Inibizione globale	k	0.3

Se, partendo dalle due reti con tre animali appena analizzate, si trascurano uno oppure due concetti (cioè si settano tutte le loro proprietà a 0), si ricade nei casi semplificati, rispettivamente, di due soli animali oppure di un solo animale. Come è lecito aspettarsi, simulazioni su queste reti hanno riportato risultati soddisfacenti e analoghi a quelli già ottenuti con la rete a tre animali: vengono richiamate le proprietà salienti di ciascun concetto, oscillano in fase e, nel caso di due animali, questi ultimi vengono correttamente desincronizzati. Le proprietà marginali invece non vengono evocate se non quando usate come input esterno (*fig. 26, 27, 28*). Nella rete con due soli animali si è inoltre riscontrata una differenza rispetto alla rete a tre animali: aumentando il numero di input salienti per uno dei due concetti, questo ultimo non viene evocato con più frequenza.

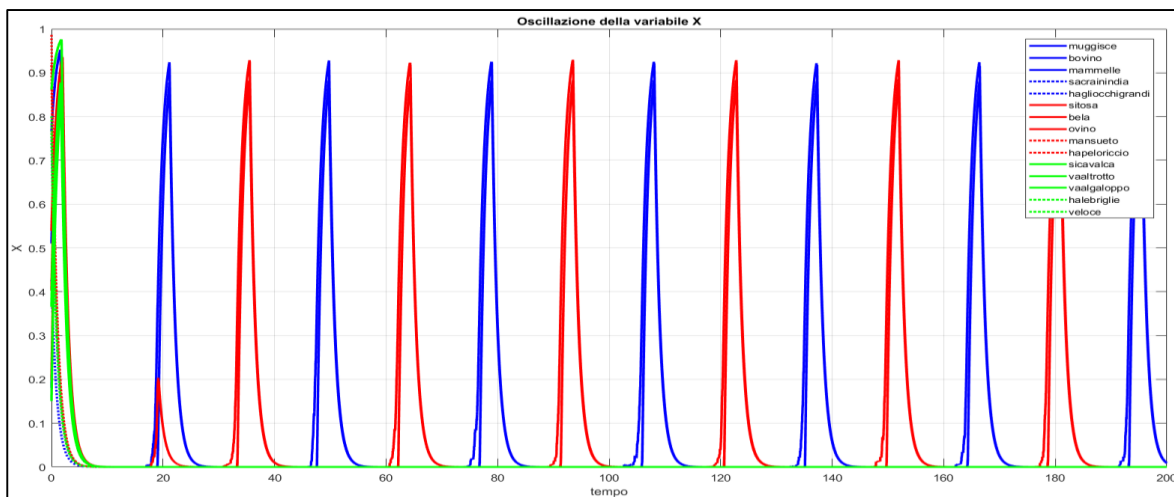


Figura 26 - Oscillazione unità eccitatoria delle proprietà salienti di mucca (blu) e pecora (rosso) quando in input viene data una proprietà saliente per ciascun animale.

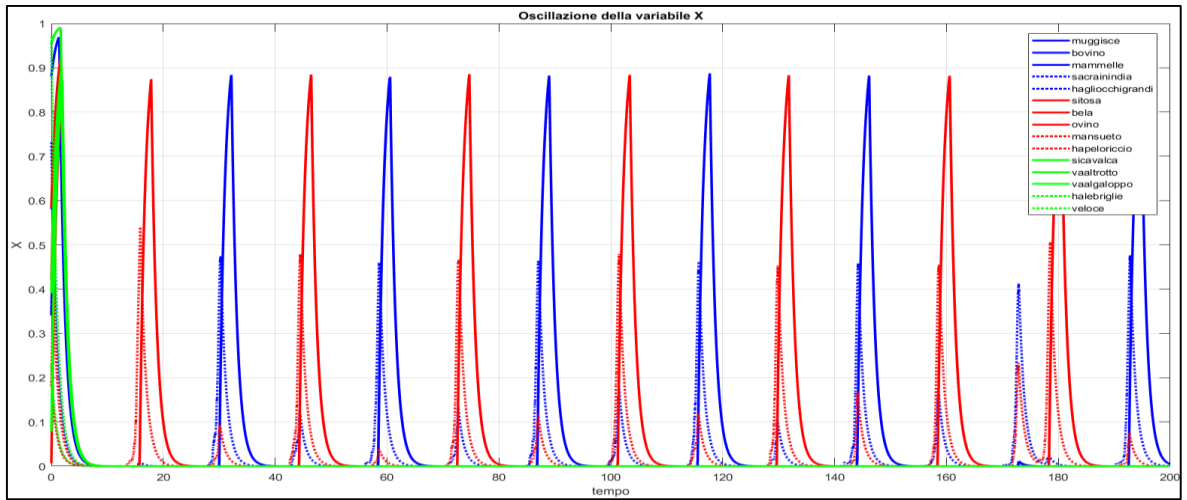


Figura 27 - Oscillazione unità eccitatoria delle proprietà salienti (linee continue) e marginali (linee tratteggiate) di mucca (blu) e pecora (rosso) quando in input viene data una proprietà marginale per ciascun animale.

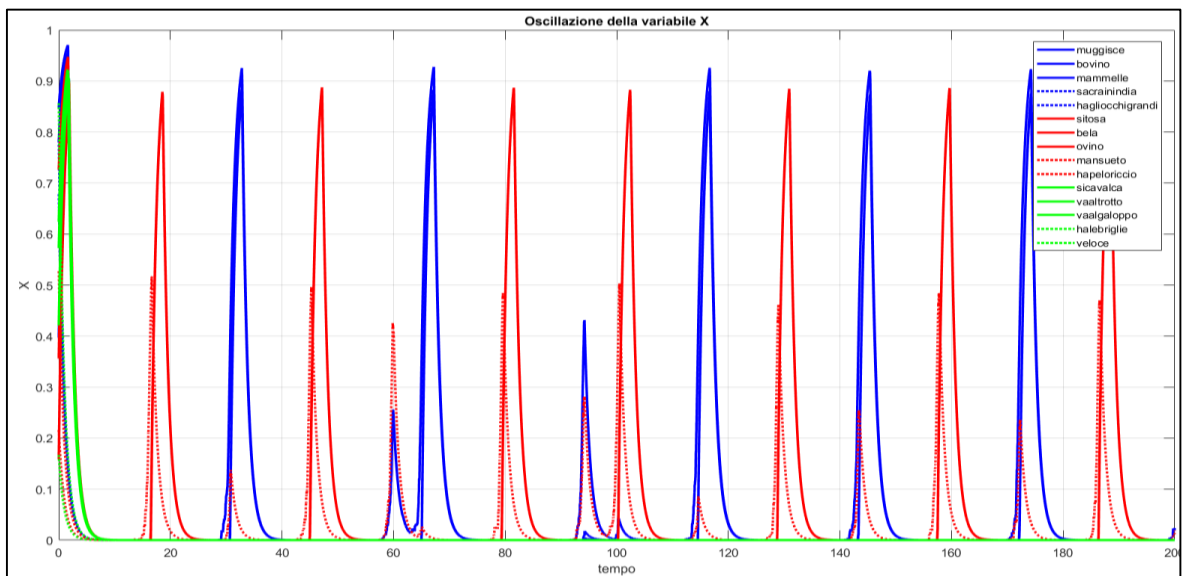


Figura 28 - Oscillazione unità eccitatoria delle feature salienti (linee continue) e marginali (linee tratteggiate) dei concetti di mucca (blu) e pecora (rosso) quando alla rete viene data una proprietà marginale come input ad un concetto e una proprietà saliente come input all'altro concetto.

3.4.2.2 Rete con 3 animali 10 proprietà

Partendo dalla rete semantica di base, si è implementata una rete contenente 3 animali con 10 proprietà ciascuno, al fine di verificare se il modello riuscisse a riconoscere più concetti simultaneamente anche nel caso in cui questi ultimi avessero un numero di caratteristiche abbastanza elevato.

Dapprima si è studiato il comportamento della rete nel caso in cui tutte e 10 le proprietà di ciascun concetto fossero salienti. Successivamente si è diminuito il numero di proprietà salienti (e aumentato quello di proprietà marginali) e si è analizzato il comportamento del modello in questi casi. Si fa notare che per le reti descritte di seguito sono state apportate alcune modifiche alla tassonomia degli animali di base. In particolare, dato che per i concetti di “pecora” e “cavallo” non erano presenti 10 proprietà salienti distintive, le *feature* “ha il pelo riccio” e “usato per fare formaggio” per il concetto di pecora, e la *feature* “ha le briglie” per il concetto di “cavallo” sono state etichettate come saliente seppure nella versione di base fossero marginali. Partendo da questa tassonomia, per analizzare i casi che prevedono anche proprietà marginali, si sono modificate le probabilità delle *feature* prima di ogni fase di addestramento portandole da un valore di 0.7 a un valore di 0.35.

➤ 10 proprietà salienti

La tassonomia utilizzata in questa rete si rifà allo schema seguente:



Nella rete ogni neurone è stato sostituito con un oscillatore di Wilson – Cowan precedentemente descritto, ottenendo così un modello che si basa complessivamente sulle equazioni (1), (2), (6) e (7'). Tale rete è stata successivamente addestrata in modo del tutto analogo a quella di base, pertanto i parametri adoperati, con i rispettivi valori, sono quelli presentati nel *cap.2*.

L'unica differenza introdotta in questa rete riguarda la soglia post sinaptica di base: il suo valore è stato abbassato a 0.55 visto il numero minore di proprietà totali (30 vs 231) e quindi la minore possibilità che proprietà marginali divengano casualmente salienti. A fine addestramento è stata costruita la matrice W_E impostando sinapsi eccitatorie pari a 0.6. Si sono inoltre settati dei nuovi valori adoperati per l'inibizione globale: è stato verificato sperimentalmente che per valori di soglia (th) pari a 0.4 e di inibizione (k) pari a 0.23 il modello presenta un comportamento ottimale. Per quanto riguarda i parametri adoperati per gli oscillatori, la deviazione standard del rumore (σ) e il parametro γ sono stati portati a valori più alti, rispettivamente 0,9 e 0,5. Gli altri parametri degli oscillatori sono invece rimasti invariati. Per completezza tutti i valori sono riportati nella tabella seguente, dove vengono evidenziati i parametri che sono stati subito modificati.

Significato	Simbolo	Valore
Parametri oscillatore Wilson-Cowan		
Forza sinaptica da unità eccitatoria a unità inibitoria	α	0.3
Forza sinaptica da unità inibitoria a unità eccitatoria	β	2.5
Parametro che influenza frequenza di oscillazione	γ	0.5
Pendenza delle sigmoidi	T	0.025
Offset di H_x	φ_x	0.7
Offset di H_y	φ_y	0.15
Deviazione standard del rumore	σ	0.9
Parametri sinapsi		
Pesi sinaptici eccitatori	W_E	0.6
Soglia	Soglia	0.7
Parametri inibitore globale		
Soglia	Th	0.4
Inibizione globale	k	0.23
Parametri modello		
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)	$\vartheta_{post-base}$	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 proprietà saliente	$P_{kj(s)}$	0.7

Tabella 11. Parametri utilizzati nella rete a 3 animali e 10 proprietà salienti ciascuno.

A fine addestramento la rete è stata testata. Tutte le prove sono state effettuate a partire da uno stato iniziale pari ad un valore random preso tra 0 e 1 per ogni oscillatore. Dando in input simultaneamente una proprietà saliente per ciascun concetto, il modello riesce a riconoscere i tre concetti, evocando tutte le altre proprietà salienti per ciascuno, e a desincronizzarli come si evince dal *pattern* in *fig. 29*. Aumentando il numero di input per ciascun concetto (da 2 a 9) si è notato invece che la rete non riesce più a desincronizzare i tre animali.

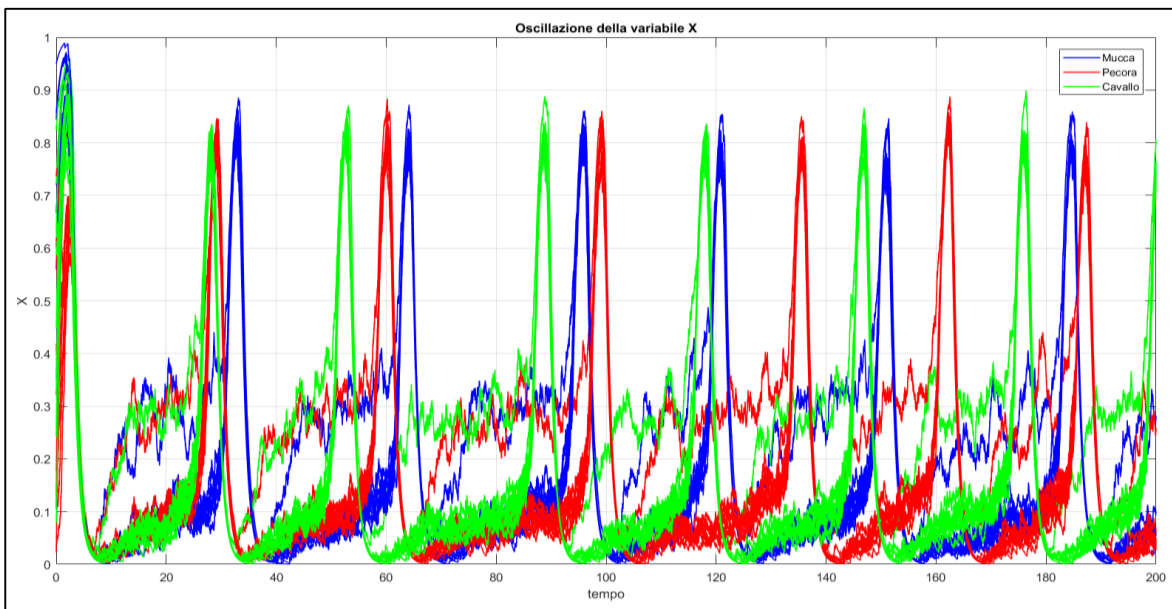


Figura 29 - Oscillazione unità eccitatoria delle proprietà salienti di mucca (blu), pecora (rosso) e cavallo (verde) quando come input viene data una proprietà saliente per ciascun animale.

➤ **9 proprietà salienti e 1 marginale / 8 proprietà salienti e 2 marginali**

Le tassonomie utilizzate in questa rete si rifanno agli schemi seguenti:



Nella rete ogni neurone è stato sostituito con un oscillatore di Wilson – Cowan precedentemente descritto, ottenendo così un modello che si basa complessivamente sulle equazioni (1), (2), (6) e (7'). Tale rete è stata successivamente addestrata tramite la regola di Hebb. La fase di addestramento, che consiste in 1000 epoche, è del tutto analoga a quella effettuata per la rete di base: i concetti vengono presentati una volta durante ogni epoca. In questo caso non tutte le *feature* hanno la stessa probabilità di verificarsi: le caratteristiche etichettate come salienti hanno una elevata probabilità di verificarsi ($P_{kj(s)} = 0.7$), mentre le caratteristiche marginali hanno una probabilità più bassa ($P_{kj(m)} = 0.7$). Anche in questo caso l'unico valore che, durante l'addestramento, ha subito delle variazioni rispetto alla rete di base è quello della soglia post sinaptica di base (0.55 vs 0.67), mentre gli altri parametri adoperati sono rimasti invariati. A fine addestramento è stata costruita la matrice W_E in corrispondenza dei collegamenti eccitatori ricalcando la matrice delle sinapsi W come descritto in seguito. Si è fissata una soglia pari a 0.7, in tutte le posizioni della matrice W_E corrispondenti a valori sinaptici in W con valori maggiore rispetto alla soglia, sono stati impostate sinapsi pari a 0.65, nel caso di 9 proprietà salienti, e pari a 0.7 nel caso di 8 proprietà salienti.

Questi valori sono stati settati dopo prove numeriche a seguito delle quali si è notato che aumentando di un fattore pari a 0.05 le sinapsi eccitatorie ogni qualvolta una proprietà saliente venisse posta come marginale (cioè passando da 10 a 9 a 8 proprietà salienti), il modello presentava un buon comportamento. Quanto appena detto si è verificato però solo fino ad un numero di proprietà salienti pari a 8. Si è notato infatti che diminuendo ulteriormente il numero di proprietà salienti, la scelta del valore numerico da dare alle sinapsi della matrice WE non rispetta più questo criterio e quindi occorre trovare nuovi valori corretti. Per completezza tutti i valori sono riportati nella tabella seguente, dove vengono evidenziati i valori che subiscono una variazione.

Significato	Simbolo	Valore
Parametri oscillatore Wilson-Cowan		
Forza sinaptica da unità eccitatoria a unità inibitoria	α	0.3
Forza sinaptica da unità inibitoria a unità eccitatoria	β	2.5
Parametro che influenza frequenza di oscillazione	γ	0.5
Pendenza delle sigmoidi	T	0.025
Offset di H_x	φ_x	0.7
Offset di H_y	φ_y	0.15
Deviazione standard del rumore	σ	0.9
Parametri costruzione sinapsi		
Pesi sinaptici eccitatori	W_E	0.65/0.7
Soglia	Soglia	0.7
Parametri inibitore globale		
Soglia	Th	0.4
Inibizione globale	k	0.23
Parametri modello		
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)	$\vartheta_{post-base}$	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 proprietà saliente	$P_{kij(s)}$	0.7
Probabilità proprietà marginale	$P_{kij(m)}$	0.35

Tabella 12. Parametri utilizzati nella rete a 3 animali e sia 9 sia 8 proprietà salienti per ciascuno.

La rete addestrata è stata successivamente testata. Per brevità si riportano solo i risultati relativi alla rete con 9 proprietà salienti e 1 marginale, ma conclusioni del tutto analoghe sono state riscontrate anche testando la rete con 8 proprietà salienti e 2 marginali. Tutte le prove sono state effettuate a partire da uno stato iniziale pari ad un valore random preso tra 0 e 1 per ogni oscillatore. Se in input vengono date simultaneamente tre proprietà salienti (1 per ogni animale), la rete riesce ad evocare correttamente tutte le altre proprietà salienti di ogni concetto e a farle oscillare in fase, mentre le proprietà marginali non vengono richiamate. Inoltre, i vari concetti riescono ad essere desincronizzati anche se, come risulta dalla *fig. 30*, la desincronizzazione non risulta perfetta.

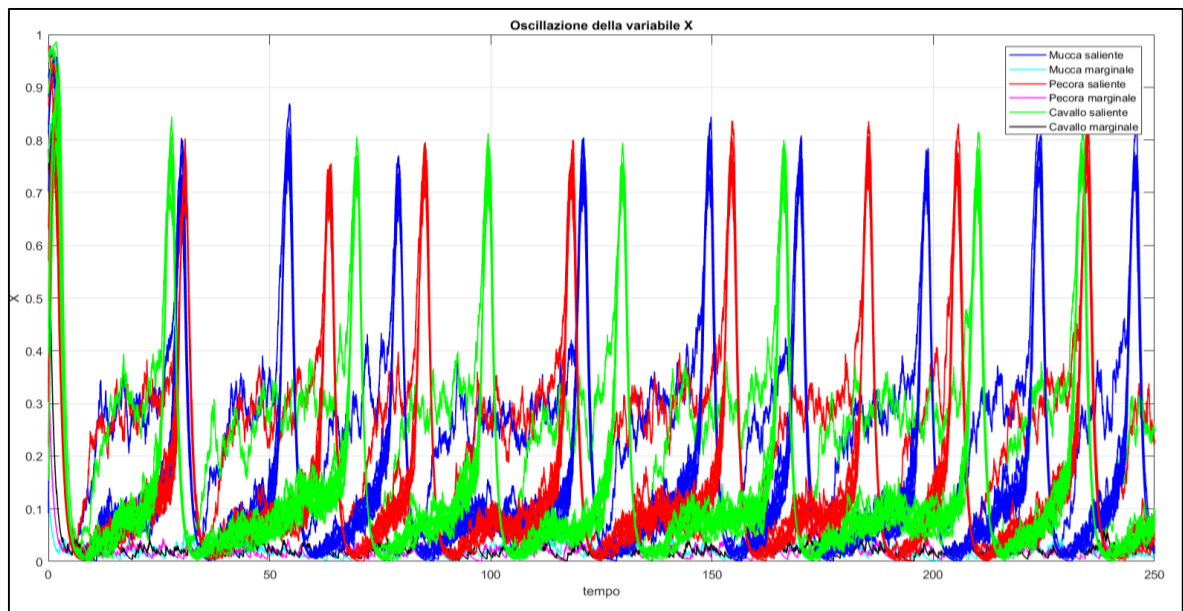


Figura 30 - Oscillazione unità eccitatoria delle proprietà salienti di mucca (blu), pecora (rosso) e cavallo (verde) quando in input viene data una proprietà saliente per ciascun animale.

Un risultato interessante si è ottenuto se ad uno dei tre concetti vengono dati due input salienti, mentre gli altri due un solo input saliente.

In particolare il concetto che riceve più input appare con una frequenza maggiore. A titolo esemplificativo si riporta il *pattern* in *fig. 31* dove il concetto di “mucca”, che riceve due input salienti (le proprietà “muggisce” e “bovino”), appare con una frequenza maggiore rispetto agli altri due concetti che ricevono un solo input saliente (“si tosa” per il concetto di “pecora” e “si cavalca” per il concetto di “cavallo”).

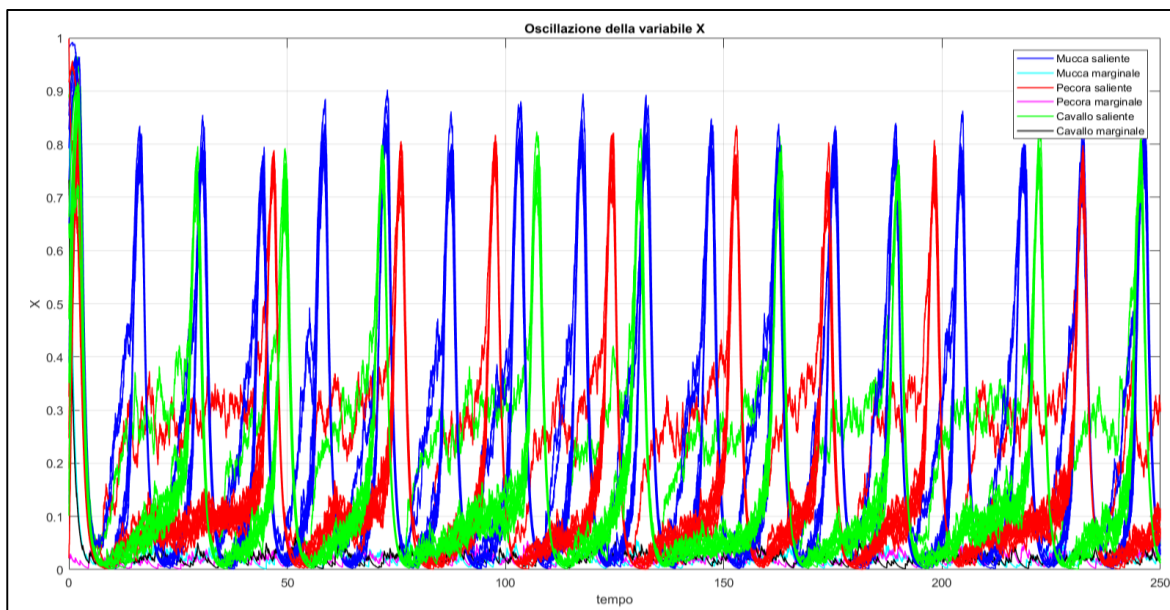


Figura 31 - Oscillazione unità eccitatoria proprietà salienti di mucca (blu), pecora (rosso) e cavallo (verde). Si noti che il concetto che riceve più proprietà salienti (mucca) appare con più frequenza.

Dando in input tre proprietà marginali simultaneamente (1 per ogni concetto), il modello riconosce correttamente i concetti e li desincronizza (*fig. 32*). In questo caso, le proprietà marginali non sono completamente nulle, poiché usate come input esterno. Si noti inoltre che le oscillazioni delle unità relative alle proprietà marginali rimangono più basse rispetto alle oscillazioni delle unità relative alle proprietà salienti. Infine se alla rete vengono dati contemporaneamente in input una proprietà saliente per un animale e una proprietà marginale per gli altri due animali, il concetto che riceve in input la proprietà saliente compare più frequentemente (*fig. 33*).

Quanto detto si è riscontrato anche nel caso in cui a due animali venga dato in input una proprietà saliente ciascuno e al terzo una caratteristica marginale: gli animali con input salienti vengono richiamati dalla rete con una frequenza più elevata (fig. 34).

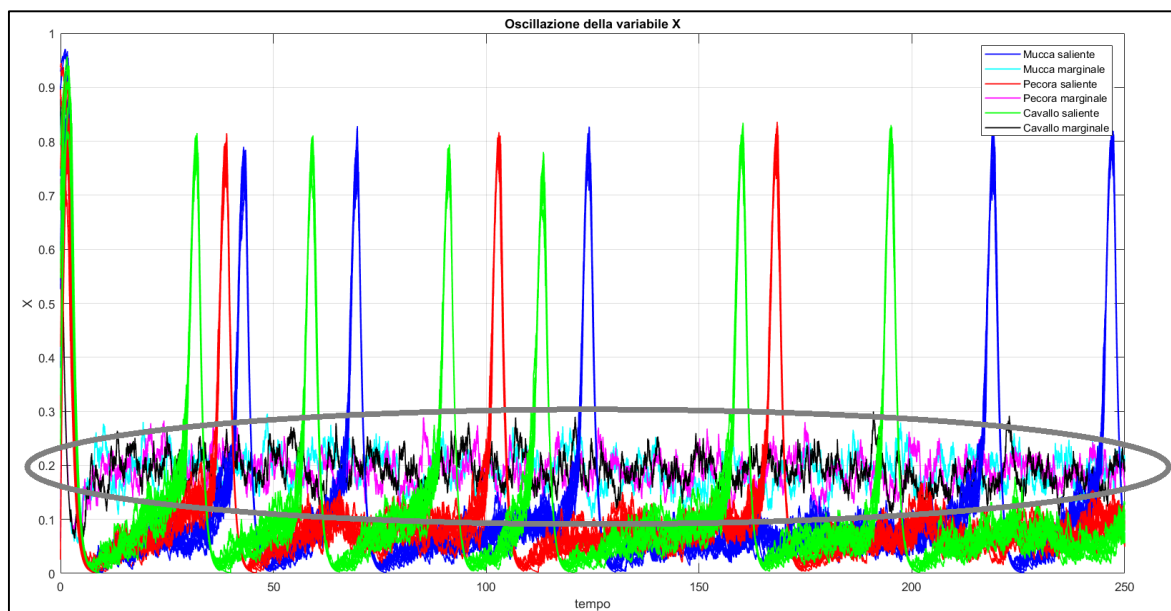


Figura 32 - Oscillazione della unità eccitatoria delle proprietà salienti di mucca (blu), pecora (rossa) e cavallo (verde). Nella bada nera sono presenti le proprietà marginali di ciascun concetto.

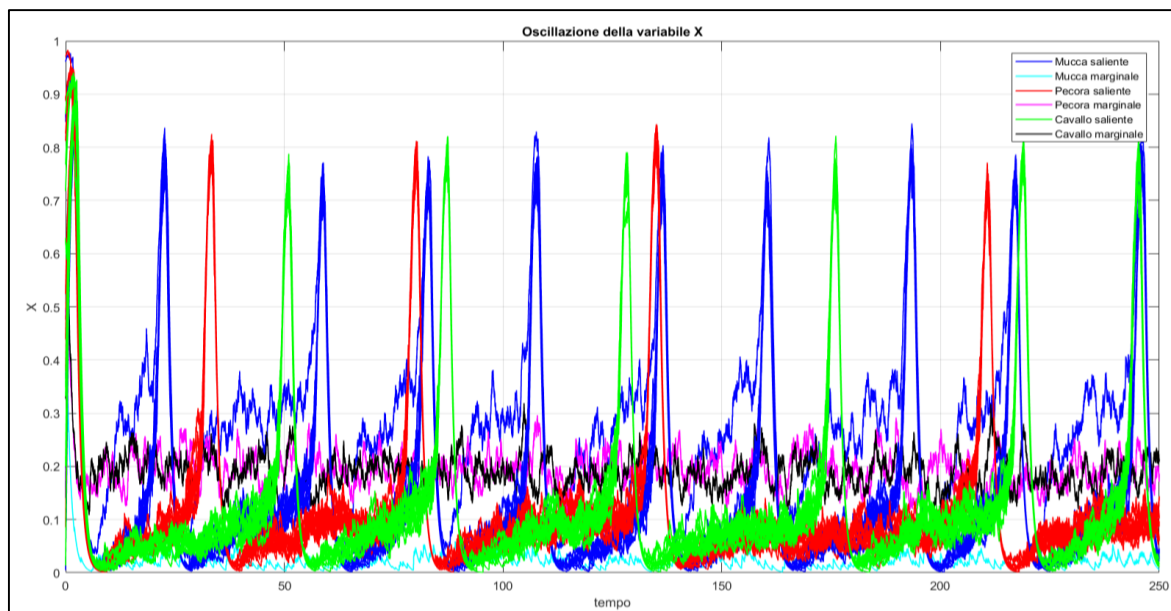


Figura 33 - Oscillazione unità eccitatoria delle proprietà salienti di mucca (blu), pecora (rosso) e cavallo (verde). L'animale che riceve in input una proprietà saliente (mucca) viene richiamato dalla rete con una frequenza più elevata.

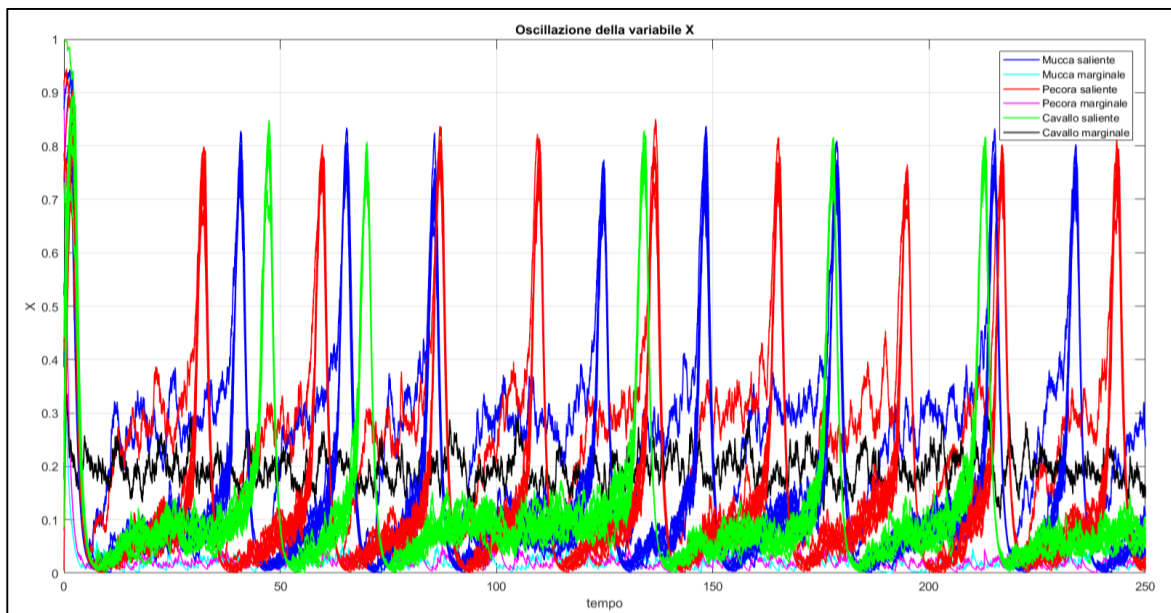


Figura 34 - Oscillazione unità eccitatoria delle proprietà salienti di mucca (blu), pecora (rosso) e cavallo (verde). Gli animali che ricevono in input una proprietà saliente (mucca e pecora) vengono richiamate dalla rete con una frequenza più elevata.

3.4.2.3 Rete con 3 oggetti 10 proprietà

Fino a questo punto della trattazione sono state eseguite simulazioni su reti provenienti dalla tassonomia degli animali. Come già messo in evidenza, risultati analoghi sono stati ottenuti anche testando il modello degli oscillatori sulla tassonomia di oggetti, contenente meno concetti rispetto alla precedente (12 vs 16). Per completezza di seguito vengono riportati i risultati del modello con oscillatori testato nel caso più “complesso” di rete con 3 oggetti e dieci proprietà ciascuno. Dapprima si è studiato il comportamento della rete nel caso in cui tutte e 10 le proprietà di ciascun oggetto fossero salienti. Successivamente si è diminuito il numero di proprietà salienti (e aumentato quello di proprietà marginali) e si è analizzato il comportamento del modello in questi casi. La rete considerata è stata estrapolata dalla tassonomia di base, dalla quale si sono scelti 3 oggetti (“tavolo”, “divano” e “libreria”). Si fa notare che anche in questo caso la tassonomia di base non disponeva di oggetti con 10 proprietà salienti pertanto sono state apportate alcune modifiche.

In particolare, per ciascun oggetto considerato, alcune proprietà che nella tassonomia di base risultavano marginali sono state considerate salienti. Per il concetto di tavolo le *feature* “duro”, “rettangolare”, “circondato da sedie”, “quadrato” e “regolabile”; per divano le proprietà “usato per sdraiarsi”, “ha i braccioli” e “imbottito”; per libreria le proprietà “alto”, “grande”, “componibile”, “si trova in biblioteca”, “si trova in ufficio” e “contiene soprammobili”. Partendo da questa tassonomia, per analizzare i casi che prevedono anche proprietà marginali, si sono modificate le probabilità delle *feature* prima di ogni fase di addestramento portandole da un valore di 0.7 a un valore di 0.35.

➤ **10 proprietà salienti**

La tassonomia utilizzata nel caso di 10 proprietà salienti si basa sullo schema seguente:



Nella rete ogni neurone è stato sostituito con un oscillatore di Wilson – Cowan precedentemente descritto, ottenendo così un modello che si basa complessivamente sulle equazioni (1), (2), (6) e (7). Tale rete è stata successivamente addestrata in modo del tutto analogo alla rete presentata nel *cap. 2*, pertanto i parametri adoperati, con i rispettivi valori, sono rimasti invariati. L’unica differenza introdotta in questa rete, rispetto a quella di base, riguarda la soglia post sinaptica di base: il suo valore è stato abbassato a 0.55 visto il numero minore di proprietà totali (30 vs 231) e quindi la minore possibilità che proprietà marginali divengano casualmente salienti.

A fine addestramento è stata costruita la matrice W_E impostando sinapsi eccitatorie pari a 0.6. Tutti i valori adoperati per gli oscillatori e per l'inibitore globale sono quelli adoprati anche per la rete basata sulla tassonomia degli animali, pertanto si faccia riferimento alla *tabella 11* nel *paragrafo 3.4.2.2*. A fine addestramento la rete è stata testata. Tutte le prove sono state effettuate a partire da uno stato iniziale pari ad un valore random preso tra 0 e 1 per ogni oscillatore. Dando in input simultaneamente una proprietà saliente per ciascun oggetto, il modello riesce a riconoscere i tre concetti, evocando tutte le altre proprietà salienti per ciascuno, e a desincronizzarli come si evince dalla figura 35. Anche in questo caso si è notato che aumentando il numero di input per ciascun concetto (da 2 a 9) la rete non riesce più a desincronizzare i tre oggetti.

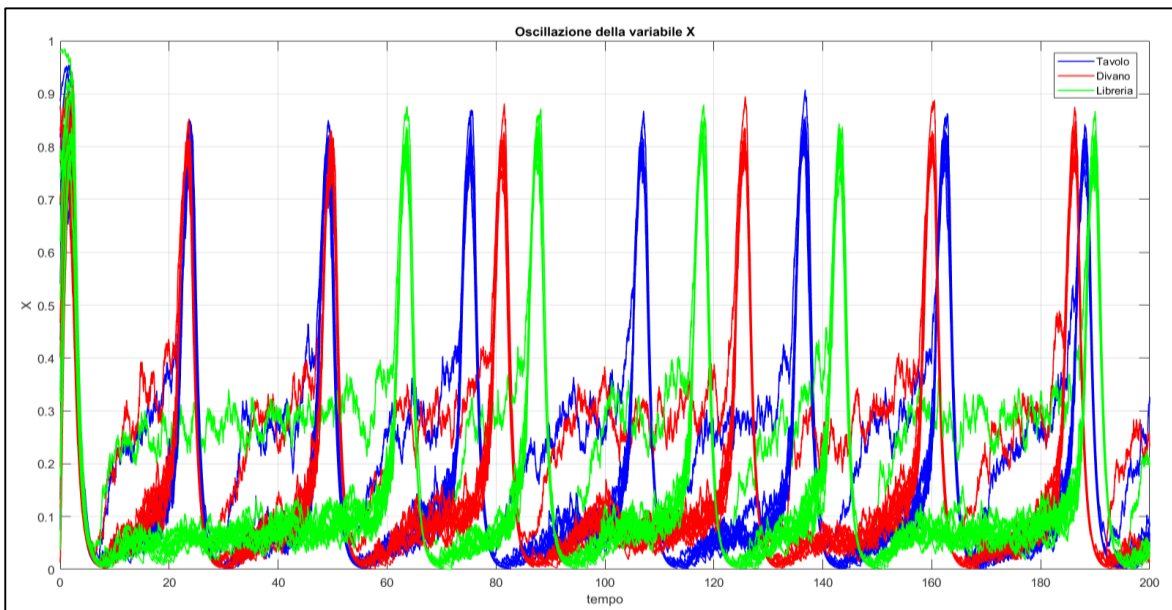
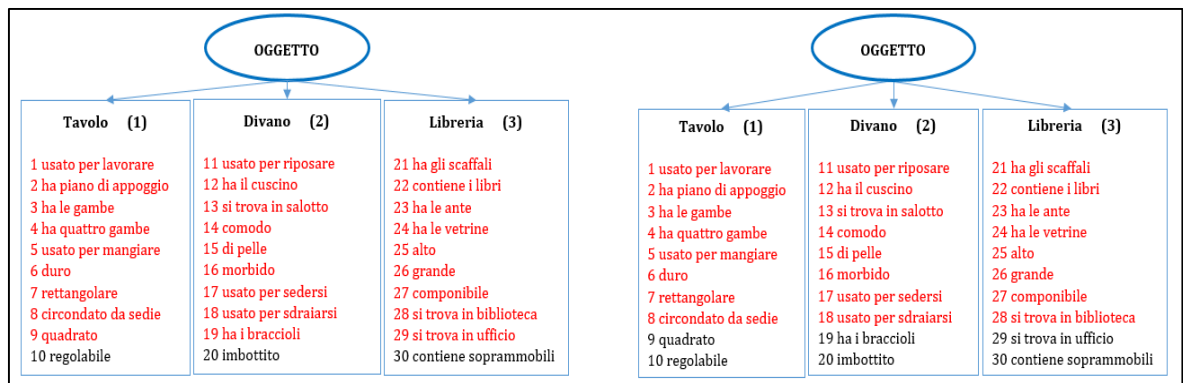


Figura 35 - Oscillazione unità eccitatorie delle proprietà salienti di tavolo (blu), divano (rosso) e libreria (verde) quando in input viene data una proprietà saliente per ciascun oggetto.

➤ **9 proprietà salienti e 1 marginale / 8 proprietà salienti e 2 marginali**

Le tassonomie utilizzate, rispettivamente, nel caso di 9 proprietà salienti e 8 proprietà salienti si basano sugli schemi seguenti:



Nella rete ogni neurone è stato sostituito con un oscillatore di Wilson – Cowan precedentemente descritto, ottenendo così un modello che si basa complessivamente sulle equazioni (1), (2), (6) e (7'). Tale rete è stata successivamente addestrata tramite la regola di Hebb. Durante la fase di addestramento, che consiste in 1000 epoche, i concetti vengono presentati una volta durante ogni epoca. In questo caso non tutte le *feature* hanno la stessa probabilità di verificarsi: le caratteristiche etichettate come salienti hanno una elevata probabilità di verificarsi ($P_{kj(s)} = 0.7$), mentre le caratteristiche marginali hanno una probabilità più bassa ($P_{kj(m)} = 0.35$). Anche in questo caso l'unico valore che, durante l'addestramento, ha subito delle variazioni rispetto alla rete del *cap. 2* è quello della soglia post sinaptica di base (0.55 vs 0.67), mentre gli altri parametri adoperati sono rimasti immutati. A fine addestramento è stata costruita la matrice W_E in corrispondenza dei collegamenti eccitatori ricalcando la matrice delle sinapsi W come descritto in seguito. Si è fissata una soglia pari a 0.7, in tutte le posizioni della matrice W_E corrispondenti a valori sinaptici in W con valori maggiore rispetto alla soglia, sono state impostate sinapsi pari a 0.65, nel caso di 9 proprietà salienti, e pari a 0.7 nel caso di 8 proprietà salienti. Questi valori sono stati settati dopo prove numeriche a seguito delle quali si è notato che aumentando di un fattore pari a 0.05 le sinapsi eccitatorie ogni qualvolta una proprietà saliente veniva posta come marginale (cioè passando da 10 a 9 a 8 proprietà salienti), il modello presentava un buon comportamento. Quanto appena detto si è verificato però solo fino ad un numero di proprietà salienti pari a 8.

Si è notato infatti che diminuendo ulteriormente il numero di proprietà salienti, la scelta del valore numerico da dare alle sinapsi della matrice WE non rispetta più questo criterio e quindi occorre trovare nuovi valori che permettano di ottenere il comportamento corretto. Per tutti i valori dei parametri utilizzati in questo caso si faccia riferimento alla *tabella 12* nel *paragrafo 3.4.2.2*. La rete addestrata è stata successivamente testata. Per brevità in questo caso si riportano solo i risultati relativi alla rete con 8 proprietà salienti e 2 marginale, ma conclusioni del tutto analoghe sono state riscontrate anche testando la rete con 9 proprietà salienti e 1 marginali. Tutte le prove sono state effettuate a partire da uno stato iniziale pari ad un valore random preso tra 0 e 1 per ogni oscillatore. Se in input vengono date simultaneamente tre proprietà salienti (1 per ogni oggetto), la rete riesce ad evocare correttamente tutte le altre proprietà salienti di ogni concetto e a farle oscillare in fase, mentre le proprietà marginali non vengono richiamate. Inoltre, i vari concetti riescono ad essere desincronizzati anche se, come si evince dalla *fig. 36*, la desincronizzazione non risulta perfetta. Un risultato interessante si è ottenuto se ad uno dei tre concetti vengono dati due input salienti, mentre gli altri due un solo input saliente. In particolare il concetto che riceve più input appare con una frequenza maggiore. A titolo esemplificativo si riporta il *pattern* in *fig. 37* nella pagina successiva.

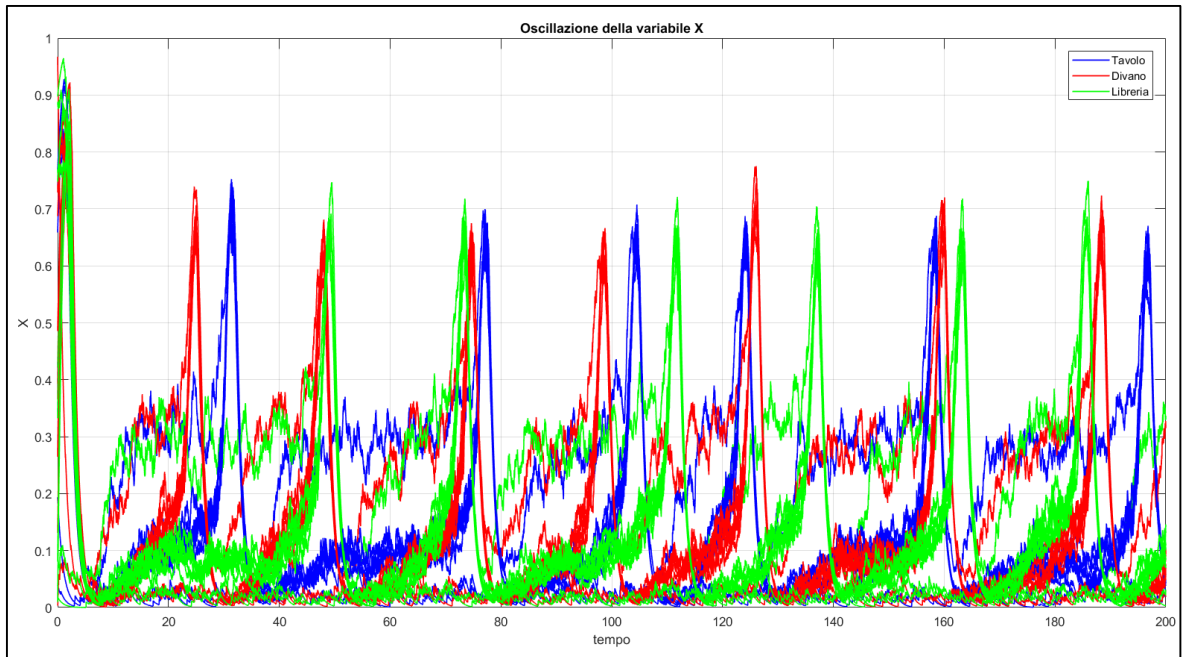


Figura 36 - Oscillazione unità eccitatoria delle proprietà salienti di tavolo (blu), divano (rosso) e libreria (verde) quando in input alla rete vengono date simultaneamente tre proprietà salienti (1 per ogni oggetto).

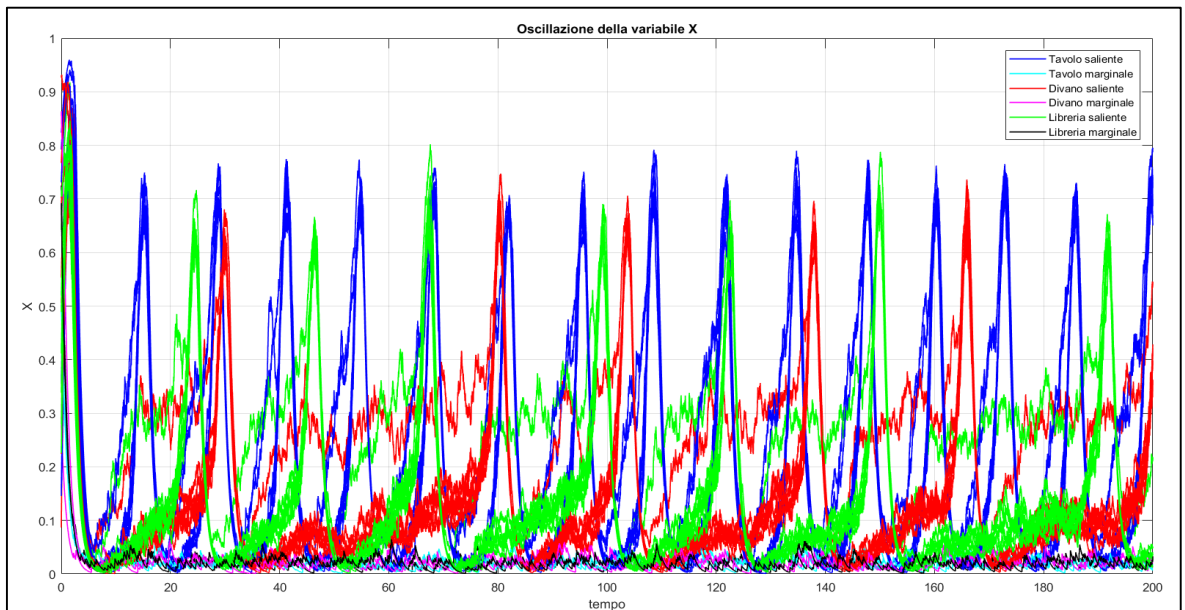


Figura 37 - Oscillazione unità eccitatoria proprietà salienti di tavolo (blu), divano (rosso) e libreria (verde). Si noti che il concetto che riceve più proprietà salienti (tavolo) appare con più frequenza.

Dando in input tre proprietà marginali simultaneamente (1 per ogni concetto), il modello riconosce correttamente i concetti e li desincronizza (fig. 38).

In questo caso, le proprietà marginali non sono completamente nulle, poiché usate come input esterno. Si noti inoltre che le oscillazioni delle unità relative alle proprietà marginali rimangono più basse rispetto alle oscillazioni delle unità relative alle proprietà salienti. Un risultato interessante si è ottenuto se ad uno dei tre concetti vengono dati due input marginali, mentre gli altri due un solo input marginale. In particolare il concetto che riceve più input appare con una frequenza maggiore. A titolo esemplificativo si riporta il *pattern* in *fig. 39*. Infine se alla rete vengono dati contemporaneamente in input una proprietà saliente per un oggetto e una proprietà marginale per gli altri due oggetti, il concetto che riceve in input la proprietà saliente compare più frequentemente (*fig. 40*). Quanto appena detto si è riscontrato anche nel caso in cui a due animali venga dato in input una proprietà saliente ciascuno e al terzo una caratteristica marginale: gli oggetti con input salienti vengono richiamati dalla rete con una frequenza più elevata.

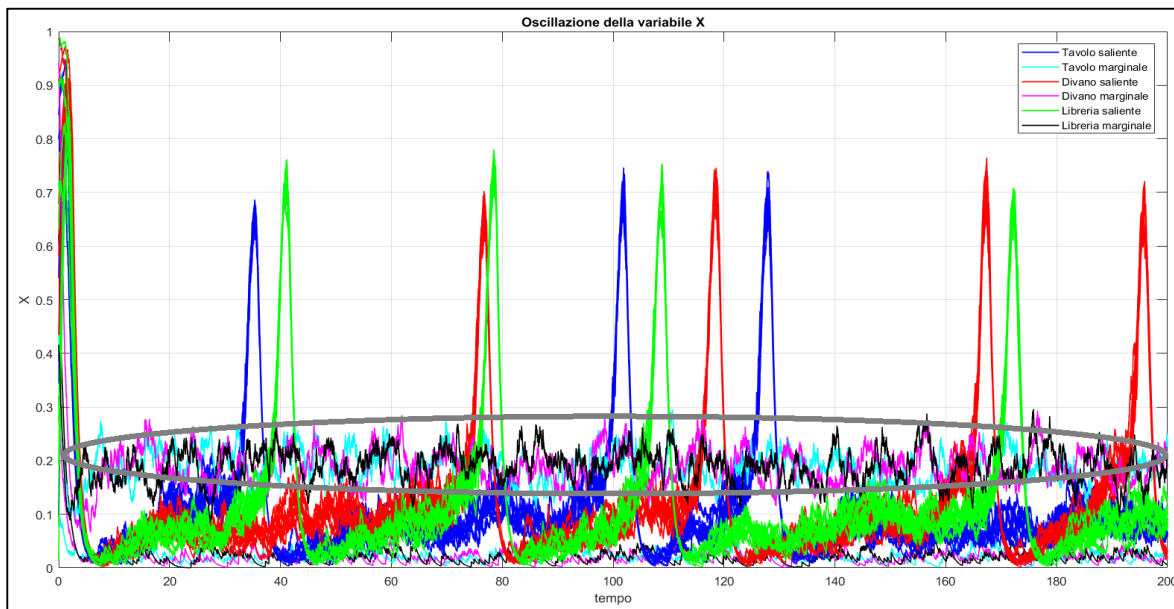


Figura 38 - Oscillazione della unità eccitatoria delle proprietà salienti di tavolo (blu), divano (rossa) e libreria (verde). Quelle cerchiato sono le proprietà marginali di ciascun concetto.

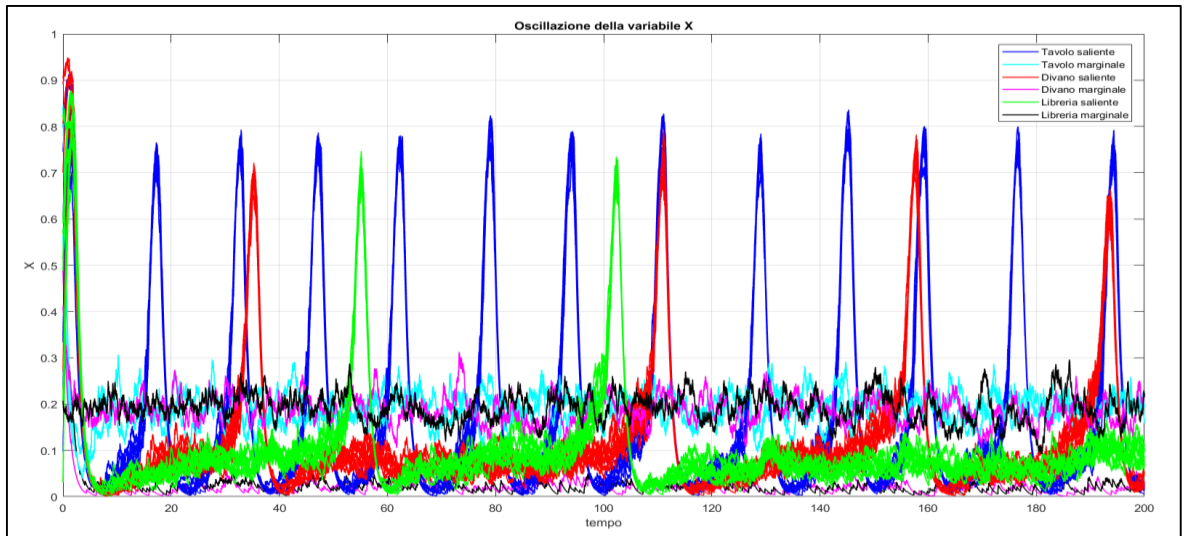


Figura 39 - Oscillazione unità eccitatoria proprietà salienti di tavolo (blu), divano (rosso) e libreria (verde). Si noti che il concetto che riceve più proprietà marginali (tavolo) appare con più frequenza.

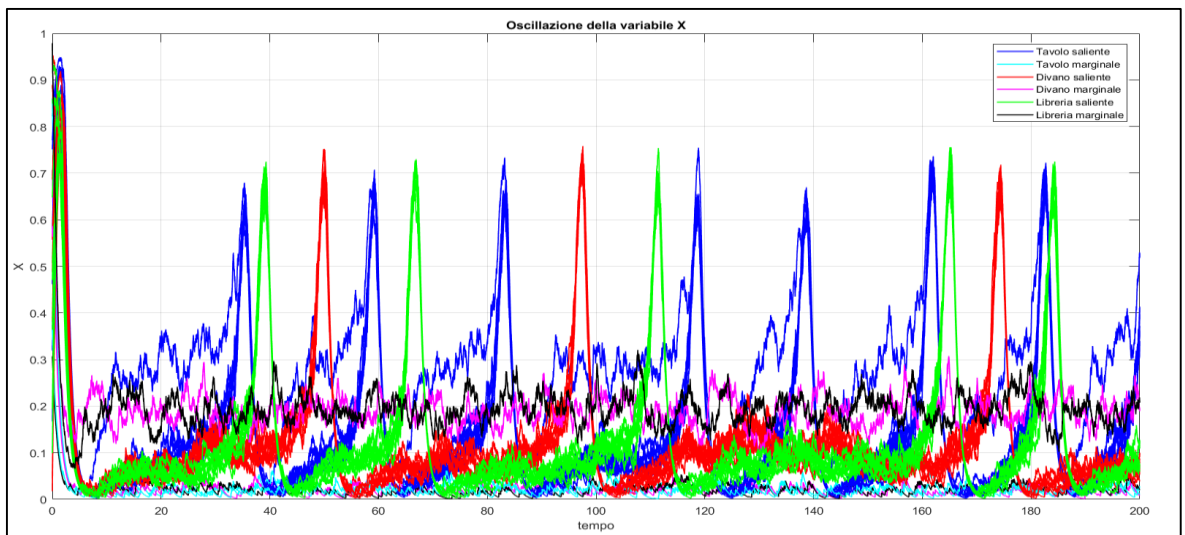


Figura 40 - Oscillazione unità eccitatoria delle proprietà salienti di tavolo (blu), divano (rosso) e libreria (verde). L'oggetto che riceve in input una proprietà saliente (tavolo) viene richiamato dalla rete con una frequenza più elevata.

Le prestazioni complessive del modello con oscillatori possono essere riassunte nel modo seguente:

Riconoscimento dei singoli concetti e sincronismo delle proprietà: se vengono stimulate le caratteristiche distintive (salienti o marginali), il modello riconosce il concetto corrispondente evocando solo le caratteristiche salienti e facendole oscillare in modo sincrono, senza evocare quelle marginali né quelle di altri concetti.

In nessun caso si è verificato che una caratteristica marginale sia stata evocata dalle altre caratteristiche. In particolare i risultati mostrano che risulta sufficiente una caratteristica distintiva affinché il modello funzioni correttamente per tutti gli animali delle reti testate.

Formazione di categoria: se vengono stimulate le caratteristiche di categoria, il modello riconosce correttamente la categoria evocando e facendo oscillare in sincronismo tutte le proprietà che vi appartengono. Le caratteristiche distintive dei singoli animali non vengono richiamate.

Riconoscimento simultaneo di più concetti: se vengono date in input contemporaneamente più proprietà appartenenti a concetti diversi, il modello riesce a riconoscere simultaneamente i concetti desincronizzando le proprietà salienti di ciascuno di essi e facendole oscillare in maniera sincrona. Questo è verificato nel caso in cui come input vengano date tutte proprietà salienti, tutte proprietà marginali oppure nel caso in cui alcuni concetti vengano stimolati con proprietà salienti e altri con proprietà marginali. I risultati mostrano che risulta sufficiente una caratteristica distintiva (marginale o salienti) affinché il modello funzioni correttamente per tutti gli animali delle reti testate.

Sebbene in linea di massima si possa affermare che il modello con oscillatori lavori correttamente, si sono riscontrati alcuni limiti. Innanzitutto, durante le simulazioni relative alle diverse reti analizzate, alcuni parametri corrispondenti agli oscillatori (γ e σ), all'inibitore globale (th e k) e al valore numerico da assegnare alle sinapsi eccitatorie (WE), hanno dovuto subire delle modifiche affinché si ottenesse il comportamento desiderato. Tali modifiche si basano su criteri che sono stati scelti caso per caso in quanto non è stato possibile trovare un criterio generale che andasse bene per tutti i casi di studio. Inoltre, si è notato che il sincronismo ottenuto non risulta essere molto robusto nel tempo e la desincronizzazione, sebbene avvenga in tutti i casi in esame, non sempre risulta perfetta.

CAPITOLO 4

DIPENDENZA DAL CONTESTO

4.1 Conoscenza concettuale: proprietà CD e proprietà CI

Tradizionalmente la memoria semantica viene descritta come *context-free* e le rappresentazioni concettuali in essa contenute (ovvero le nozioni interne che ogni individuo ha riguardo ad un certo concetto) sono considerate statiche. Ad esempio, si ha l'impressione che il concetto di "limone" abbia un certo significato (un frutto acido, giallo, a forma di elissoide) e che questo significato non vari, anche quando si pensa al concetto di "limone" in contesti diversi. Studi recenti sulla conoscenza concettuale hanno invece permesso di affermare che le rappresentazioni concettuali possano essere considerate dinamiche, quindi in continuo cambiamento e intrinsecamente legate al contesto in cui appaiono. Quanto appena detto può trovare un riscontro positivo nell'architettura distribuita delle rappresentazioni concettuali: i concetti sono rappresentazioni multidimensionali distribuite tra regioni cerebrali che stanno alla base di funzioni sensoriali, percettive e motorie. Pertanto il concetto di "limone" contiene caratteristiche distribuite che, seppur interconnesse le une con le altre, possono essere più o meno attive (ad esempio, a seconda delle circostanze). [8]

In generale alcune teorie sostengono che tutte le proprietà di un concetto siano attive in tutte le occasioni, altre ritengono invece che le proprietà di un concetto siano completamente determinate dal contesto e che quindi si attivino solo in determinate circostanze [5]. L'approccio di Barsalou si può collocare tra queste due distinte concezioni. Tale approccio ad oggi risulta essere il più convincente, in quanto è in grado di spiegare il maggior numero di dati sperimentali e di superare i punti critici delle altre teorie. Esso attesta che i concetti contengano due differenti tipi di proprietà: proprietà indipendenti dal contesto (CI) e proprietà dipendenti dal contesto (CD). Le prime si attivano sempre quando si pensa ad un determinato concetto, le seconde invece si attivano solo se si pensa al concetto collocato in un determinato contesto [5].

Ad esempio, si prenda in considerazione il concetto di “pianoforte”. Se si pensa a tale concetto, senza collocarlo in nessun particolare contesto, proprietà che possono venire frequentemente in mente (cioè attivarsi) possono essere “ha i tasti bianchi e neri” oppure “usato per suonare”. In generale cioè vengono in mente tutte quelle proprietà che identificano il concetto in esame e che quindi vengono attivate ogni volta che si pensa a quest’ultimo (proprietà CI). Ora, se si pensa allo stesso concetto di “pianoforte” collocandolo all’interno di un particolare contesto (es. trasloco), oltre alle proprietà sopra citate, potrebbero attivarsi anche proprietà come “pesante”. In generale cioè si attivano anche quelle proprietà che sono attinenti al concetto in esame, ma alle quali si pensa solo in particolari circostanze (proprietà CD). La distinzione tra proprietà indipendenti dal contesto (CI) e proprietà dipendenti dal contesto (CD) risulta utile soprattutto nei casi in cui vi sia una grande quantità di informazioni o di molteplici significati associati un dato concetto. In questi casi i processi di controllo cognitivi svolgono un ruolo centrale nel garantire che vengano richiamate solo le giuste informazioni, pertanto senza le informazioni contestuali la comprensione potrebbe risultare meno adeguata alle particolari esigenze di volta in volta presentate dall’ambiente esterno. Diversi test dimostrano infatti che il processo di comprensione sia fortemente supportato dall’aggiunta di informazioni contestuali, che a loro volta sono strettamente legate all’esperienza.

4.2 Modello con contesto

4.2.1 Descrizione quantitativa e qualitativa

Il modello con contesto descritto di seguito si basa sulle stesse equazioni del modello di base (Eq. (1) e (2)) e i concetti sfruttati nelle simulazioni sono presi dalle tassonomie descritte nel *cap. 2*. Nella rete adoperata in questo modello però, a differenza di quella di base, per ogni concetto sono state scelte 3 proprietà marginali distintive che complessivamente potessero rappresentare un certo contesto e sono quindi state considerate proprietà CD.

Nel presente lavoro di tesi la dipendenza di un concetto dal contesto viene studiata sfruttando la probabilità condizionata. Quest'ultima indica come varia la probabilità al crescere delle informazioni disponibili a priori sugli eventi coinvolti. In particolare la probabilità condizionata risulta utile quando si vuole analizzare un evento A (evento condizionato), se si dispone di una certa informazione B (evento condizionante):

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (8)$$

dove $P(A|B)$ indica la probabilità che si verifichi A, dato l'evento B. $P(A \cap B)$ indica la probabilità congiunta, ovvero la probabilità del verificarsi simultaneo dei due eventi. $P(B)$ è la probabilità marginale di B, ovviamente deve valere $P(B) > 0$ in quanto non è possibile definire la probabilità condizionata rispetto ad un evento impossibile.

Dalla definizione di probabilità condizionata (Eq. (8)) è possibile inoltre esprimere la probabilità congiunta come segue:

$$P(A \cap B) = P(A|B) P(B) \quad (9)$$

A questo punto, sfruttando la definizione di probabilità congiunta, è ammissibile interpretare la probabilità marginale in modo diverso. Ad esempio, dati due eventi B e \bar{B} , sotto certe condizioni la probabilità dell'evento A, $P(A)$, può essere calcolata sommando la probabilità congiunta di A e B alla probabilità congiunta di A e \bar{B} [10]:

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(B)P(A|B) + P(\bar{B})P(A|\bar{B}) \quad (10)$$

La dipendenza dal contesto di alcune proprietà comporta un addestramento più complesso in quanto utilizzeremo la probabilità condizionata fra i vari concetti per determinare proprietà CD. Per chiarire quanto appena detto, si considerano tre proprietà dipendenti dal contesto che genericamente vengono nominate B, A e C. Al termine dell'addestramento esse dovranno essere tra loro legate come rappresentato nello schema seguente:

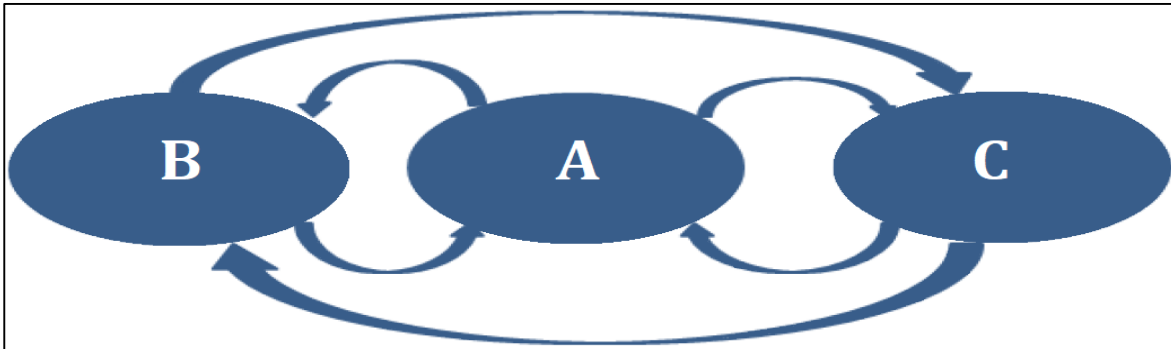


Figura 41 - Schema che rappresenta come le tre proprietà sono tra loro connesse. Le frecce rappresentano le sinapsi che si creano a fine addestramento.

Come nel modello di base, le proprietà salienti e marginali presentano rispettivamente una probabilità del 70% e del 35%. Alla proprietà condizionante (B) viene assegnata la stessa probabilità che contraddistingue le caratteristiche marginali ($\text{perc}_B=0.35$), mentre la probabilità delle proprietà condizionate viene calcolata tramite l'Eq. (10). In generale, quello che dovrebbe accadere, dopo un buon addestramento, è che quando si verifica la prima proprietà (B), anche la seconda (A) e la terza (C) devono attivarsi. Quando invece la prima proprietà non si verifica, neppure le altre due devono attivarsi (a meno che non vengano stimolate da un input). In particolare nella nostra rete, quando la proprietà B si verifica (cioè nel 35% dei casi), le proprietà A e B devono avere un'alta probabilità di verificarsi, mentre quando B non si verifica (cioè nel 65% dei casi) le proprietà A e C devono avere una bassa probabilità di verificarsi. Per ottenere questo comportamento devono essere assegnati dei valori opportuni alle probabilità condizionate. La probabilità che la *feature* condizionate A si presenti dal momento che si è verificata B, indicata come $P(A|B)$, e la probabilità che C si presenti dal momento che si è verificata A, indicata come $P(C|A)$, devono essere abbastanza elevate e sono state settate ad un valore massimo pari a $P_{max} = 0.85$. Di contro, la probabilità che A si presenti dal momento che non si è verificata B, indicata come $P(A|\bar{B})$, e la probabilità che C si presenti dal momento che non si è verificata A, indicata come $P(C|\bar{A})$, devono essere molto basse e si è scelto un valore pari a $P_{min} = 0.05$.

Con questi valori, la proprietà A si verifica nel 85% dei casi quando c'è B, formando una sinapsi, e solo nel 5% dei casi quando non c'è. La proprietà C si verifica nel 85% dei casi quando c'è A e solo nel 5% dei casi quando non c'è. Utilizzando tali valori nell'Eq. (10), si ottengono le probabilità marginali delle proprietà condizionate A e C:

$$P(A) = P(B)P(A|B) + P(\bar{B})P(A|\bar{B}) = 0.35 * 0.85 + 0.65 * 0.05 = 0.33$$

$$P(C) = P(A)P(C|A) + P(\bar{A})P(C|\bar{A}) = 0.33 * 0.85 + 0.67 * 0.05 = 0.31$$

I valori ottenuti sono concordi con il modello dato che risultano minori della soglia post sinaptica di base utilizzata (0.67). Per riprodurre quanto appena detto all'interno del modello neuro-computazionale, sono stati adoperati i seguenti codici matlab durante l'addestramento:

$$I(B) = 1 * (\text{rand}(1,1) < \text{perc}_B)$$

$$I(A) = I(B) * (\text{rand}(1,1) < P_{max}) + (1 - I(B)) * (\text{rand}(1,1) < P_{min})$$

$$I(C) = I(A) * (\text{rand}(1,1) < P_{max}) + (1 - I(A)) * (\text{rand}(1,1) < P_{min})$$

4.2.2 Addestramento

La rete viene addestrata attraverso la regola di Hebb. Durante la fase di addestramento, che consiste in 1000 epoche consecutive, ogni concetto viene presentato una sola volta durante ogni epoca. Le caratteristiche non hanno tutte la stessa probabilità adoperata come input per il dato concetto: le proprietà salienti hanno un'alta probabilità di verificarsi ($P_{kj(s)} = 0.7$), le proprietà marginali e quelle CD hanno invece una bassa probabilità di verificarsi ($P_{kj(m)} = P_{kj(CD)} = 0.35$). Tutti i parametri con i rispettivi valori sfruttati durante questa fase di addestramento sono analoghi a quelli usati durante l'addestramento del modello di base (cap. 2).

4.3 Risultati

Di seguito vengono proposti i risultati ottenuti dopo l'addestramento sinaptico e dopo le simulazioni. Per i risultati ottenuti si è sfruttata la tassonomia di animali, ma risultati affini si sono ottenuti anche testando la tassonomia degli oggetti. Per semplicità ci si è concentrati su un solo concetto della tassonomia degli animali. In particolare, è stato scelto l'animale "cavallo" per il quale è stato supposto che un contesto specifico sia determinato dalle proprietà "veloce", "ha le briglie, e "usato per gare sportive" (racchiuse nel riquadro nero nello schema seguente):



4.3.1 Risultati addestramento

I risultati ottenuti al termine della fase di addestramento sinaptico sono del tutto analoghi a quelli ottenuti per la rete di base: proprietà salienti ricevono forti sinapsi in input da tutte le altre proprietà del concetto tranne che da sé stesse, mentre proprietà marginali non ricevono sinapsi in input dalle altre caratteristiche del concetto. Per quanto riguarda le proprietà dipendenti dal contesto il risultato è il seguente: esse ricevono sinapsi dalle altre proprietà legate al contesto e non da sé stesse, inoltre non ricevono sinapsi dalle altre proprietà del concetto, quindi non sono richiamate al di fuori di quel particolare contesto.

Ogni figura di seguito riportata rappresenta la forza delle sinapsi che entrano in una proprietà CD dalle altre proprietà CD dopo un addestramento di 1000 epoche. In ascissa sono indicate le posizioni relative a ciascuna proprietà mentre in ordinata è indicato il peso sinaptico. In particolare, quello che si evince dai grafici in *fig. 42* è che la proprietà “ha le briglie” (109) riceve sinapsi solo dalle altre due proprietà dipendenti dal contesto (110 e 112), “è veloce” riceve solo dalla 109 e dalla 112, mentre “usato per gare sportive” riceve solo dalla 110 e dalla 109.

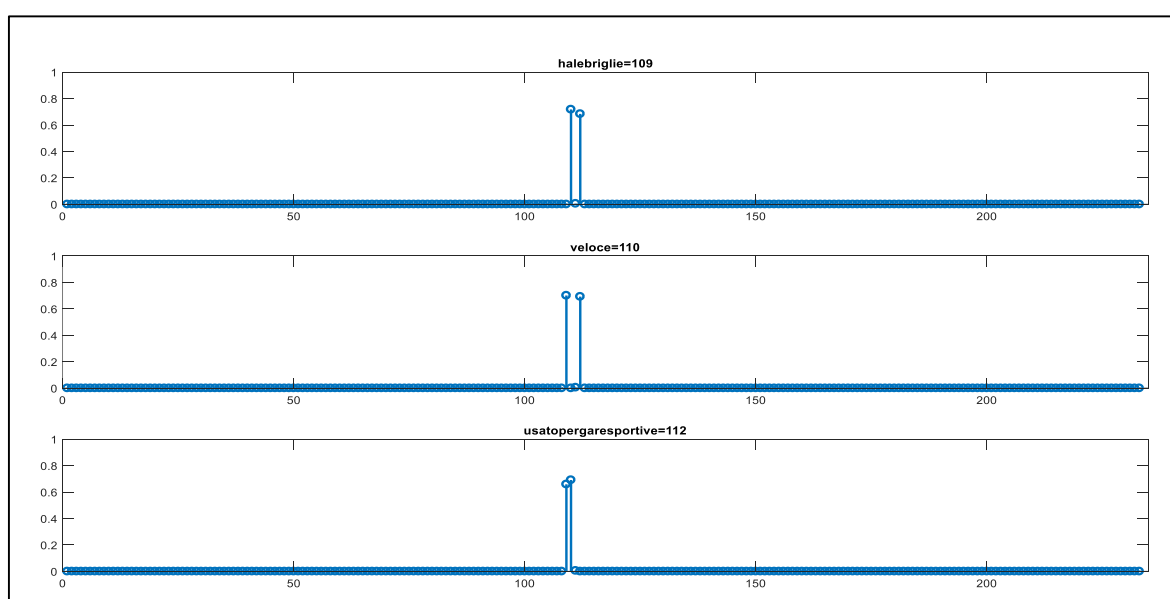


Figura 42 - Forza delle sinapsi entranti nelle proprietà dipendenti dal contesto da ciascuna delle altre due proprietà CD.

4.3.2 Risultati simulazioni

Il modello correttamente addestrato è stato successivamente testato. I risultati di seguito riportati si concentrano sui comportamenti ottenuti quando si considerano proprietà CD e la loro relazione con proprietà CI. I risultati riguardanti il comportamento della rete in tutti gli altri casi, per quanto concerne il riconoscimento di animali e la formazione di categorie, non sono stati riportati in quanto del tutto analoghi a quelli già presentati nel *cap. 2*. In particolare si è verificato che dando in input una proprietà CD, la rete richiama le altre proprietà *context dependent* (oltre alla proprietà salienti del concetto in esame).

A titolo esemplificativo vengono riportati di seguito (*fig.43*) i *pattern* (a 10, 20 e 50 *step* della simulazione) del comportamento della rete se in input viene data la proprietà “ha le briglie” relativa al concetto “cavallo”. In particolare, si evince che vengono richiamate le altre due proprietà CD ovvero la 110 e la 112 e le proprietà salienti del concetto (2-9, 34, 54, 84, 85, 104-108, 203, 214, 218, 219, 223 e 228). Nessuna proprietà marginale viene invece evocata.

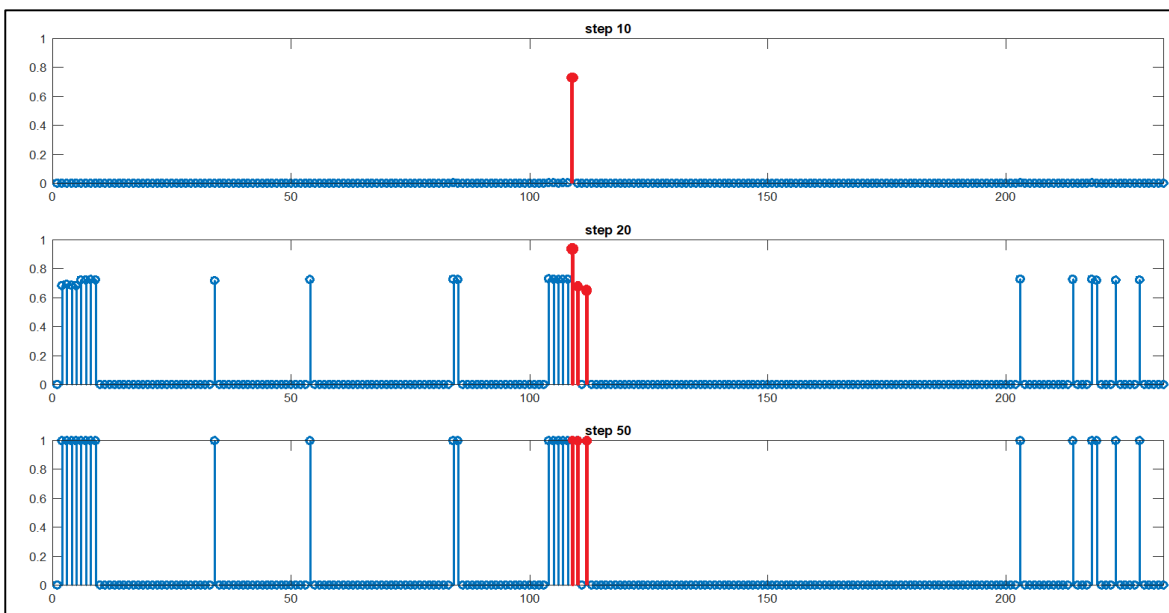


Figura 41 - Pattern relativi al ruolo delle proprietà CD. Richiamando la prima proprietà CD, la rete richiama anche le altre due proprietà CD (rosse), oltre alle proprietà salienti del concetto in esame.

Un comportamento simile si ottiene anche quando in input vengono date la seconda (“veloce”) o la terza (“usato per gare sportive”) proprietà CD: la rete richiama le altre proprietà *context dependent* (oltre alla proprietà salienti del concetto in questione), senza richiamare caratteristiche marginali. Di seguito viene riportato quello che accade se viene data in input una proprietà CI distintiva (saliente o marginale) del concetto. La rete evoca tutte le proprietà salienti del concetto mentre le caratteristiche CD non sono evocate. A titolo esemplificativo vengono riportati di seguito (*fig.44*) i *pattern* (a 10, 20 e 50 *step* della simulazione) del comportamento della rete se in input viene data la proprietà “si cavalca” relativa al concetto “cavallo”.

La rete richiama le proprietà salienti del concetto (2-9, 34, 54, 84, 85, 104-108, 203, 214, 218, 219, 223 e 228). Nessuna proprietà marginale o CD viene invece evocata.

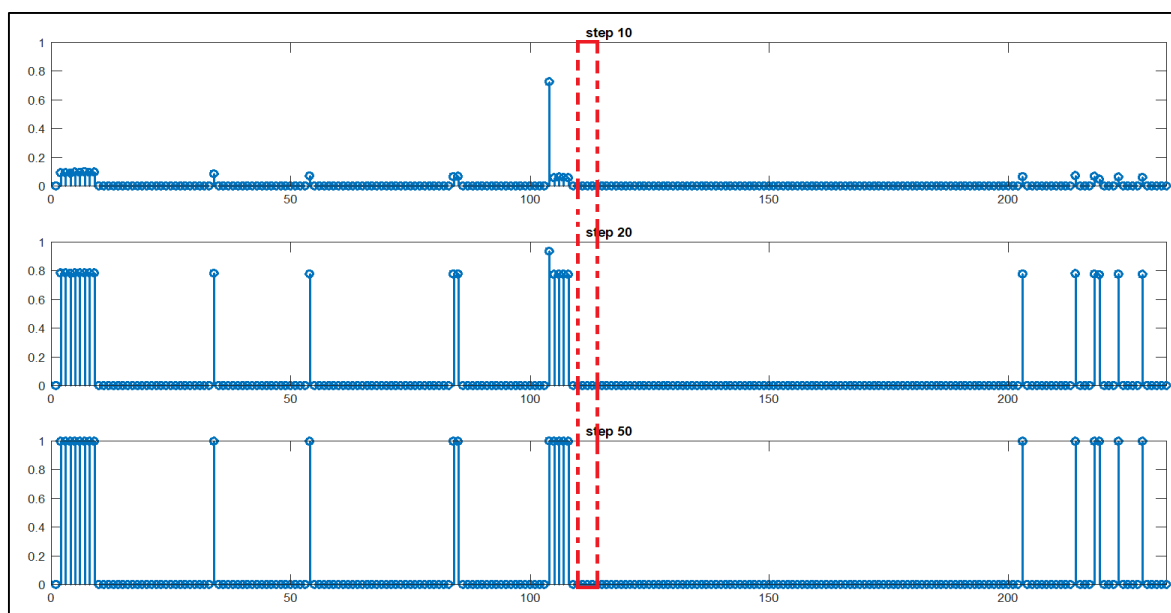


Figura 24 - Pattern relativi alla relazione tra proprietà CD e CI. Richiamando una proprietà CI, la rete evoca tutte le proprietà salienti del concetto in esame, mentre le proprietà dipendenti dal contesto non sono evocate (riquadro rosso tratteggiato).

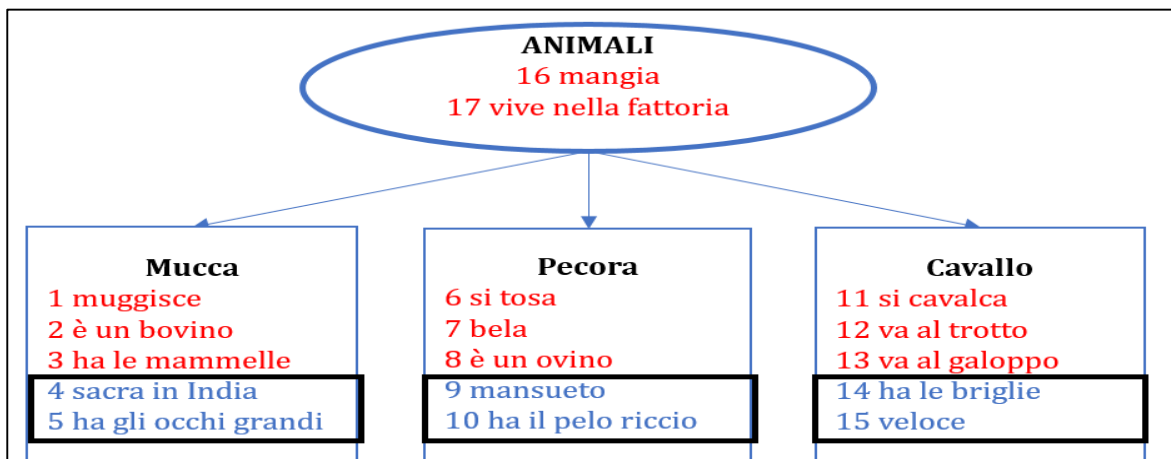
Le prestazioni complessive della presente rete risultano coerenti con il concetto di proprietà *context dependent*: quando al modello viene data in input una proprietà CD, esso richiama correttamente anche le altre proprietà dipendenti dal contesto (oltre a richiamare le proprietà salienti del concetto in esame). Se invece in input viene data una proprietà distintiva CI (saliente o marginale), le proprietà dipendenti dal contesto non sono evocate e, in linea con il comportamento del modello di base, la rete richiama solo le proprietà salienti del concetto in esame.

4.4 Modello contesto e oscillatori

4.4.1 Descrizione e addestramento

Di seguito è stata applicata la dipendenza dal contesto trattata nei paragrafi precedenti ad una delle reti con oscillatori analizzate nel *cap. 3*.

In particolare è stata scelta la rete composta da tre animali e 7 proprietà ciascuno in quanto disponeva di almeno due proprietà marginali da potere adoperare per testare la dipendenza contestuale. Per ciascuno animale della rete le due caratteristiche marginali sono state poste dipendenti dal contesto (proprietà racchiuse nel riquadro nero nello schema seguente):



Come nel modello con oscillatori, le proprietà salienti presentano una probabilità del 70%. Alla proprietà condizionante viene attribuita una probabilità del 40% ($\text{perc}_B = 0.4$). La dipendenza dal contesto ha comportato un addestramento più complesso che prevede di sfruttare la probabilità condizionata per determinare le proprietà CD. In particolare, per ciascun concetto, quando la probabilità B (la prima tra quelle racchiuse nei riquadri neri nello schema) si verifica (cioè nel 40% dei casi), la proprietà A ha un'alta probabilità di verificarsi. Quando invece la proprietà B non si verifica (cioè nel 60% dei casi), la proprietà deve avere una bassa probabilità di verificarsi. Per ottenere questo comportamento devono essere assegnati dei valori opportuni alle probabilità condizionate. La probabilità che la *feature* condizionata A si presenti dal momento che si è verificata B, indicata come $P(A|B)$, deve essere abbastanza elevata ed è stata settata ad un valore massimo pari a $P_{max} = 0.85$. Di contro, la probabilità che A si presenti dal momento che non si è verificata B, indicata come $P(A|\bar{B})$, deve essere molto basse e si è scelto un valore pari a $P_{min} = 0.05$.

Con questi valori, la proprietà A si verifica nel 85% dei casi quando c'è B, formando una sinapsi, e solo nel 5% dei casi quando non c'è. Utilizzando tali valori nell'Eq. (10), si ottiene la probabilità marginale della proprietà condizionata A:

$$P(A) = P(B)P(A|B) + P(\bar{B})P(A|\bar{B}) = 0.40 * 0.85 + 0.60 * 0.05 = 0.37$$

I valori ottenuti sono concordi con il modello con oscillatori dato che risultano minori della soglia post sinaptica di base utilizzata (0.55). Per riprodurre quanto appena detto all'interno del modello neuro-computazionale, sono stati adoperati i seguenti codici matlab durante l'addestramento:

```
I(B)=1*(rand(1,1)<perc_B)
I(A)=I(B)*(rand(1,1)<P_max)+(1-I(B))*(rand(1,1)<P_min)
```

La rete viene addestrata attraverso la regola di Hebb. Durante la fase di addestramento, che consiste in 1000 epoche consecutive, ogni concetto viene presentato una sola volta durante ogni epoca con una diversa frequenza di occorrenza. Tutti i parametri, con i rispettivi valori, sfruttati durante questa fase di addestramento sono analoghi a quelli usati durante l'addestramento del modello con oscillatori (*cap. 3*).

4.4.2 Risultati

I risultati ottenuti dall'addestramento possono essere considerati conformi a quelli ottenuti nella rete semantica di base: le proprietà salienti ricevono forti sinapsi in input da tutte le altre proprietà tranne che da sé stesse, mentre le proprietà marginali ricevono in input sinapsi praticamente nulle da tutte le proprietà. Per quanto riguarda le proprietà dipendenti dal contesto i risultati dell'addestramento sono analoghi a quelli descritti nei paragrafi precedenti: per ogni concetto, ogni proprietà CD riceve sinapsi dall'altra proprietà dipendente dal contesto e non da sé stessa né dalle proprietà CI.

Il modello addestrato è stato successivamente testato. I risultati di seguito riportati si concentrano sui comportamenti ottenuti quando si considerano proprietà CD e la loro relazione con proprietà CI. I risultati riguardanti il comportamento della rete per quanto concerne il riconoscimento di più oggetti simultaneamente (con la rispettiva attivazione in sincrono di tutte le proprietà salienti) e la loro desincronizzazione, non sono stati riportati in quanto del tutto analoghi a quelli già presentati nel *cap. 3*. In particolare si è verificato che dando in input proprietà CD simultaneamente (una per ogni concetto), la rete ha il comportamento atteso: per ogni concetto vengono richiamate sia l'altra proprietà CD, sia le proprietà salienti CI (distintive e condivise). Tutte le proprietà richiamate per ciascun concetto risultano oscillare in maniera sincrona tra di loro e complessivamente risultano sfasate rispetto a quelle degli altri concetti. A titolo esemplificativo in *fig. 45* viene riportato il comportamento della rete quando in input vengono date le proprietà CD “sacra in India”, “mansueto”, e “ha le briglie” rispettivamente per i concetti di mucca, pecora e cavallo.

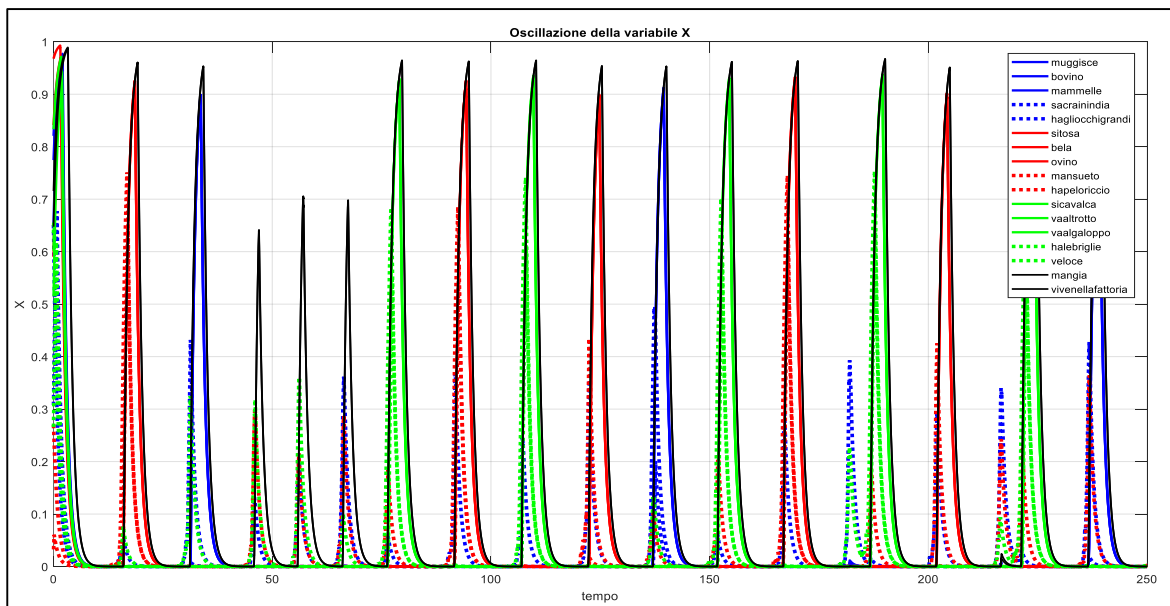


Figura 43 - Comportamento della rete quando in input a ciascun animale viene data una proprietà CD. Per ciascun concetto viene evocata anche l'altra proprietà CD (linee tratteggiate) e le proprietà salienti (linee continue) distintive (blu per mucca, rosso per pecora e verde per cavallo) e condivise (nere).

Di seguito (*fig.46*) è riportata una sezione del grafico precedente per meglio mettere in evidenza il comportamento ottenuto dalla rete. Si nota meglio che, per ciascun concetto, le 2 proprietà CD (linee tratteggiate) si richiamano a vicenda ed oscillano in sincronismo, anche se con ampiezza minore rispetto alle proprietà salienti. In linea con quanto detto precedentemente, inoltre, anche le caratteristiche CI salienti distintive (colorate) e condivise (nere) di ciascun concetto sono attivate simultaneamente in sincrono. Coerentemente con quanto espresso sulla dipendenza contestuale, se in input alla rete viene data una proprietà CI, nessuna proprietà dipendente dal contesto viene invece evocata ed il comportamento della rete può essere considerato ad esempio del tutto analogo a quello riportato in *fig. 21* del Cap. 3 (se le proprietà CI scelte come input sono salienti distintive).

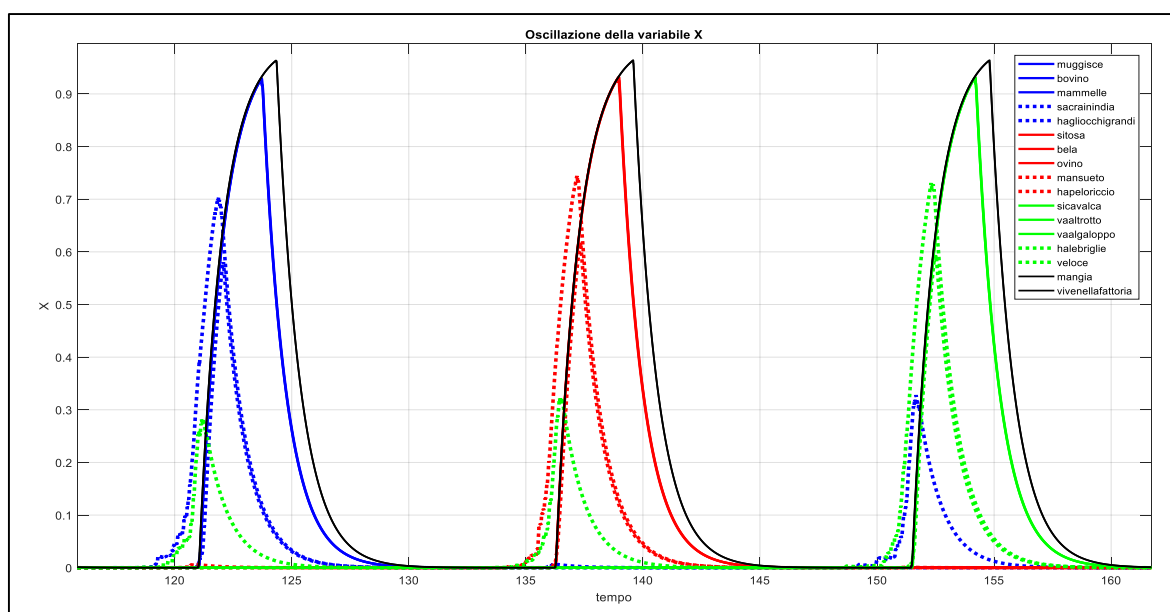


Figura 44 - Dettaglio del comportamento della rete quando in input a ciascun animale viene data una proprietà CD.

CONCLUSIONI

Nel presente lavoro di tesi, sono partita da un modello neurocomputazionale in grado di acquisire automaticamente il significato di concetti e di categorie, sfruttando la probabilità con cui si verificano le diverse proprietà. Per tenere conto della differenza tra caratteristiche distintive e condivise, e per tenere in considerazione il diverso ruolo di salienza e di marginalità, le sinapsi tra i vari neuroni devono essere asimmetriche. A tale fine le sinapsi sono state create attraverso la regola di Hebb utilizzando differenti soglie pre e post sinaptiche. Inoltre, si è addestrato il modello facendo uso di una soglia post-sinaptica variabile che dipende esclusivamente dall'attività del neurone pre-sinaptico e cambia automaticamente in base a quanto frequentemente una determinata feature appare. In questo modo si è ottenuta una buona discriminazione tra salienza e marginalità di proprietà sia distintive che condivise. In particolare, usando una soglia post-sinaptica variabile, è stato possibile evitare che una proprietà condivisa da molti membri della stessa categoria (come "vola" per gli uccelli) divenga saliente e sia attribuita erroneamente all'intera categoria. Con questa assunzione, inoltre, il modello riesce a distinguere concetti che sono molto vicini semanticamente, ovvero che condividono molte caratteristiche (come ad esempio divano e poltrona).

Il modello è stato testato su due ampie tassonomie contenenti rispettivamente 16 animali e 12 oggetti; ciascun concetto è caratterizzato da proprietà condivise (in modo parziale o in modo totale) distintive, marginali o salienti. In linea di massima si è riusciti a verificare che il modello è in grado di codificare le caratteristiche dei vari concetti, riuscendo a risolvere semplici *task* cognitivi sul riconoscimento dei concetti e sulla formazione di categorie. Un limite del modello è legato alle proprietà condivise tra più concetti: talvolta, se per certi concetti una proprietà risulta saliente (sulla base della frequenza di occorrenza durante l'addestramento) e per altri marginale, dopo la fase di addestramento, il modello tende a considerare tale proprietà saliente per tutti i concetti.

Questo non avviene per tutte le proprietà condivise: nelle nostre simulazioni questo fenomeno si è riscontrato solo per la caratteristica “si trova allo zoo” condivisa da giraffa, elefante, orso, zebra e pappagallo, saliente per i primi due e marginale per i restanti. Tale fenomeno è legato alla creazione in modo casuale delle sinapsi. Si è dimostrato che è possibile superare questo limite, abbassando ulteriormente la frequenza di occorrenza della proprietà marginali ad un valore pari 0.3 (invece che 0.35). Partendo dai risultati sopra descritti, al fine di analizzare il sincronismo neuronale e il riconoscimento simultaneo di più concetti, nel modello sono state sostituite le unità neurali con unità oscillanti. Tali simulazioni sono state effettuate su reti contenenti meno concetti rispetto a quelle di base, studi con reti semantiche più ampie possono essere oggetto di lavori futuri. Le simulazioni fatte hanno permesso di verificare la capacità di fare oscillare in maniera sincrona le proprietà salienti di un medesimo concetto e di sfasarle rispetto a quelle di altri concetti presentati in simultanea. Sebbene, in linea di massima, il modello con oscillatori abbia fornito buoni risultati, sono stati riscontrati anche alcuni limiti. Innanzitutto durante le simulazioni relative alle diverse reti analizzate, alcuni parametri corrispondenti agli oscillatori (γ e σ), all'inibitore globale (th e k) e al valore numerico da assegnare alle sinapsi eccitatorie (WE), hanno dovuto subire delle modifiche affinché si ottenesse il comportamento desiderato. Tali modifiche si basano su criteri che sono stati scelti *ad hoc* per ogni rete in quanto non è stato possibile trovare un criterio generale che andasse bene per tutti i casi in esame. Inoltre, si è notato che il sincronismo ottenuto non risulta essere molto robusto nel tempo e la desincronizzazione, sebbene avvenga in tutti i casi analizzati, non sempre risulta perfetta. Successivamente, si è voluta simulare la dipendenza della semantica di un concetto dal contesto. Per raggiungere tale scopo, è stato usato un addestramento sinaptico che facesse uso delle probabilità condizionate fra i vari concetti.

Le simulazioni svolte sulle reti di base hanno fornito risultati positivi: le proprietà dipendenti dal contesto ricevono sinapsi dalle altre proprietà legate al contesto, mentre non ricevono sinapsi dalle altre proprietà non dipendenti dal contesto, quindi non vengono evocate al di fuori di quella particolare circostanza. Inoltre, simulando la dipendenza contestuale anche in modelli con oscillatori, si è verificato che le proprietà legate al contesto riescono ad oscillare in sincronismo e riescono ad essere sfasate rispetto alle proprietà contesto-dipendenti di altri concetti presentati simultaneamente, verificando che diversi contesti non si influenzano tra loro. In conclusione possiamo affermare che i modelli descritti in questo lavoro di tesi hanno rispecchiato gli obiettivi che ci siamo prefissati a inizio lavoro e che forniscono uno strumento promettente per successive applicazioni sia in ambito teorico, per una migliore comprensione dei meccanismi alla base della memoria semantica, in ambito neurologico.

APPENDICE

Di seguito vengono riportati gli schemi a blocchi che descrivono in maniera dettagliata le due tassonomie utilizzate nel presente lavoro di tesi, in particolare la tassonomia degli oggetti (fig.1) e la tassonomia degli animali (fig. 2). In generale ad ogni proprietà è stato associato un numero che corrisponde alla posizione che ogni attributo assume all'interno della rete. Le proprietà contenute all'interno delle ellissi formano delle categorie e sono totalmente condivise da tutti i concetti appartenenti a quella specifica classe ed etichettate come salienti (rosse). Le proprietà contenute nei rettangoli invece appartengono ai singoli concetti e possono essere salienti (rosse), marginali (blu), distintive (non sottolineate) o condivise (sottolineate).

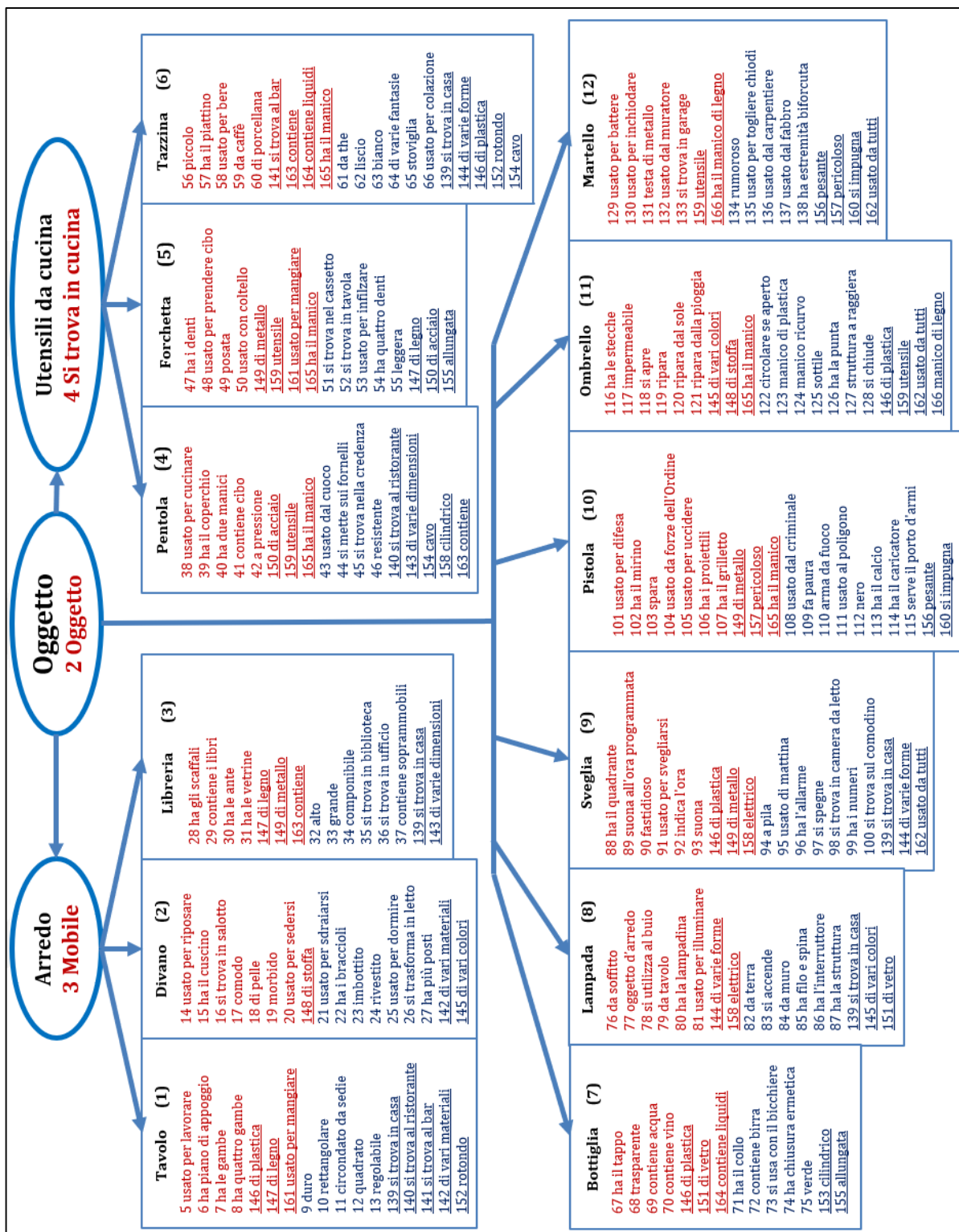


Figura A 1 Schema a blocchi Oggetti

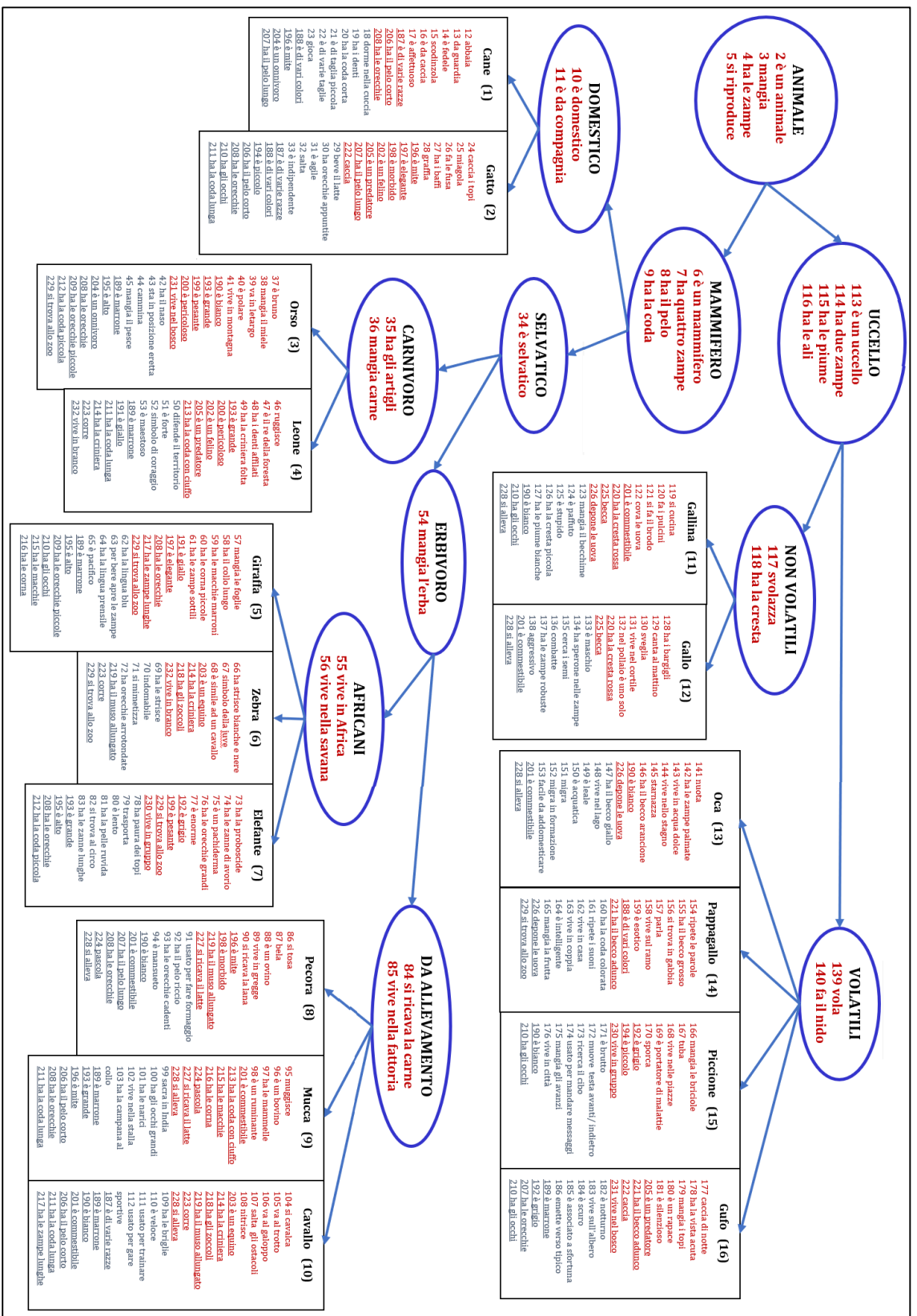


Figura A 2 Schema a blocchi Animali

BIBLIOGRAFIA

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] Barsalou L.W. (1982), **“Context-independent and context-dependent information in concepts”**, *Memory & Cognition*, Vol. 10 (1), pp. 82-93.
- [6] Ursino M. (2017), **“Dispense di Sistemi Neurali”**, Facoltà di Ingegneria, Corso di Laurea Magistrale in Ingegneria Biomedica, Università di Bologna.
- [7] 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.

- [8] Eiling Yee, Sharon L. Thompson-Shill (2016), “**Putting concepts into context**”, *Psychonomic Bulletin & Review*, Vol. 23 (4), pp. 1015-1027.
- [9] 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.
- [10] Levine D. M., Krehbiel D. C., Berenson M. L., *Statistica*, Milano, APOGEO, 2006 (II edizione), pp. 117-150.
- [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] 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.