

Basilio Calderone

L'emergenza del Paradigma.

Un modello di apprendimento non-supervisionato applicato al sistema verbale dell'italiano

Il presente lavoro si pone come obiettivo primario un'indagine sulla natura emergente degli schemi associativi che organizzano, in linea teorica, il lessico mentale di un parlante. In questa prospettiva, sarà condotta un'applicazione sperimentale delle Self-Organizing Maps (SOMs) ai dati della flessione verbale dell'italiano, considerati come input al dispositivo neurale.

Una SOM, come si specificherà meglio in seguito, si presenta (nel nostro caso) come una struttura connessionistica, capace di organizzare in categorie linguistiche dati di ingresso di natura morfologica.

L'organizzazione è non supervisionata, nel senso che non vi è presenza di un supervisore esterno che informi il sistema sulle corrette modalità della ristrutturazione categoriale.

In tali dispositivi neurali, l'apprendimento di dati di input viene a coincidere con la realizzazione di una configurazione spaziale degli stessi su una superficie (mappa) generalmente bi-dimensionale. In tal modo una SOM è in grado di rilevare l'esistenza di similarità pertinenti tra i tratti che descrivono i dati in ingresso e al contempo di categorizzarli in base alla scoperta di tale pertinenza. La pertinenza categoriale è ottenuta mediante principi di rilevanza statistica.

Una simulazione dell'apprendimento morfologico mediante SOMs mira così a riprodurre la realizzazione di un possibile schema di auto-organizzazione, emergente e funzionale (ai fini organizzativi e di ristrutturazione), dei dati cui la rete è esposta in fase di addestramento. Questo schema sarà rappresentato visivamente come una distribuzione topologica dei dati di input sulla mappa, in base ad un qualche criterio emergente di pertinenza linguistica, individuato dalla SOM in modo incrementale.

1. Introduzione

Definire cosa si intenda esattamente per "apprendimento" significherebbe oggi sintetizzare tre quarti di resoconti e risultati speculativi raggiunti all'interno del panorama di ricerca degli studi cognitivi. Una definizione operativa che possa intendersi come punto di partenza che accenti e accomuni gran parte degli addetti ai lavori, deve far riferimento, di necessità, al processo di ristrutturazione al quale il soggetto (o l'automa) discende sottopone il sistema di conoscenze acquisite, per far fronte a specifiche esigenze dettate dall'ambiente esterno. Questo assunto teorico è alla base della realizzazione, in vari ambiti interdisciplinari, di sistemi d'apprendimento che, in qualche modo, tendono a simulare - a volte in una prospettiva biologica-organica, più spesso in termini funzionali-sistemici - il comportamento intelligente esibito da umani e animali.

Per poter parlare di apprendimento da parte di sistemi artificiali occorre avere una teoria che chiarisca in che modo la conoscenza acquisita viene "rappresentata" al loro interno. Nell'ambito delle scienze cognitive in genere, il problema della *rappresentazione della conoscenza* ha trovato espressione nella dicotomia terminologica, molte volte abusata ed enfaticizzata a torto, tra *rappresentazione simbolica* e *rappresentazione sub-simbolica* (o

neurale).¹ La prima basa il suo assetto teorico e metodologico sulla nozione di *simbolo*² (nel senso tecnico adottato dalla Teoria dell'Informazione³), inteso come elemento scomponibile dotato a priori di una valenza semantica e di proprietà sintattiche. In quanto unità "pronta all'uso" con forte potere compositivo, il simbolo, il cui ruolo è definito in base a regole e criteri di associazione con altri simboli, entra a far parte di macro-rappresentazioni, aumentando così il proprio potere descrittivo. Di contro, una rappresentazione sub-simbolica fa riferimento al concetto di rete neurale, ossia a un dispositivo di elaborazione parallela, composto da unità – preposte ad elaborare segnali in ingresso - e connessioni, definite da un coefficiente di attivazione tra le varie unità.

In quest'ultimo quadro metodologico, noto anche sotto il nome di connessionismo, il termine "apprendimento" indica un processo di modifica dei coefficienti di connessione, dipendente in qualche modo dalla natura delle configurazioni (pattern) di stimolazione esterna.

2. Rete neurale. Una definizione

Una rete neurale⁴ è un sistema composto da:

- 1) *unità elementari* (chiamate, con una terminologia non uniforme, *moduli*, *neuroni*, *nodi* o semplicemente *unità*) dotate di *linee di ingresso* e di *uscita*.
- 2) *connessioni* tra queste unità. Le connessioni sono caratterizzate da una precisa direzione, nel senso che veicolano il segnale in uscita da un'unità verso un'altra unità. A ciascuna di esse è associato un numero, detto *peso* o *coefficiente di connessione*. Si assume convenzionalmente che, se il peso è positivo, la connessione è *eccitatoria*, se è negativo è *inibitoria* e se è nullo non esiste connessione.

La struttura delle connessioni definisce l'architettura di una rete. Tra i tipi di architettura più usati vi sono quelle *monostrato*, in cui tutte le unità della rete appartengono al medesimo strato, e quelle *multistrato*, in cui vi è una ripartizione in più strati. Se esistono, oltre a quelli di ingresso e di uscita, altri strati, le cui unità non ricevono da e non inviano segnali direttamente verso l'esterno, essi si definiscono *strati intermedi* o *nascosti* (vedi Figura 1). Le unità che li compongono vengono chiamate anche *unità nascoste*.

La rete può essere classificata anche a seconda della struttura di connessioni che presenta al proprio interno. Distinguiamo: le connessioni *feedforward*, tutte quelle

¹ Si veda Chalmers (1992), Dyer (1990), Dinsmore (1997) per una prospettiva epistemologica al problema. Testi di impronta più specificatamente cognitiva sono Pessa & Penna (1993), Fodor & Pylyshyn (1988). Per la diatriba *Simbolico Vs Sub-simbolico* all'interno delle scienze del linguaggio, tra i molti riferimenti, si veda Miikkulainen (1997), Pinker & Mehler (1988), Smolensky (1992).

² Newell & Simon (1976).

³ Shannon & Weaver (1971).

⁴ La bibliografia sulle reti neurali è piuttosto nutrita. A titolo di esempio si veda Hebb (trad. it. 1975), Feldman & Ballard (1982), Anderson & Rosenfeld (1988), Parisi (1989), Pessa (1993), Rumelhart & McClelland (1986), Gallant (1995), Plunkett & Elman (1998). Un testo introduttivo e di carattere storico è Penna & Pessa (1993).

connessioni, cioè, che sono dirette dall'ingresso verso l'uscita, ovvero dagli strati inferiori verso quelli superiori; le connessioni *lateral*i, che collegano unità appartenenti allo stesso strato e infine le connessioni *feedback*, che dall'uscita rientrano verso l'ingresso, cioè vanno dagli strati superiori verso gli strati inferiori.

In alcuni tipi di rete, se esiste una connessione da un'unità A ad un'unità B, caratterizzata da un apposito peso, vi è necessariamente anche una connessione inversa dall'unità B all'unità A, caratterizzata dallo stesso peso. Reti dotate di questa proprietà vengono chiamate reti *simmetriche*, o dotate di connessioni simmetriche. Le altre tipologie di reti sono chiamate *asimmetriche*.

I segnali che viaggiano lungo le linee di connessione possono essere solo di due tipi ed assumere valore 0 e 1, oppure -1 e $+1$. Tali reti vanno sotto il nome di *reti binarie*.

In altri tipi di reti i segnali possono assumere qualsiasi valore o valori compresi entro un certo intervallo. Questo è il caso, fra l'altro di reti contenenti unità lineari o sigmoidali (ne parleremo più avanti). Generalmente l'evoluzione temporale di una rete avviene in modo discreto, cioè secondo una scala dei tempi del tipo $t, t+1, t+2, \dots$

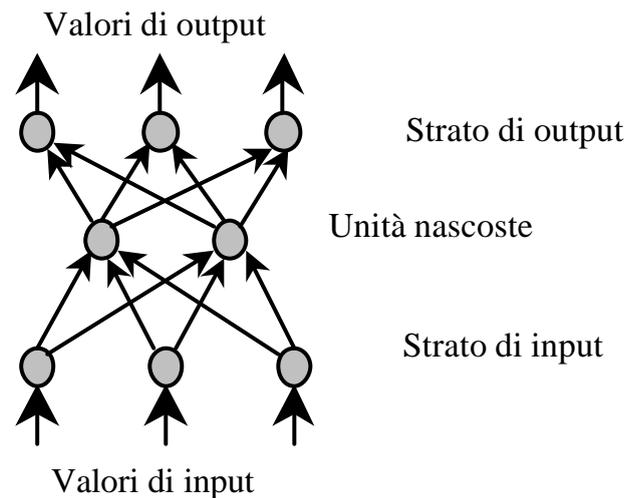


Figura 1. Rete multistrato di tipo feedforward.

A sua volta, ogni unità è caratterizzata da una *legge di attivazione*, che può essere formulata facendo uso di due entità: il *potenziale* e la *funzione di uscita*. Il potenziale rappresenta, in un certo senso, lo stato interno momentaneo della singola unità e dipende, a sua volta, dai segnali provenienti dalle linee di ingresso e dai pesi che le caratterizzano; la funzione di uscita stabilisce in che modo il segnale trasmesso lungo la linea di uscita di una data unità dipende dal potenziale assunto da quella stessa unità nell'istante precedente.

Se indichiamo con $P(t)$ il potenziale di una data unità al tempo t e con $f(P(t))$ la funzione di uscita di quella unità, ovvero la sua uscita al tempo $t+1$ (supponiamo di aver a che fare con una scala dei tempi discreta), esempi tipici di funzioni di uscita sono:

$$-f(P(t)) = P(t) \quad \text{funzione lineare}$$

$$-f(P(t)) = \frac{1}{1+e^{-P(t)}} \quad \text{funzione sigmoide}$$

$$-f(P(t)) = \begin{cases} 1 & \text{se } P(t) \geq 0 \\ 0 & \text{se } P(t) < 0 \end{cases} \quad \text{funzione a gradino}$$

Gli stessi esempi sono illustrati graficamente in Figura 2.

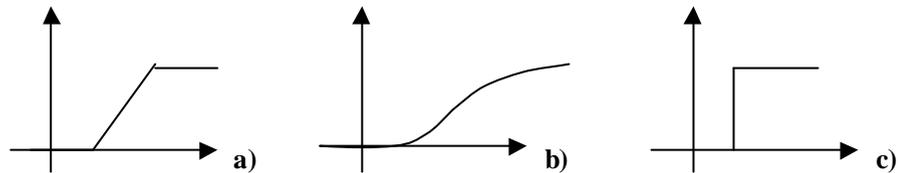


Figura 2. Esempi di funzioni di attivazione a) funzione lineare b) funzione sigmoide c) funzione a gradino.

“In pratica quello che succede è che ciascuna delle unità interne fa la somma (algebraica) delle eccitazioni ed inibizioni che arrivano dalle diverse unità di input e in base a questa somma determina il proprio livello di attivazione. Questa simula la cosiddetta funzione integrativa del sistema nervoso. Durante la trasmissione l’informazione sensoriale viene trasformata (elaborata) dai pesi e dal carattere eccitatorio o inibitorio delle connessioni che collegano le unità di input alle unità interne. In altre parole il pattern d’attivazione delle unità di input viene ricodificato nel pattern di attivazione delle unità interne. A loro volta le unità interne trasmettono l’attivazione alle unità del terzo strato, quello più in alto. Lo strato di output⁵”.

Generalmente, il valore del potenziale di attivazione di una unità ad un certo istante si ottiene sottraendo il valore della soglia da quello della somma pesata degli ingressi.

Indichiamo ora con il simbolo w_{ij} il coefficiente di connessione della linea di ingresso all’unità contrassegnata con l’indice i e proveniente dalla linea di uscita dell’unità contrassegnata con l’indice j ; sia poi $u_j(t)$ il valore del segnale in uscita da quest’ultima al tempo t e s_i la soglia dell’unità i . Il potenziale, secondo quanto detto sopra, sarà:

$$P_i(t) = \sum w_{ij} u_j(t) - s_i$$

⁵ Parisi (1999: 83).

Accanto alla legge di attivazione delle unità, molti modelli connessionistici sono dotati di una legge di apprendimento, ovvero di variazione nel tempo dei coefficienti di connessione in funzione di vari processi che possono interessare la rete.

Nel paragrafo seguente si provvederà a fornire maggiori dettagli sul processo di apprendimento che una rete subisce per rispondere, in seguito, in maniera appropriata alla presentazione di determinati patterns di ingresso e presentare, quindi, comportamenti omogenei e regolari nel tempo.

In particolare verranno tracciati brevemente, a titolo di esempio, due tipologie di processi di apprendimento: a) apprendimento supervisionato (*supervised*) e b) apprendimento non-supervisionato (*unsupervised*). Quest'ultimo avrà ruolo introduttivo per l'esposizione della rete SOM (facente parte, appunto, del paradigma non-supervisionato), che sarà utilizzata, nel corso del presente lavoro, per la scoperta di specifiche categorie linguistiche su base morfologica.

2.1. L'apprendimento

La capacità di apprendere rappresenta forse l'elemento di maggiore attrazione dei modelli neurali perché permette di impiegare una rete neurale per risolvere problemi senza dover individuare direttamente la soluzione analitica, ma semplicemente esponendo il modello neurale ad una serie di esempi. In generale, non è facile definire esattamente ciò che s'intenda per apprendimento. In questo concetto vi è il riferimento ad un processo che determina un cambiamento strutturale della conoscenza acquisita, richiesto per far fronte alle esigenze dell'ambiente esterno. Dato che, quando si parla della struttura di una rete neurale, ci si riferisce di solito alla struttura e dinamica delle sue connessioni, nella maggior parte dei casi il termine "apprendimento", riferito ad una rete neurale, indica un processo di modifica dei pesi di connessione.

Nell'approccio sub-simbolico l'apprendimento svolge un ruolo fondamentale, ma è, nello stesso tempo, tanto facile da realizzare quanto difficile da descrivere. Infatti, non è facile individuare che cosa esattamente cambia nella complicata matrice dei coefficienti di connessione in seguito ad un processo di apprendimento, dato che si ha a che fare con una modalità di elaborazione dell'informazione di tipo parallelo e distribuito e non è possibile sapere, in linea di principio dove siano immagazzinate le informazioni necessarie possedute dalla rete, dato che sono distribuite in modo sparso nell'intera struttura di questi coefficienti. Questo è un grosso svantaggio dell'approccio sub-simbolico, che non riesce in alcun modo a tener conto delle differenze qualitative tra le diverse informazioni disponibili a un sistema, mentre ciò può essere fatto in modo molto semplice utilizzando un approccio di tipo simbolico (Pinker & Mehler, 1988).

Le leggi di apprendimento si possono ripartire in due grandi classi:

- *Apprendimento supervisionato* (Supervised learning), in cui la rete è addestrata usando un insieme di coppie di vettori (\mathbf{x} , \mathbf{d}) che rispettivamente sono il vettore d'ingresso ed il vettore di uscita desiderato. Tali coppie sono fornite alla rete dallo sperimentatore. L'addestramento consiste nella modifica dei pesi in modo da minimizzare una particolare funzione di errore che tenga conto della differenza tra

l'uscita fornita dalla rete e quella desiderata. In altri termini la rete deve "imparare" una generica funzione $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$ (con n e m rispettivamente numero degli ingressi e delle uscite della rete) interpolando le coppie ingresso-uscita $(x, F(x))$ che le vengono presentate durante l'apprendimento. L'apprendimento ha termine quando il segnale di errore scende al di sotto di un valore prefissato dallo sperimentatore⁶.

- *Apprendimento non-supervisionato* (Unsupervised learning). Non esiste una definizione precisa di quale sia la risposta corretta della rete. In questo tipo di paradigma di addestramento la rete è sottoposta ad un insieme di vettori di cui è in grado di rilevare alcune caratteristiche statistiche. Uno dei compiti che la rete è capace di svolgere è la classificazione degli ingressi suddividendoli in gruppi omogenei (*clusters*). Una rete di questo tipo è capace di apprendere una funzione $F: \mathbf{R}^n \rightarrow \mathbf{R}^n$ dello spazio di ingresso in sé stesso che associ ad ogni ingresso l'elemento rappresentativo del gruppo di appartenenza.

In accordo con le finalità del presente lavoro, si focalizzerà l'attenzione sulla struttura ed i meccanismi propri di una particolare tipologia di rete che utilizza, appunto, un apprendimento di tipo non supervisionato, le cosiddette *SOMs* (Self-Organizing Maps⁷).

3. La SOM (Self-Organizing Map)

La Self-Organizing Map è una struttura connessionistica i cui neuroni sono organizzati all'interno di un *array* di superficie mono-dimensionale o più spesso bi-dimensionale (Figura 3). La mappa riceve in ingresso dei vettori di input e, ad addestramento ultimato, è in grado di esibire sulla sua superficie una compressione topologica di tutti i dati vettoriali acquisiti. In tal modo similarità ed analogie (spaziali) dei vettori presentati in ingresso, riceveranno medesimo trattamento dalla SOM che provvederà ad ordinarli su intorno spaziali prossimi. Ad una prima analisi, la SOM svolge il ruolo di *clusterizzatore* di dati d'ingresso, i quali vengono accomunati spazialmente in base ad una qualche loro caratteristica più o meno soggiacente. Nelle migliori delle ipotesi, ad ogni neurone della rete corrisponderebbe un *cluster* e cioè un raggruppamento omogeneo dei dati di input, localizzati in quella zona per un qualche criterio di similarità individuato dalla rete⁸.

Il processo di addestramento (*training process*) coincide, quindi, con una fase di adattamento spaziale che la SOM compie, tramite modifica dei propri coefficienti di connessione, ad una presentazione di patterns vettoriali considerati segnale di input.

Come già detto, il processo di apprendimento è non-supervisionato, ciò significa che non vi è definizione in anticipo sulla configurazione spaziale cui perverrà la SOM. La variazione dei coefficienti di connessione, e quindi la configurazione della mappa,

⁶ Un'esposizione dettagliata sugli algoritmi per l'apprendimento supervisionato prescinde dalle finalità del presente lavoro. Si veda per maggiori dettagli Gallant (1995), Floreano (1996), Rumelhart & McClelland (1986), Pessa (1992; 1993).

⁷ Testo di riferimento delle SOMs è Kohonen (1995). Si veda anche Kohonen (1984), Honkela (1997).

⁸ Non è esclusa però la possibilità, più frequente per la verità, che sia un assemblaggio di neuroni a definire un unico *cluster*. Ciò è dovuto alla risoluzione della mappa. La risoluzione di una mappa è in relazione con la sua dimensione, aumentando, infatti, i nodi (e quindi anche la dimensione) questi individueranno una categoria più ristretta e più granulare

dipenderà soprattutto dalle caratteristiche statistiche del campione di dati che costituisce la base di conoscenza su cui si informerà tutto il processo di apprendimento.

Ogni qualvolta si presenta un pattern di ingresso, viene infatti eseguito un *matching* tra quel pattern (definito come un vettore a n -dimensioni) e tutte le unità neuronali formanti la mappa bi-dimensionale (definite anch'esse in termini vettoriali a n -dimensioni). Il criterio adottato per stimare la diversità tra i due vettori fa capo alla nozione di "distanza". La "distanza" in questione può essere calcolata in vari modi, solitamente si considera la distanza euclidea. E' importante notare che questo calcolo è reso possibile dal fatto che i due vettori, quello che rappresenta il pattern in ingresso e quello che rappresenta l'unità della mappa, contengono il medesimo numero di componenti.

Si considera ora come unità di uscita "vincitrice" quella in cui la "distanza" precedentemente calcolata assume il minimo valore. L'algoritmo di Kohonen adotta la strategia della "bolla", ciò significa che non vi è una sola unità "vincitrice", ma si considerano "vincitrici" (si recuperano, cioè) anche tutte le altre unità di uscita che si trovano entro una certo "intorno" dall'unità "vincitrice" principale. Tale "intorno", nella terminologia di Kohonen, è detto *raggio della bolla*.

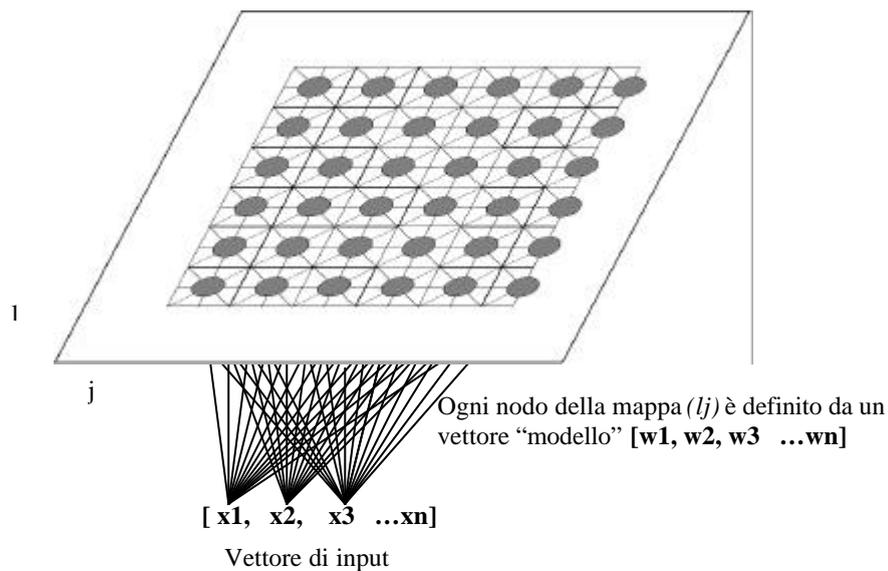


Figura 3. SOM 6 x 6.

Dopo aver individuato le unità "vincitrici", si modificano solo i coefficienti di connessione delle linee che fanno capo a queste unità, in modo da diminuire la "distanza" tra essi ed il segnale presentato in ingresso. Si passa poi al vettore di input successivo. Man mano che il processo di apprendimento va avanti il "raggio" della bolla ed il tasso di modifica dei coefficienti vengono fatti progressivamente decrescere, con una legge definita dallo sperimentatore. Quando quest'ultimo decide di interrompere il processo di

apprendimento, se i patterns di ingresso presentavano al proprio interno qualche regolarità statistica, si può osservare che si è formata spontaneamente una differenziazione nei ruoli delle unità di uscita, nel senso che ognuna di esse risponde (cioè diventa “vincitrice”) solo in corrispondenza a particolari classi di pattern di ingresso.

Quanto detto finora può tornare utile per avere una panoramica ed una visione d’insieme del carattere delle SOMs e del loro funzionamento. Si vuole ora entrare più nel dettaglio, presentando magari degli esempi di applicazione ed implementazione delle SOMs a problemi di *clustering*. Al contempo si forniranno informazioni di carattere più tecnico riguardante l’algoritmo di apprendimento e di definizione della “bolla”.

3.1. Una Mappa di vettori binari

Consideriamo ora una semplice caso di mappatura tramite SOM di vettori binari considerati input della rete.

La Tabella 1 sintetizza la descrizione di 16 animali definiti in base l’assenza o la presenza di 13 differenti attributi.

		d	d	g	o	h	e	a	w	t	l	h	z
		o	h	u	o	o	a	g	f	d	o	c	g
		v	e	c	s	w	w	l	o	o	l	a	e
		e	n	k	e	l	k	e	x	g	f	t	r
		n	e	a	w								
is	small	1	1	1	1	1	1	0	0	0	0	1	0
	medium	0	0	0	0	0	0	1	1	1	1	0	0
	big	0	0	0	0	0	0	0	0	0	0	1	1
has	2 legs	1	1	1	1	1	1	1	0	0	0	0	0
	4 legs	0	0	0	0	0	0	0	1	1	1	1	1
	hair	0	0	0	0	0	0	0	1	1	1	1	1
	hooves	0	0	0	0	0	0	0	0	0	0	0	1
	mane	0	0	0	0	0	0	0	0	0	0	0	1
	feathers	1	1	1	1	1	1	1	0	0	0	0	0
likes to	hunt	0	0	0	0	1	1	1	1	0	1	1	1
	run	0	0	0	0	0	0	0	0	1	1	1	0
	fly	1	0	1	1	1	1	1	0	0	0	0	0
	swim	0	0	1	1	0	0	0	0	0	0	0	0

Tabella 1. Animali e loro attributi codificati in maniera binaria (tratto da Kohonen, 1995: 164).

Tutti i membri del data set sono stati presentati come input, in maniera iterativa, ad una SOM di 10 x 10 neuroni soggetta al processo di adattamento (come si è accennato in precedenza, l’adattamento spaziale ai pattern d’ingresso è il compito primario di una rete SOM). L’inizializzazione della mappa è stata definita in maniera *random*. Dopo 2000 cicli di presentazioni, in cui il raggio della “bolla” e l’indice di apprendimento sono fatti

decrescere, ogni neurone diventa più o meno rappresentativo di una delle combinazioni dei 13 attributi e quindi di uno dei 16 animali. La mappa ottenuta è mostrata in Figura 4.

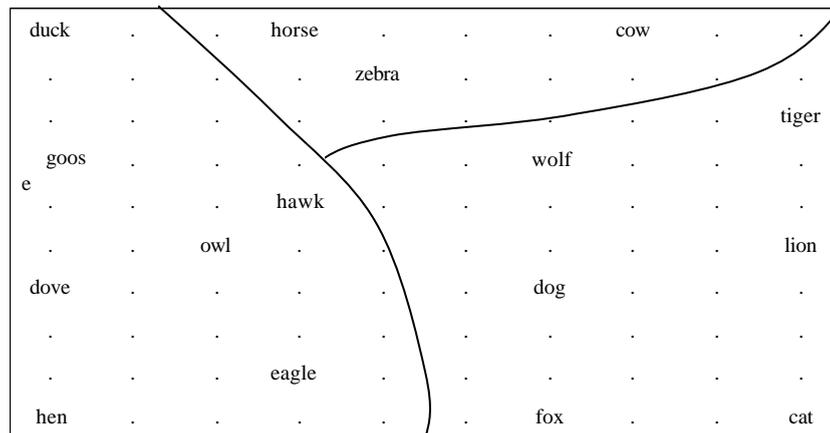


Figura 4. Una SOM 10 x 10 addestrata tramite il data set della Tabella 1. All'interno della mappa è possibile discernere dei clusters formatisi in base al rapporto di similarità dei pattern di ingresso (tratto da Kohonen, 1995: 164).

La mappa mostra un ordine spaziale che in un certo modo ha catturato le “relazioni di famiglia” essenziali tra i diversi animali. Ad esempio le unità corrispondenti alla famiglia “uccelli” occupano la parte sinistra della mappa, i “predatori” sono categorizzati verso destra e gli animali più pacifici come la zebra, il cavallo e la mucca si trovano, invece, in alto al centro. All'interno di ogni classe sono possibili ulteriori sottocategorie basati sul rapporto di similarità tra coppie e terne di animali.

Dall'esempio appena trattato, si evince che la SOM provvede ad una *quantizzazione vettoriale*, nel senso che il valore di ingresso fornito alla rete è uno spazio continuo e deve essere discretizzato su una superficie bi-dimensionale. Se, quindi, gli ingressi della rete sono vettori n -dimensionali \mathbf{x} , la SOM opererà una rappresentazione spaziale discreta e puntuale di tali dati multi-dimensionali, organizzandoli in raggruppamenti omogenei con dimensionalità inferiore a quella d'inizio che tuttavia presentano una coerenza topologica con i dati di ingresso in base ai quali tali raggruppamenti sono stati realizzati.

La rete deve fornire una rappresentazione ottimale della discretizzazione del *continuum* spaziale dei dati in input limitatamente però alla sua struttura fisica. Reti bi-dimensionali permetteranno solo una visualizzazione della mappatura lungo due coordinate, reti tridimensionali rappresenteranno il vettore discretizzato lungo tre coordinate e così via⁹.

⁹ E' intuitivo capire che dei problemi di visualizzazione dei dati emergeranno qualora la mappa avrebbe numero dimensionale pari a 4 o superiore. Per tal motivo il presente testo farà riferimento solamente a mappe bi-dimensionali affinché l'indagine visiva dei dati rimanga agile e confortevole.

3.2. L'algoritmo di addestramento

Il set degli esempi di input è descritto da un vettore $\mathbf{x}(t) \in R^n$ dove t indica una discretizzazione temporale. Ogni nodo i della rete SOM contiene un vettore “modello” $\mathbf{m}_i(t) \in R^n$, che ha lo stesso numero di elementi del vettore di input $\mathbf{x}(t)$.

L'algoritmo della SOM raggiunge uno stato di equilibrio, creando quindi degli ordini spaziali (*clusters*), tramite la ripetizione dei seguenti passaggi:

1. Ogni vettore di input $\mathbf{x}(t)$ è confrontato con tutti i vettori “modello” $\mathbf{m}_i(t)$ al fine di trovare la BMU (*best-matching unit*), l'unità cioè che presenta maggiori similarità (in termini di distanza euclidea) con il vettore di input. Come si è visto in precedenza a proposito del *competitive learning*, il BMU è spesso chiamata l'unità “vincitrice” e si indica con \mathbf{c} .
2. Ogni vettore “modello” delle unità “vincitrici” e un numero di unità circostanti \mathbf{c} (definiti la già conosciuta nozione di “bolla”) sono adattati¹⁰ al vettore di ingresso secondo una regola di apprendimento che adesso andremo ad esporre.

Come già ripetuto, l'idea fondamentale nel “SOM learning process” è che per ogni vettore di input $\mathbf{x}(t)$ (ordinato in un apposito *training set*), il raggio della “bolla” decresce gradualmente (fase di apprendimento) e l'unità “vincitrice” selezionata subisce una modifica (vettoriale) al fine di adattarsi il più possibile a $\mathbf{x}(t)$, il vettore spaziale di ingresso.

Durante il processo di apprendimento possono verificarsi contraddizioni individuali, ma nel complesso, il risultato della rete è l'emergere di valori ordinati per $\mathbf{m}_i(t)$ localizzati sulla mappa.

Se il numero degli esempi di input disponibili è basso, essi devono essere presentati alla rete in maniera re-iterativa.

La definizione di \mathbf{c} (l'unità “vincitrice”) è data dalla differenza della distanza euclidea tra il vettore d'ingresso ed il vettore “modello”, ottenuta grazie all'espressione:

$$\mathbf{c} : \mathbf{m}_c(t) = \min \|\mathbf{x}(t) - \mathbf{m}_i(t)\| \quad (1)$$

L'adattamento è implementato come una graduale riduzione della differenza tra i componenti del vettore di input e i componenti del vettore “modello”. L'algoritmo è espresso nel seguente modo:

$$\mathbf{m}_i(t + 1) = \mathbf{m}_i(t) + \alpha(t)[\mathbf{x}(t) - \mathbf{m}_i(t)] \quad \text{per ogni } i \in N_c(t), \quad (2)$$

dove t rappresenta la discretizzazione temporale, il parametro $\alpha(t) \in [0, 1]$ è uno scalare, detto anche *indice di apprendimento* e decresce durante le fasi dell'apprendimento.

¹⁰ Nei termini dei propri coefficienti di connessione, si intende.

La grandezza della “bolla”, per il recupero e l’aggiornamento delle unità circostanti c , è definito da $N_c(t)$ e anch’esso decresce durante l’adattamento.

Uno degli aspetti più interessanti della regola di apprendimento è che essa contiene due parametri variabili col tempo: la grandezza definente la “bolla”, N_c , e l’indice di apprendimento, α .

Come si è detto, generalmente essi vengono fatti diminuire col tempo anche se in alcune applicazioni α viene lasciato costante. Solitamente la grandezza della “bolla” è sintetizzata dal suo raggio $r(t)$. Esempi di leggi di variazione del raggio della “bolla” con il tempo sono:

$$r(t) = A - B \times t$$

in cui il valore tipico di A è 4 e di B è 0.0002, oppure:

$$r(t) = A/(1 + B \times t)$$

in cui i valori tipici sono : A = 4, B = 0.003.

Leggi esattamente analoghe a queste possono essere utilizzate per far diminuire col tempo l’indice $\alpha(t)$. In sostanza esse servono a far sì che, nella fase iniziale dell’apprendimento, vengano interessati dal cambiamento molti coefficienti di connessione, i quali subiscono variazioni abbastanza di rilievo (valori di α e $r(t)$ elevati), mentre, al termine del processo di apprendimento, quando ormai ogni unità si è specializzata nel rispondere a particolari pattern in ingresso, conviene che sia il raggio della “bolla” che l’indice di apprendimento siano molto modesti.

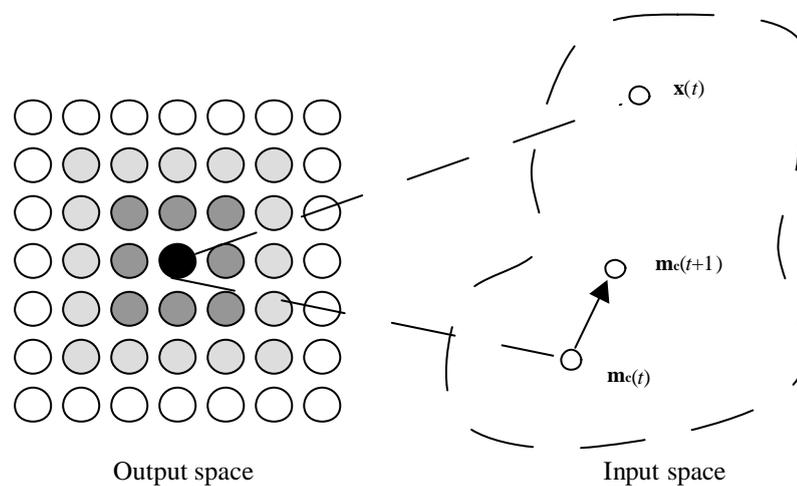


Figura 5. Istantanea della fase di adattamento tra il vettore d’ingresso, $x(t)$ e quello della mappa $m_c(t)$.

La Figura 5 mostra una rappresentazione grafica di una *Self-Organizing Map* durante il processo di adattamento. Il cerchietto nero indica che l'unità è stata selezionata come unità "vincitrice" per la rappresentazione del pattern di input $\mathbf{x}(t)$. Il vettore dell'unità vincente, $\mathbf{m}_c(t)$, è adattato (durante la fase di apprendimento) al pattern di input e quindi $\mathbf{m}_c(t + 1)$ sarà più vicino al vettore input $\mathbf{x}(t)$ di quanto lo era $\mathbf{m}_c(t)$. Tale processo di adattamento, sappiamo, coinvolge anche un numero di unità prossime all'unità "vincitrice". In Figura 5 queste unità sono rappresentate mediante una scala di tonalità di grigio. Maggiore è l'intensità di grigio, maggiormente queste unità saranno coinvolte nel processo di adattamento (apprendimento).

3.3. Aspetti multipli di una SOM

In questo paragrafo saranno considerate le proprietà e potenzialità di una SOM per le sue possibili applicazioni. Le considerazioni qui sviluppate avranno carattere propedeutico a quanto verrà detto in seguito, ma allo stesso tempo sono da considerarsi sunto di quanto detto in precedenza.

1) La SOM può essere ritenuta un *modello di apprendimento non-supervisionato*. Come tale risponde bene alla risoluzione di problemi quali:

- **Clustering**. Fornito un insieme di elementi in ingresso, questi devono poter essere raggruppati in *clusters* (gruppi omogenei) ed il compito della rete, dopo l'apprendimento, è individuare, per ogni ingresso, il gruppo di appartenenza.
- **Quantizzazione vettoriale**. Il valore di ingresso fornito alla rete è uno spazio continuo che deve essere discretizzato. Gli ingressi della rete sono vettori n -dimensionali \mathbf{x} , l'uscita è una rappresentazione discreta dello spazio d'ingresso. La rete deve fornire una rappresentazione ottimale della discretizzazione dello spazio d'ingresso.
- **Estrazioni delle caratteristiche**. La SOM estrae delle caratteristiche dall'ingresso iniziale, questo spesso implica una riduzione della dimensionalità.

2) La SOM fornisce una "adaptive and contextual knowledge representation". I formalismi tradizionali per la rappresentazione della conoscenza – reti semantiche, logica dei predicati, *frames*, eccetera- rimangono fissi nella loro staticità e la loro relazione di referenza dei loro elementi è di natura arbitraria e prevede l'intervento umano. Una rete SOM non solo è da considerarsi come modellizzatore *fuzzy*, permette rappresentazioni più granulari all'interno delle categorie, ma le categorie stesse sono create *in situ* durante il processo di apprendimento e successivamente sottoposte a *gradient modelling*.

3) La SOM è usata spesso utilizzata come *dispositivo per analisi statistiche*. Una rete SOM, infatti, costituisce un metodo di riduzione e proiezione delle dimensionalità, capace di mappare (e categorizzare) "high-dimensional data space into low-dimensional data

space” riuscendo a preservare però le proprietà dinamiche e topologiche dello spazio iniziale.

4. Modelli psicolinguistici per l'apprendimento della morfologia

Nella letteratura psicolinguistica vengono tradizionalmente riconosciuti due meccanismi di base per la rappresentazione e l'apprendimento dell'informazione morfologica da parte del parlante. Il primo presuppone l'esistenza di un sistema di *regole generative* (Taft, 1988), attraverso le quali le forme flesse di una lingua sono analizzate o prodotte in tempo reale (e cioè al momento del bisogno) sulla base di costituenti morfologici minimi (i morfemi). Il secondo meccanismo si richiama alla nozione di *rete associativa* (Bybee, 1995) lessicale, attraverso la quale le forme flesse memorizzate nel lessico mentale sono messe in relazione tra loro, in modo tale da garantire procedure di accesso all'informazione lessicale efficienti e robuste, e la possibilità di interpretare e produrre parole nuove sulla base delle parole già acquisite.

Nel primo caso, alle regole è demandato il duplice compito di rappresentare generalizzazioni morfologiche in modo simbolico e alleggerire il lessico mentale del parlante dal carico informativo rappresentato da forme flesse regolari e prevedibili. Infatti, secondo questo modello, nel lessico mentale vengono memorizzate in modo permanente a) le basi lessicali (generalmente le cosiddette radici) sulle quali le regole operano e b) quelle forme flesse idiosincratiche che non possono essere prodotte dalle regole stesse (ad esempio una forma irregolare come è). Categorie morfo-sintattiche quali persona, numero, tempo e modo sono generate per concatenazione di uno o più suffissi, secondo i dettami della morfologia cosiddetta “Item and Process” (Hockett, 1954), oppure sono listate come attributi delle entrate lessicali relative agli stessi suffissi, che entrano a far parte del lessico, secondo l'approccio morfologico denominato “Item and Arrangement” (Hockett, 1954; Jensen, 1990), a tutti gli effetti. Altre informazioni, come quella relativa alla classe di coniugazione di un verbo, sono associate direttamente alla radice del verbo in questione.

Nel quadro del secondo meccanismo di apprendimento, le regolarità nella morfologia di una lingua sono rappresentate come *schemi emergenti di associazione parallela di forma e contenuto tra parole flesse memorizzate nel lessico mentale dal parlante*. La funzione di questi schemi è di produrre ed interpretare forme che il parlante non ha (ancora) memorizzato. La produttività di uno schema può essere espressa come una funzione diretta della varietà degli esempi memorizzati attestanti lo schema.

Si noti che i due meccanismi non sono incompatibili, dal momento che ricorrono a modalità complementari di apprendimento e di organizzazione dell'informazione morfologica. Esiste tuttavia in letteratura una tendenza diffusa ad assegnare alle regole la funzione di modellare regolarità linguistiche frequenti e produttive, e a riservare l'uso di schemi lessicali associativi al trattamento di regolarità minori, rare e non produttive (Pinker & Prince, 1988; Bertinetto, 1999). D'ora in avanti chiameremo *modello duale* un modello ibrido di apprendimento che si avvalga di entrambe le modalità di generalizzazione. Una variante del modello duale è contenuta nella proposta avanzata da Dressler (1985): alternanze minoritarie e sotto-regolari riceverebbero una rappresentazione mentale sia in forma di regole che di schemi associativi lessicali. Questo modello si avvale dunque di un meccanismo *ternario* di apprendimento.

Esiste infine un'ipotesi *monistica* dell'apprendimento morfologico (Vennemann, 1974; Rumelhart & McClelland 1986; Bybee, 1988), secondo la quale non c'è bisogno di assumere che processi regolari, sotto-regolarità morfologiche ed eccezioni siano appresi, prodotti e percepiti secondo modalità cognitivamente distinte. I meccanismi di associazione lessicale possono, da soli, prendersi carico di tutte e tre le tipologie di fenomeni in questione. Questo perché la differenza tra una regola e un'eccezione non è categoriale, ma risiede nella diversa forza associativa degli schemi lessicali che le rappresentano nel lessico mentale del parlante. Una regola è uno schema associativo il cui dominio include una grande varietà di parole diverse tra loro. Una sotto-regolarità ha semplicemente un dominio di associazioni più ristretto. Un'eccezione assoluta è uno schema "degenere", non riconducibile a nessun'altra parola attestata.

La nozione di paradigma riecheggia in molti di questi modelli. Nell'ambito dell'ipotesi monistica dell'apprendimento morfologico, il paradigma sembra rivestire il ruolo di principio organizzatore delle famiglie di parole memorizzate in una rete associativa. Tutte le forme flesse associate ad una stessa cella del paradigma (ad es. *amiamo, siamo, contiamo, veniamo* ecc.) dovrebbero tendere a disporsi intorno ad un attrattore comune, riproducendo a livello topologico la struttura di un macro-paradigma o schema flessionale. Sembra tuttavia difficile stabilire una connessione diretta tra vincoli strutturali di tipo paradigmatico e il modo in cui le forme acquisite dal parlante si localizzano topologicamente in aree contigue della memoria. L'ordine di scala tra i due fenomeni non è, infatti, paragonabile. Da una parte, le associazioni tra neuroni adiacenti che danno vita a fenomeni di attrazione reciproca rispondono tipicamente a principi di riorganizzazione locale: neuroni che presentano pattern di attivazione simili tendono a disporsi intorno ad attrattori comuni. D'altra parte, il paradigma sembra rispondere a principi di minimizzazione globale dei raggruppamenti prodotti.

Anche gli approcci a regole fanno implicitamente appello alla nozione di paradigma. Come vedremo più nel dettaglio in seguito, il punto principale di contatto è dato dal requisito meta-grammaticale di adeguatezza esplicativa, secondo il quale il numero di morfemi individuati e il numero di regole acquisite devono rispondere al principio di massima economia descrittiva. Goldsmith (2001) dimostra che il paradigma si presta a rispondere a requisiti meta-grammaticali di adeguatezza descrittiva come il "minimum description length".

5. Modelli computazionali per l'apprendimento della morfologia

Il problema logico dell'acquisizione della conoscenza morfologica da parte del parlante è stato tradotto, in vari ambiti disciplinari e con diversi obiettivi sia pratici che teorici, in procedure algoritmiche valutabili empiricamente simulando al computer una o più sessioni di apprendimento. In tutti i casi, si parte dall'analisi di una serie di forme che costituiscono il *campione di addestramento* o corpus di input della procedura. Le forme in questione possono essere corredate dall'insieme di tratti morfologici a loro associati (ad es. 1s del presente indicativo), oppure no. Nel primo caso si parla comunemente di addestramento supervisionato, e di addestramento non supervisionato nel secondo caso.

L'obiettivo della fase di addestramento varia secondo il tipo di approccio al problema. La procedura di acquisizione, per alcuni studiosi, deve essere in grado di fornire, alla fine del periodo di addestramento, un'analisi morfologica corretta delle forme attestate nel

corpus. L'analisi consiste generalmente nell'individuazione dei confini di morfema tra i costituenti morfologici delle forme in questione (Harris, 1968; Hafer & Weiss, 1974; van den Bosh, 1997), oppure nel loro raggruppamento in classi morfologicamente coerenti (Skousen, 1989), o ancora in schemi (o regole) di relazione morfo-fonologica tra forme. Secondo altri studiosi, invece, l'obiettivo dell'apprendimento morfologico risiede nell'acquisizione della capacità di generare forme flesse a partire da una base lessicale nota (Rumelhart & McClelland, 1986; Ling, 1994).

Un altro parametro di variazione è dato dai criteri di valutazione della bontà dei risultati ottenuti. La stragrande maggioranza degli approcci al problema mette in campo criteri interni di valutazione, relativi ad esempio all'accuratezza della ripartizione in classi morfologiche naturali o al numero di forme corrette automaticamente generate dalla procedura. La misura del successo dell'algoritmo è data solamente dal risultato finale, non dal modo in cui sia stato conseguito, o da quanto rapidamente è stato raggiunto.

In altri casi, invece, gli studiosi hanno fatto appello a criteri meta-grammaticali di "adeguatezza descrittiva", che impongono una valutazione globale della risposta della procedura nei termini del numero delle classi o dei morfemi individuati, o ancora delle regole morfo-fonologiche estratte (Kazakov, 1997; Goldsmith, 2001). In questi casi, l'interesse dello studioso risiede soprattutto nel tentativo di emulare le "procedure di scoperta" ('discovery procedures') che portano un linguista a preferire una certa analisi morfologica rispetto ad un'altra, quando quest'ultima risulti "debolmente equivalente" ma meno compatta (o economica) rispetto alla prima.

In altri ambiti, lo scopo principale della simulazione al computer è di riprodurre il comportamento linguistico del bambino nelle sue fasi evolutive, piuttosto che la metodologia descrittiva del linguista. L'enfasi, in questo caso, è posta sull'aspetto *incrementale* del problema cognitivo: tipicamente, il bambino non aspetta che l'evidenza attestata gli consenta di convergere su una generalizzazione sicuramente corretta, ma tenta comunque, al bisogno, di generare una forma sconosciuta, indipendentemente dal suo stadio di maturazione linguistica e dalle probabilità di generare una forma erronea.

5.1. Paradigmi morfologici ed emergentismo

All'interno del modello monistico dell'apprendimento morfologico, si può ragionevolmente supporre che le celle del paradigma svolgano il ruolo di attrattori tra forme flesse con lo stesso contenuto morfo-sintattico. Forme che occupano la stessa posizione nel paradigma tenderebbero ad essere memorizzate in aree contigue di memoria. In un modello di memoria associativa (Kohonen 1986 e 1995), la topologia della rete finirebbe così col riprodurre la struttura astratta di uno schema flessionale. Nel quadro di un approccio duale all'apprendimento morfologico, d'altra parte, possiamo ragionevolmente immaginare che la struttura del paradigma imponga un vincolo globale sul numero complessivo dei morfemi da acquisire e delle regole per la loro combinazione (Carstairs, 1987), riducendo considerevolmente lo spazio delle ipotesi, e minimizzando sia le dimensioni del lessico dei morfemi che il carico dell'analisi morfologica in termini di elaborazione ("processing").

Se queste conclusioni sembrano evidenziare aspetti non banali del problema logico dell'apprendimento della morfologia, d'altra parte non è chiaro che cosa ci possano insegnare sul piano strettamente cognitivo. Dobbiamo immaginare che il bambino si

comporti come un “piccolo linguista”, ragionando in termini di criteri globali di minimizzazione del repertorio di unità di base (morfemi) e di massima economia descrittiva della grammatica?

La risposta non può che essere scettica: le cosiddette “procedure di scoperta” dei linguisti non si prestano ad essere tradotte meccanicamente in processi di acquisizione plausibili sul piano cognitivo. E’ difficile infatti immaginare, sulla base di quanto sappiamo della maturazione linguistica del bambino, che quest’ultimo riesca ad adattare una strategia “top down” di ottimizzazione globale ad una serie incrementale di fasi di addestramento incompleto (all’interno delle quali, cioè, sono presentate solo alcune forme di alcuni lessemi). Questa strategia dovrebbe comportare ripetute ristrutturazioni complessive della conoscenza acquisita all’istante t rispetto all’organizzazione della stessa conoscenza all’istante $t-1$, secondo una dinamica tipicamente discontinua.

Le nostre conoscenze attuali sull’insorgenza di strutture complesse a livello cerebrale evidenziano, al contrario, la natura graduale e incrementale di questi processi. Se i processi di organizzazione dinamica della conoscenza linguistica fossero governati da principi di minimizzazione globale delle unità di base, la trasmissione della morfologia di una lingua da una generazione alla successiva tenderebbe ad una riduzione sistematica del repertorio delle unità di base (o morfemi), unitamente ad altri effetti unidirezionali ed irreversibili di semplificazione quali, ad esempio, l’estensione generalizzata di un processo fonologico a tutto l’insieme delle forme di un dato paradigma. Anche sul piano del cambiamento linguistico diacronico, tuttavia, l’evidenza in nostro possesso sembra deporre a favore di modelli emergenti di ristrutturazione selettiva, con fasi di oscillazione e repentine transizioni tra dinamiche continue. Certamente, nozioni quali la frequenza con la quale un bambino è esposto a certe forme flesse piuttosto che ad altre gioca un ruolo importante in questi processi emergenti (Bybee & Hopper, 2000). Tuttavia, lo ripetiamo, non sembra plausibile, sul piano cognitivo, rendere conto di queste interazioni nei termini di funzioni di ottimizzazione globale.

Dobbiamo a questo punto confrontarci con un apparente paradosso. Da una parte sappiamo che se i paradigmi non esistessero, la morfologia sarebbe molto più difficile da apprendere. D’altra parte, tuttavia, non sembra possibile, ad oggi, vedere nei paradigmi il prodotto spontaneo del modo in cui le forme si auto-organizzano nella memoria del parlante. La prospettiva di organizzazione globale imposta dal paradigma mal si concilia con un approccio incrementale all’apprendimento.

Le procedure di acquisizione automatica finora considerate non affrontano questo paradosso. In tutti i modelli da noi considerati nei quali svolge un ruolo diretto o indiretto, il paradigma costituisce un principio di organizzazione linguistica astratta che il processo di apprendimento conosce a priori e tende a replicare. In nessun caso ci è dato di osservare un processo di auto-organizzazione paradigmatica delle forme flesse a partire da dati sparsi e incompleti. Goldsmith (2001) modella il principio noto come “minimum description length” con l’ausilio di schemi flessionali (“signature”), che ricordano molto da vicino la nozione di paradigma. Analogamente, Gaussier (1999) cerca di replicare l’organizzazione paradigmatica del dato morfologico ricorrendo a procedure di *clustering* automatico. Pirrelli & Yvon (1999) fondano a loro volta la nozione di paradigma sull’analogia proporzionale, ma non è chiaro in che misura quest’ultima possa ritenersi una proprietà inerente al dato.

Il paradigma teorico dell’emergentismo offre, a nostro avviso, una chiave di lettura alternativa. Molti fenomeni linguistici “complessi” possono essere analizzati come il

prodotto di una dinamica temporale, vale a dire come il risultato dell'interazione e della riorganizzazione attraverso il tempo di unità linguistiche semplici. Ad esempio, è stato suggerito che il bambino memorizza le forme flesse di una lingua come nodi di una rete associativa. Le relazioni tra nodi esercitano una forza strutturale che ne facilita la memorizzazione, e consente al tempo stesso l'interpretazione e la produzione di forme sconosciute. Una caratteristica fondamentale di questi processi di strutturazione dinamica è che essi non rispondono a regole o criteri di ottimizzazione e strutturazione globali, ma obbediscono esclusivamente a principi di interazione locale i cui effetti si ripercuotono tuttavia a livello dell'intera struttura dei dati. Proprietà strutturali con queste caratteristiche vengono definite in letteratura come "emergenti". Nelle pagine che seguono cercheremo di definire una dinamica di apprendimento che possa rendere conto della graduale, spontanea insorgenza di strutture paradigmatiche a partire da famiglie sparse di parole non correlate sul piano lessicale. La simulazione al computer di questa dinamica costituirebbe, a nostro avviso, un passo in avanti verso la comprensione dei meccanismi mentali di auto-organizzazione del dato morfologico e del loro possibile effetto globale sulla strutturazione dell'informazione morfologica nel repertorio di forme flesse memorizzate dal parlante.

7. SOM e Morfologia

L'intenzione del presente lavoro è quella di utilizzare un sistema auto-organizzante, come la SOM, al fine di classificare un *insieme* di dati,¹¹ costituito da voci verbali flesse, codificato e strutturato per il processo di addestramento. In particolare, il nostro obiettivo primario consisterà nell'evidenziare l'insorgenza di principi di organizzazione globale dell'informazione morfologica a partire da routine incrementali di apprendimento basate su un'evidenza e su procedure di generalizzazione puramente locali.

Per una SOM, acquisire automaticamente informazione morfologica da un repertorio non supervisionato di forme flesse significa assegnare a ciascuna forma le sue coordinate su una mappa (generalmente un piano bi-dimensionale). La mappa è una rete di neuroni, rappresentati come vettori multi-dimensionali. Forme adiacenti sulla mappa corrispondono a vettori che presentano schemi di attivazione simili, cioè valori simili in corrispondenza delle stesse dimensioni. La presenza di aree densamente popolate da forme, separate tra loro da zone di confine relativamente vuote sta ad indicare che il repertorio di addestramento si ripartisce in classi morfologiche naturali.

Dato l'alto numero di dimensioni di cui si compongono generalmente i vettori, la prossimità sulla mappa bi-dimensionale non può rendere conto della loro similarità globale, ma rifletterà piuttosto un punto di vista emergente dall'apprendimento: vettori contigui saranno simili relativamente ad un sottoinsieme delle loro dimensioni, e cioè sotto certi rispetti, considerati più rilevanti di altri. Dal nostro punto di vista, uno degli aspetti di maggior interesse nell'uso delle SOM in morfologia è osservare il modo in cui la mappa arriva autonomamente a sviluppare una nozione di rilevanza morfologica.

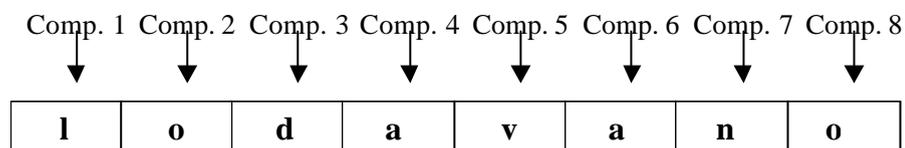
¹¹ Nel corso di queste pagine faremo riferimento in maniera indistinta a termini quali "data set", "corpus", "training set", "campione d'addestramento". A dispetto della terminologia usata, l'accezione di riferimento, in tutti questi casi, rimarrà sempre la stessa e indicherà sostanzialmente la collezione (o insieme) di dati morfologici utilizzati come base primaria di conoscenza per l'addestramento della SOM.

7.1. Alla ricerca di una codifica morfologica

Non esiste un codice morfologico per mezzo del quale codificare agilmente e rappresentare vettorialmente i dati di input alla rete. Così come non esiste una scala di “morfologicità” per mezzo della quale è possibile stabilire dei gradi discreti di vicinanza o lontananza morfologica tra due o più forme verbali.

Una possibile soluzione è considerare la forma verbale in tutta la sua lunghezza al pari di una sequenza di componenti. Facciamo un esempio: si consideri la forma verbale “lodavano”.

lodavano



Così strutturata, la forma “lodavano” è codificata attraverso otto componenti, ognuna corrispondente ad ogni singola lettera. Una codifica del genere enfatizza, com’è ovvio, l’ordine sequenziale delle lettere e quindi considera la forma verbale come una sequenza di componenti di numero eguale al numero di lettere che la compongono¹².

Tale codifica (che possiamo chiamare codifica “da sinistra a destra”) risulta abbastanza grezza, e sollecita una serie di osservazioni. In particolare, il criterio di considerare le voci verbali al pari di stringhe di variabili corrispondenti alle sequenze di lettere, appare parecchio efficace per evidenziare similitudini ed uguaglianze di porzioni morfologiche che si localizzano nella parte sinistra della forma verbale, in quella parte, cioè, che coinciderebbe, in termini linguistici, con la *radice verbale*. Proponiamo un nuovo esempio: consideriamo una codifica “da sinistra a destra” per quattro voci verbali: “lodavano”, “amante”, “lodasti” e “amavano”.

¹² Per adesso la nostra attenzione si sta focalizzando sulla struttura della matrice dei dati di input. Si provvederà più avanti ad assegnare a tale struttura un’attribuzione numerica-vettoriale.

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5	Comp. 6	Comp. 7	Comp. 8
lodavano	l	o	d	a	v	a	n	o
amante	a	m	a	n	t	e	x	x
lodasti	l	o	d	a	s	t	i	x
amavano	a	m	a	v	a	n	o	x

Una mappa addestrata con una codifica così strutturata, coglierà in ogni caso le regolarità statistiche che riguardano la parte iniziale delle stringhe di lettere, quella generalmente invariabile (la radice). Si noti tuttavia che la codifica vettoriale non è in grado di evidenziare il fatto che due parole hanno la stessa terminazione, a meno che esse non abbiano anche la stessa lunghezza. Nel nostro esempio, *lodavano* e *amavano*, pur avendo entrambe la stessa terminazione flessionale, non presentano gli stessi valori per le stesse dimensioni del vettore, in quanto il suffisso *-vano* si colloca nei due casi in due diverse posizioni a partire dal primo elemento del vettore. Le ultime quattro lettere della voce “*lodavano*” occupano rispettivamente le componenti in posizione 5, 6, 7 e 8. Le medesime ultime quattro lettere nella voce “*amavano*” sono invece definite rispettivamente dalle posizioni 4, 5, 6 e 7.

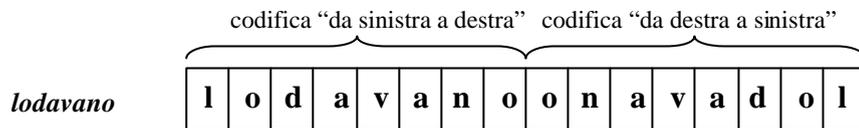
Questo vuol dire che “*lodavano*” ed “*amavano*”, che pur presentano quattro variabili in comune (le ultime quattro lettere, cioè) verranno trattati dalla mappa in maniera distinta e separata, senza recuperare quanto di morfologicamente rilevante le due forme condividono.

In altre parole una codifica siffatta ignora quella parte delle forme verbali più soggetta a variabilità nel corso della flessione, generalmente nota sotto il nome di *desinenza*. Se pur con gravi difetti la codifica “da destra a sinistra” presenta una sua legittimità linguistica ed epistemologica. Si tratta infatti di una pura trascrizione della voce verbale all’interno della struttura “a griglia” della matrice di dati.

Nella migliore delle ipotesi dovrebbe essere la SOM, infatti, ad attribuire una pertinentizzazione categoriale ai dati codificati, pertinentizzazione che dovrebbe riflettere una qualche similarità (morfologica, si spera) soggiacente alle forme verbali considerate input d’ingresso della rete.

Non deve essere lo sperimentatore, ma la rete a dare salienza e capacità organizzativa all’insieme dei dati tutto. Vale la pena dunque affinare e cercare di complementare la codifica “da destra a sinistra” in modo tale che si recuperino le similarità provenienti anche dalla parte destra della voce verbale.

Una buona soluzione a riguardo sembra essere l’integrazione della codifica “da sinistra a destra” con una codifica di eguale natura ma stavolta ordinata da destra a sinistra. Riproponiamo la voce “*lodavano*” sotto questa nuova codifica.



La forma verbale “lodavano” appare ora gestita da sedici componenti anziché da otto.

Così facendo cerchiamo di ottenere un compromesso tra un recupero di similarità proveniente dalla parte sinistra della stringa (la radice), e quello proveniente dalla parte destra della stringa e cioè la desinenza.

Tale nuova sintassi di codifica è motivata quindi da esigenze pratiche ed anche tecniche, ma la linea teorica che la regola è identica a quella usata per la codifica “da sinistra a destra”: operare una trascrizione *naif* del dato di input, lasciando intatta la sua natura morfologica, preservandola da allettanti processi esogeni di segmentazione ed analisi in genere.

Proviamo adesso, utilizzando la nuova codifica complementare “da sinistra a destra” e “da destra a sinistra”, a confrontare insieme le già considerate quattro forme verbali: “lodavano”, “amante”, “lodasti” e “amavano”.



Il nuovo ordine di codifica presenta, come si vede, anche il vantaggio di considerare la desinenza come parte rilevante per il recupero di analogie morfologiche tra le varie forme verbali. In questo modo la porzione morfologica *-vano* (presente nella prima e nell’ultima voce dell’esempio qui riportato) viene presa in considerazione e resa comune alle due forme verbali¹³.

La codifica complementare costituirà la “codifica ufficiale” per una serie di esperimenti che ci accingiamo a condurre.

¹³ Non secondariamente, c’è da dire che la codifica complementare, pur permettendo il recupero di similarità tra desinenze e basi di forme verbali, aumenta notevolmente il “rumore computazionale” introducendo nuove variabili per le quali vale il medesimo problema della coincidenza tra la posizione della lettera ed il numero della variabile.

7.1.1. Considerazioni sulla codifica vettoriale

La scelta della componente numerica che sostituirà ogni singola lettera è stata nel nostro caso parecchio semplice e di facile applicazione. Per la verità, diversi criteri di scelta potevano essere adottati: riscrivere ciascuna lettera utilizzando l'estrazione random di ventuno valori presi su un segmento di superficie uniforme, oppure utilizzare un valore fonetico per caratterizzare ciascuna lettera ed altri criteri ancora.

La nostra preferenza si è indirizzata sulla frequenza della singola lettera all'interno del corpus considerato. Si è creduto infatti che la frequenza possa riflettere, in un qualche modo, la funzionalità della singola lettera per la composizione morfologica dei verbi tutti. Si è proceduto, quindi, rilevando la frequenza assoluta di ogni lettera all'interno del corpus.

Come appena detto, ciascun vettore contiene una transcodifica numerica della sequenza di caratteri che compongono la forma associata a quel vettore. La prossimità di due vettori nello spazio della mappa indicherà che le forme a loro associate hanno in comune un certo numero di caratteri nello stesso ordine. La nozione di rilevanza morfologica che la mappa è in grado di acquisire a partire da questa codifica è dunque relativa soltanto al modo in cui le forme in questione si realizzano superficialmente a livello ortografico.

E' chiaro che solo alcune generalizzazioni morfologiche possono essere espresse a questo livello di descrizione. Ad esempio, potremo con buona probabilità riconoscere che *cantiamo* e *corriamo* sono significativamente più simili tra loro che, mettiamo, *cantiamo* e *corro*. Tuttavia, non è affatto ovvio, dal punto di vista di una SOM addestrata con codifiche numeriche del tipo appena illustrato, che due forme flesse per lo stesso tempo e modo quali *canto* e *cantiamo* siano significativamente più vicine tra loro di quanto non lo siano due forme, come *canto* e *canteremo*, relative a tempi diversi dello stesso paradigma. Analogamente, forme suppletive come *è/sono* o *ho/abbiamo*, ma anche allomorfe radicali del tipo *muoio/moriamo* o *vado/andiamo*, risulteranno ragionevolmente localizzate in punti remoti della mappa, nonostante la loro appartenenza allo stesso paradigma verbale.

Ciononostante, è sensato chiedere a una SOM di evidenziare alcune porzioni morfologicamente coerenti del paradigma di un verbo se queste ultime sono sistematicamente marcate a livello superficiale. Per quanto detto finora, questo vale ovviamente per le terminazioni flessionali. Meno banale è osservare che le forme flesse di uno stesso verbo possono condividere uno stesso tema verbale. Il tema dell'imperfetto indicativo, ad esempio, si ricava a partire dal tema del presente con la semplice aggiunta di una *-v*: *ama-v*, *corre-v*, *senti-v* ecc. Analogamente il tema del futuro indicativo e del presente condizionale dei verbi della seconda e terza coniugazione si forma a partire dal tema del presente con l'aggiunta di *-r*: *corr-er*, *sent-ir*. Laddove dunque i processi di formazione del tema verbale sono regolari e sistematici, è ragionevole attendersi che la topologia delle forme nella mappa sia in grado di rappresentare questi processi raggruppando in aree adiacenti le forme flesse che li contengono.

Si noti inoltre che la mappa non ha nessuna ragione a priori per considerare morfologicamente simili due forme di lunghezza diversa ma entrambe terminanti in *-mo* come *cantiamo* e *correvamo*, piuttosto che *masticavo* e *menti*, che pure presentano le stesse lettere nello stesso ordine, ma in sequenze discontinue. E' ragionevole tuttavia attendersi che la mappa, una volta addestrata sulla morfologia dell'italiano, possa manifestare una maggiore sensibilità verso sottosequenze continue di caratteri, piuttosto che verso serie discontinue. L'osservazione non è peregrina. La rete, infatti, deve essere, almeno in linea di principio, predisposta in egual misura all'apprendimento di lingue dalla morfologia

concatenativa (come l'italiano) e dalla morfologia non concatenativa, come ad esempio l'arabo. Non avrebbe dunque senso introdurre, nella rete o nella rappresentazione morfologica, l'informazione che, in italiano, i morfemi sono costituiti da sequenze continue di caratteri. Ancora una volta, anche da questo punto di vista il criterio di rilevanza morfologica dovrebbe emergere come una proprietà induttivamente acquisita a partire dai dati.

7.2. Esperimenti condotti

7.2.1. L'Auto-organizzazione del verbo

Negli esperimenti è stata utilizzata una SOM (16 X 12) addestrata tramite un *corpus* di 1560 forme verbali, all'interno del quale ad ogni forma, codificata grazie alla sintassi complementare ("da sinistra a destra" e "da destra a sinistra"), è associato un vettore di *input*.

Prima di passare all'identificazione di probabili *clusters* morfologici distinti dalla mappa (Figure 6, 7, 8, 9, 10, 11, 12 e 13), è utile compiere un'osservazione generale di carattere tecnico. Si noti, infatti, che spesso uno stesso neurone è deputato a "gestire" diverse forme verbali allo stesso tempo. Ciò è dovuto al limitato numero di neuroni utilizzati cui fa riscontro un considerevole campione di addestramento che determina, quindi, un'alta densità di forme verbali per neurone. In breve, il medesimo neurone è stato prescelto, durante la fase d'apprendimento, per "ospitare" diversi vettori d'ingresso.

Intuitivamente, il sovraccarico di un neurone prescelto a rappresentare più di un vettore in input è il segno che, all'interno della rete, il neurone in questione si sta specializzando a rappresentare quanto di comune è condiviso da tutte le forme "attratte" dal neurone stesso. Questo processo è il sintomo di una generalizzazione morfologica emergente: il neurone tende a rappresentare un'intera classe di forme morfologicamente correlate. Parliamo di generalizzazione emergente e non di generalizzazione *tout-court* per il semplice fatto che non è ancora del tutto chiaro, alla luce degli esperimenti qui illustrati, quali dimensioni morfologiche governano la topologia di ogni singola mappa ottenuta. Che tutte le forme del PARTICIPIO PRESENTE si dispongano in un'area relativamente circoscritta della mappa (Figura 7), è segno evidente che la rete ha individuato la presenza di una terminazione rilevante condivisa dalle varie forme del participio presente usate in fase di addestramento. Meno chiaro, tuttavia, è come tutte queste forme si distribuiscono all'interno dell'area individuata e perché l'area stessa resta pur sempre così ampia rispetto all'intera superficie della mappa. La rete manifesta da questo punto di vista un comportamento parzialmente caotico, raggruppando tra loro forme che presentano la stessa lunghezza, o che iniziano per la stessa lettera. Altri effetti apparentemente arbitrari di adiacenza sono stati introdotti dalla codifica da noi adottata.

E' chiaro che raggruppamenti "casuali" di questa natura non hanno alcuna valenza morfologica, ma riflettono un ventaglio di punti di vista ancora non perfettamente allineati sul piano dell'analisi linguistica. Altri esperimenti (Pirrelli, Calderone, Herreros, Virgilio, 2004) indicano con sufficiente chiarezza che le mappe riportate qui di seguito sono l'istantanea di un processo di organizzazione morfologica ancora in atto, nel corso del quale la rete oscilla dinamicamente tra una classificazione delle forme verbali per radice (o

lessicale) e una classificazione delle stesse forme per terminazione flessionale (o morfo-sintattica). Quella cui assistiamo non è dunque una situazione di stallo, ma il singolo fotogramma di una dinamica in atto che porta la rete a dare sempre maggiore rilevanza a processi di categorizzazione morfologica di tipo macro-paradigmatico, secondo i quali forme di verbi diversi che tuttavia occupano la stessa cella paradigmatica sono disposti in zone adiacenti della mappa. Possiamo anticipare che i principali fattori responsabili di questa dinamica sono:

- i) il rapporto numerico tra esponenti lessicali e forme flesse nel corpus di addestramento
- ii) una transcodifica “random” dei caratteri in valori numerici
- iii) una tendenza della rete a minimizzare la quantità di “rumore” presente nella proiezione dei dati di input sulla mappa bi-dimensionale, privilegiando l’insorgenza di vaste aree di vettori relativamente omogenei rispetto ad un’organizzazione topologica a “macchia di leopardo”.

Il punto i) e iii) sono strettamente correlati. Data la tendenza espressa in iii), la mappa è in grado di orientarsi verso una categorizzazione macro-paradigmatica dei vettori in input se e solo se il corpus di addestramento riflette, dal punto di vista statistico, la natura inerentemente aperta del repertorio di radici verbali a fronte della classe chiusa di terminazioni flessionali. Sul piano cognitivo, ha senso ipotizzare che la riorganizzazione dei dati morfologici per raggruppamenti morfo-sintattici (piuttosto che lessicali) ha luogo nel momento in cui l’organizzazione lessicale per radici è messa in crisi da un numero sempre crescente di verbi diversi. Quanto al punto ii), è evidente che gli effetti di similarità accidentale introdotti dalla nostra codifica finiscono con l’introdurre un rumore intollerabile nell’organizzazione finale della mappa.

Come accennato in precedenza, la SOM, una volta addestrata, ha ordinato i valori di input sotto una qualche pertinenza morfologica che merita maggiori indagini. Non è difficile trovare una coincidenza tra la rilevanza morfologica che la SOM ha attribuito al corpus d’ingresso e alcune *categorie grammaticali*, definite da forme flesse con medesimi attributi morfo-sintattici (come il tempo, il modo, la persona ed il numero), costituenti nel loro insieme parte del loro paradigma verbale. La Figura 6, ad esempio, organizza spazialmente la categoria grammaticale IMPERFETTO INDICATIVO. La Figura 7 riproduce, invece, a livello topologico, le forme flesse appartenenti a due diverse celle paradigmatiche, identificate dagli attributi PARTICIPIO PRESENTE SINGOLARE e PARTICIPIO PRESENTE PLURALE. Seguendo questa logica, individuiamo nelle figure 8, 9, 10 e 11 dei *clusters* di forme flesse definiti in base alle seguenti opposizioni paradigmatiche: 3 P dell’IMPERFETTO CONGIUNTIVO, 1 S del PRESENTE CONDIZIONALE, 3 S del PRESENTE CONDIZIONALE e 2 P del FUTURO INDICATIVO.

Le sei disposizioni spaziali appena descritte sono organizzazioni emergenti all’interno di una dinamica temporale. La SOM, infatti, in quanto sistema non supervisionato di auto-organizzazione, ha operato una “pertinentizzazione” all’interno del corpus di addestramento portando in evidenza alcune relazioni grammaticali che le singole forme flesse intrattengono nel loro insieme.

Ciò che risulta importante per i nostri fini è il grado di salienza che la SOM attribuisce ad alcuni tratti morfologici anziché altri, salienza che ha portato alla costituzione delle suddette mappe. Ci sembra importante enfatizzare che, nonostante le indecisioni cui

abbiamo già fatto cenno, la rete è stata indotta a dare più peso alla parte terminale delle forme flesse rispetto alla parte iniziale, a dispetto del fatto che la loro rappresentazione vettoriale fosse del tutto omogenea. E' inoltre interessante notare come la SOM individui dei sotto-ordini all'interno delle disposizioni categoriali sinora considerate. Allomorfe come *dovr-*, *berr-*, *verr-*, *terr-* e *rimarr-* subiscono degli ulteriori raggruppamenti spaziali all'interno del medesimo *cluster* paradigmatico (Figure 9, 10 e 11) riuscendo così quasi ad evidenziare differenti gradi di pertinenza morfologica. Tutto questo a dispetto di una codifica vettoriale intenzionalmente grossolana e "brutale": una semplice transcodifica numerica della sequenza di caratteri che compongono la forma verbale, organizzata mediante sintassi complementare ("da sinistra a destra" e "da destra a sinistra"). Siamo pertanto autorizzati a concludere che la causa principale dell'insorgenza di queste organizzazioni spaziali di natura morfo-sintattica sia da rintracciare principalmente nelle proprietà strutturali esibite dalle forme usate in fase di addestramento e dalla capacità delle SOM di cogliere queste proprietà in termini di auto-organizzazione.

Sinora si sono considerati solo alcuni esiti positivi dei processi di auto-organizzazione morfologica di forme flesse, ottenuti mediante l'addestramento di una SOM. Una conferma indiretta della natura non casuale dei nostri risultati viene dai casi di fallimento. Volendo indagare su altre nuove e possibili organizzazioni paradigmatiche che la mappa potrebbe presentare, si è scoperto, infatti, che non sempre è possibile "ricomporre" strutture morfo-sintattiche omogenee. Consideriamo il caso di voler indagare se la SOM abbia ben organizzato una dimensione paradigmatica come la 1 S del PASSATO REMOTO INDICATIVO (Figura 12).

Risulta fin troppo chiaro dalla distribuzione spaziale delle forme flesse (si veda la Figura 12) che in questa circostanza la SOM non è stata in grado di porre dei confini di ordine categoriale all'organizzazione spaziale relativa agli attributi morfo-sintattici 1 S del PASSATO REMOTO INDICATIVO. Caso analogo accade per l'indagine su gli attributi PARTICIPIO PASSATO SINGOLARE e PARTICIPIO PASSATO PLURALE MASCHILE: la mappa non riesce a determinare un attrattore spaziale sul quale disporre le forme flesse appartenenti alle celle (del paradigma) definite da tali attributi (Figura 13).

In una prospettiva monistica, e quindi in termini di schemi associativi presenti nel lessico mentale del parlante, l'apprendimento di attributi morfo-sintattici quali il PARTICIPIO PASSATO (sia singolare che plurale) è regolato da domini associativi (relazioni tra forme flesse) che coinvolgono un numero sufficientemente elevato di parole. Da questo punto di vista, il corpus da noi utilizzato in fase di addestramento risulta troppo ristretto e disomogeneo per consentire alla rete di evidenziare le regolarità sistematiche e le sotto-regolarità che sono alla base della flessione del PARTICIPIO PASSATO in italiano. I processi di formazione del tema verbale del participio passato italiano, infatti, sono soggetti a una considerevole variabilità formale (Pirrelli, 2000; Pirrelli e Battista, 2000a e 2000b) nel senso di Aronoff (1994). Inoltre, va ricordato che nelle forme del participio passato la formazione irregolare del tema verbale (come in *fatto*, *perso*, *ridotto*, *fuso* ecc.) è quasi sempre (con la sola eccezione della forma *vissuto*) accompagnata da un suffisso di lunghezza minima, generalmente limitato a un carattere (-o, -a, -i, -e). Lo scarso peso in termini di lunghezza del suffisso in senso stretto rende ancora più difficile per la rete acquisire l'informazione morfologicamente rilevante del participio passato in modo puramente induttivo. Entrambe queste peculiarità del participio passato, variabilità nei processi di formazione del tema e scarsa lunghezza del suffisso, lo accomunano ad altre aree della flessione verbale italiana, come il presente indicativo e il passato remoto (Pirrelli

& Battista, 2000a e 2000b). Non è dunque un caso che le forme della 1 S PASSATO REMOTO siano soggette alla stessa dispersione topologica di quelle del PARTICIPIO PASSATO a fronte del corpus di addestramento scelto per i nostri scopi. La difficoltà della rete nel categorizzare forme flesse soggette a processi di formazione variabile riflette dunque, in una qualche misura, l'opacità iconica (nel senso di Dressler) della loro struttura morfologica, ma anche, al tempo stesso e forse in misura ancora più rilevante, la difficoltà di estendere induttivamente il processo di formazione acquisito ad altri casi e, dunque, la variabilità formale del processo stesso.

Una verifica sperimentale (non svolta in questo lavoro) complementare rispetto a quanto abbiamo appena detto consisterebbe nell'aumentare il campione di addestramento in modo tale da portare in luce le sotto-regolarità che soggiacciono ai processi di formazione del PARTICIPIO PASSATO italiano (singolare e plurale maschile). Il risultato di ciò si tradurrebbe, nei termini dell'organizzazione spaziale del paradigma, in una SOM che presenta al proprio interno tanti piccoli *clusters* (di diversa estensione) di forme flesse appartenenti tutte alla medesima cella paradigmatica.



Figura 6. Organizzazione spaziale emergente degli attributi morfo-sintattici 1 S (2 S, 3 S, 1 P, 2 P, 3 P) dell' IMPERFETTO INDICATIVO.

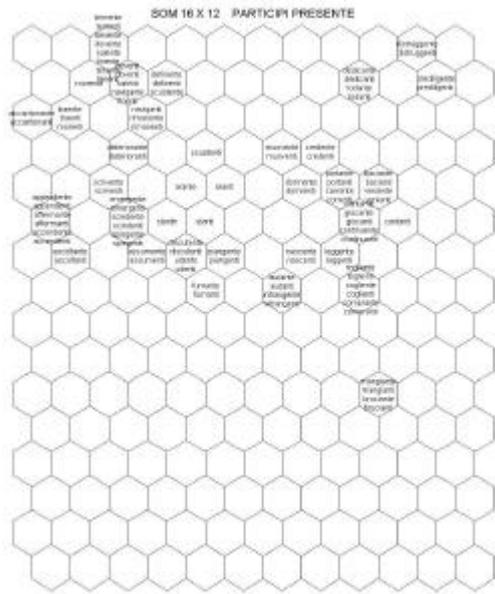


Figura 7. Organizzazione spaziale emergente degli attributi morfo-sintattici PARTICIPIO PRESENTE SINGOLARE e PARTICIPIO PRESENTE PLURALE.

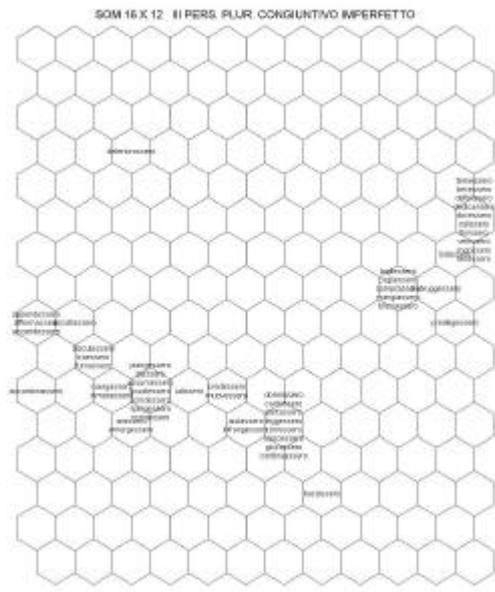


Figura 8. Organizzazione spaziale e emergente degli attributi morfo-sintattici 3 P dell'IMPERFETTO CONGIUNTIVO.

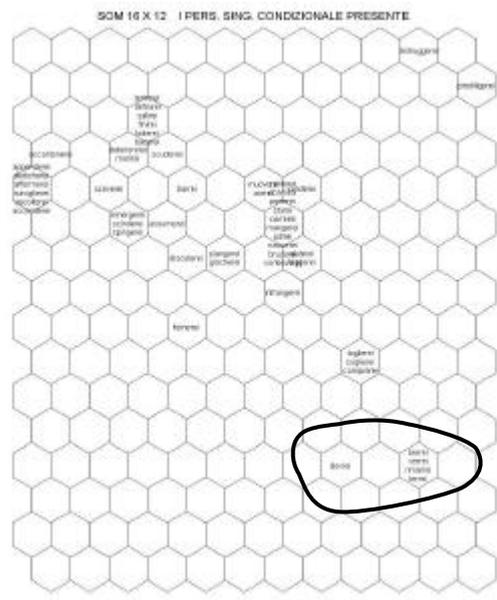


Figura 9. Organizzazione spaziale emergente degli attributi morfo-sintattici 1 S del PRESENTE CONDIZIONALE. In evidenza, casi di allomorfia.

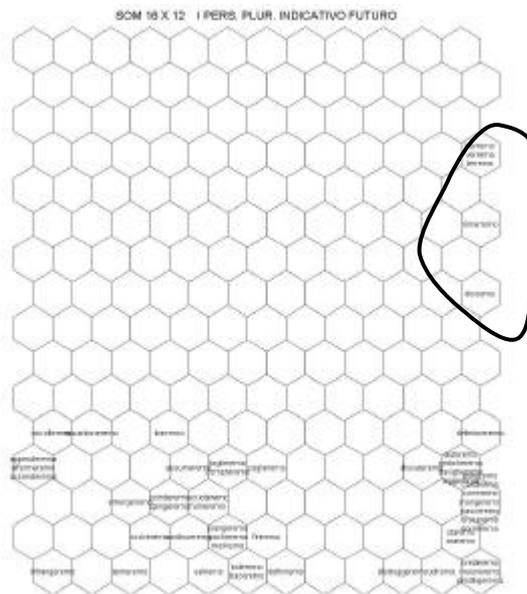


Figura 10. Organizzazione spaziale emergente degli attributi morfo-sintattici 1 P del FUTURO INDICATIVO. In evidenza, casi di allomorfia.

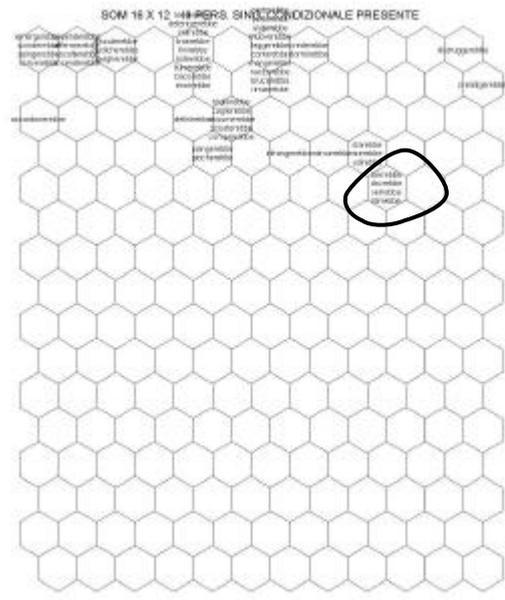


Figura 11. Organizzazione spaziale emergente degli attributi morfo-sintattici 3 S del PRESENTE CONDIZIONALE . In evidenza, casi di allomorfia.

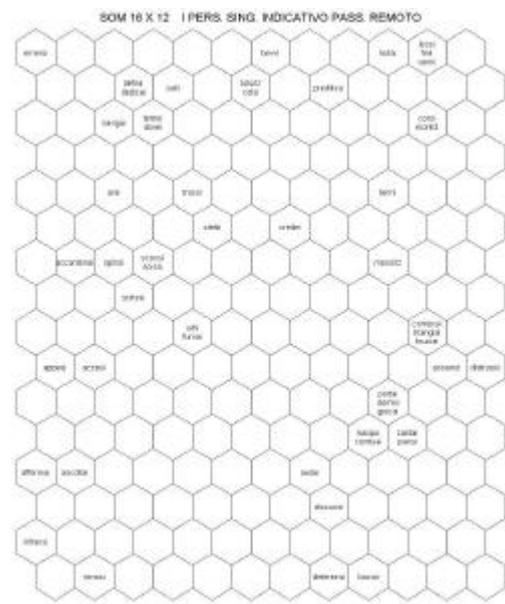


Figura 12. Organizzazione (poco soddisfacente) spaziale emergente degli attributi morfo-sintattici 1 S del PASSATO REMOTO INDICATIVO.

Nella migliore delle ipotesi le nuove forme verbali dovrebbero localizzarsi all'interno dei medesimi intorni spaziali su cui sono distribuite forme con gli stessi attributi morfo-sintattici. Proviamo, quindi, ad immettere nella SOM alcune forme inedite (dodici per l'esattezza) di un verbo come SAPERE, nella speranza che queste vengano analizzate in modo coerente col resto della struttura già acquisita dalla SOM. Il risultato è riportato in Figura 14.

E' agevole accorgersi che le nuove forme flesse si dispongono lungo la mappa ricalcando la topologia spaziale delle sei organizzazioni macro-paradigmatiche già trovate. In particolare un tema verbale come *sapr-* segue il medesimo comportamento spaziale di processi di formazione del tema analoghi cui la SOM è già stata esposta (in Figura 14 questa comunanza spaziale tra allomorfie è evidenziata mediante diagrammi di tipo insiemistico). La SOM riserva, quindi, alle forme flesse inedite la medesima rilevanza morfologica adottata per le forme già apprese, riuscendo ad esibire così un efficiente grado di generalizzazione e di estensione di conoscenza morfologica.

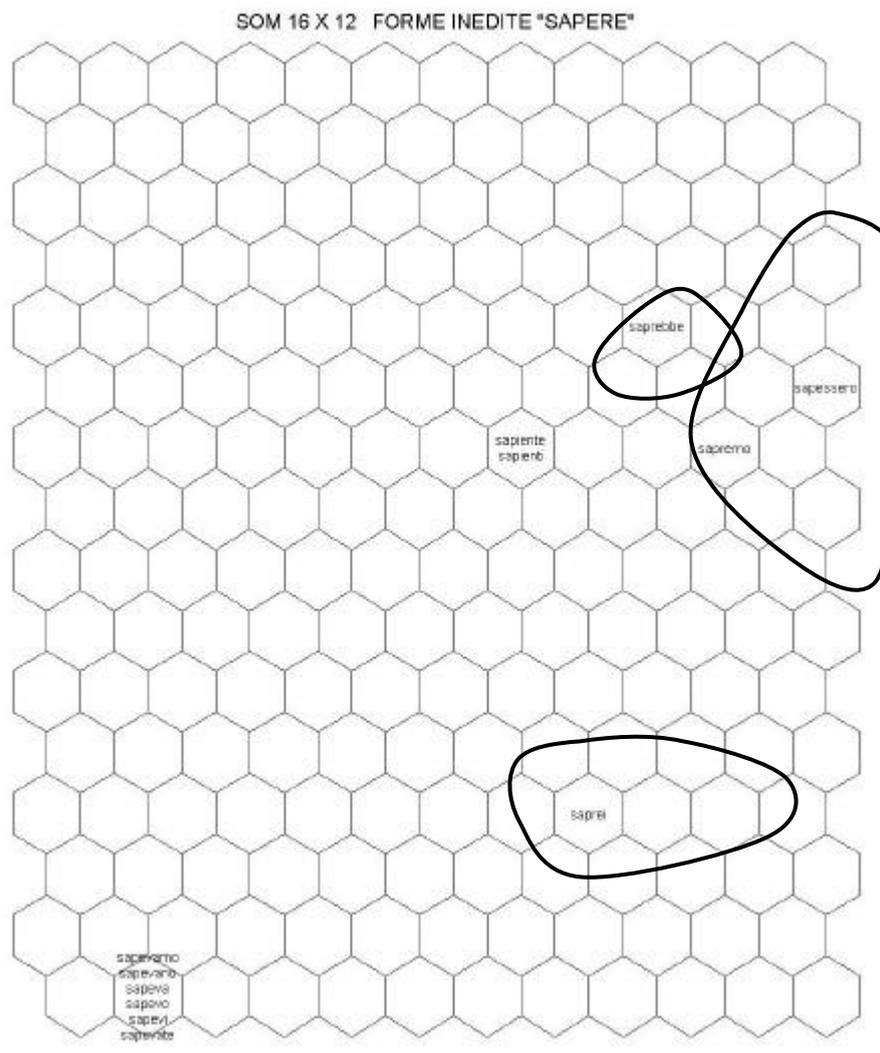


Figura 14. Risultato della categorizzazione di forme inedite. In evidenza, casi di allomorfie. Per un confronto visivo e un rapido controllo sulla corretta locazione delle nuove forme di SAPERE si rimanda alle organizzazioni paradigmatiche ritrovate in fig. 6, 7, 8, 9, 10 e 11.

8. Conclusioni e sviluppi

Il percorso di ricerca che abbiamo cercato di delineare nel presente lavoro, può essere sinteticamente descritto come il tentativo, in buona parte riuscito, di caratterizzare una struttura linguistica complessa ed altamente astratta come il *paradigma* nei termini di una proprietà emergente in una rete associativa di forme flesse. Le nostre conoscenze attuali

sull'insorgenza di strutture complesse a livello cerebrale evidenziano la natura graduale e incrementale di questo tipo di generalizzazioni. Tuttavia nessun modello computazionale di apprendimento automatico è stato finora in grado, per quanto ci è dato di sapere, di caratterizzare il macro-paradigma morfologico come il prodotto collaterale ed emergente del modo in cui i dati (le forme flesse) sono memorizzati nel lessico mentale del parlante. Tutto questo nonostante l'esistenza di alcuni modelli teorici dell'apprendimento della morfologia che si sono variamente ispirati a questa ipotesi (Bybee, 1995).

A ben guardare, il nostro studio cerca di andare al cuore di un apparente paradosso. Da una parte sappiamo che se i paradigmi non imponessero un vincolo globale e strutturale sul modo in cui i morfemi di una lingua entrano a far parte di unità linguisticamente più complesse quali le forme flesse, la morfologia sarebbe assai più difficile da apprendere. D'altra parte, tuttavia, non sembra, a prima vista, possibile vedere nei paradigmi il prodotto spontaneo del modo in cui le forme si auto-organizzano nella memoria del parlante. In altri termini, la prospettiva di organizzazione globale imposta dal paradigma mal si concilia con un approccio incrementale all'apprendimento. Inoltre, se i processi di organizzazione dinamica della conoscenza linguistica fossero davvero governati da principi di minimizzazione globale delle unità di base (come ad esempio suggerito dal principio della *minimum description length* o dai criteri di adeguatezza descrittiva di una grammatica adottati in linguistica teorica soprattutto di scuola generativa), la trasmissione della morfologia di una lingua da una generazione alla successiva dovrebbe tendere ad una riduzione sistematica del repertorio di allomorfi (o morfemi alternanti), unitamente ad altri effetti unidirezionali ed irreversibili di semplificazione quali, ad esempio, l'estensione generalizzata di un processo fonologico a tutto l'insieme delle forme che appartengono allo stesso paradigma. Al contrario, l'evidenza in nostro possesso sul cambiamento linguistico diacronico sembra deporre a favore di modelli emergenti di ristrutturazione selettiva, con fasi di oscillazione e repentine transizioni tra dinamiche continue.

Il paradigma teorico dell'emergentismo offre una chiave di lettura coinvolgente di questo apparente paradosso. Alcune strutture globali possono emergere "gratuitamente" come il prodotto di una dinamica temporale, vale a dire come il risultato dell'interazione e della riorganizzazione attraverso il tempo di unità linguistiche semplici, quando queste raggiungano una soglia critica dal punto di vista cognitivo. L'evidenza accumulata in queste pagine riguardo a una possibile ontogenesi della struttura (macro)paradigmatica del sistema verbale dell'italiano rappresenta a nostro avviso una conferma empirica indiretta dell'ipotesi emergentista. A questo scopo, abbiamo utilizzato le cosiddette SOMs come modello computazionale di una memoria associativa, che si sono rivelate strumento molto duttile e trasparente per i nostri scopi di ricerca. Attraverso l'uso delle SOMs abbiamo cercato di simulare al computer il modo in cui le strutture paradigmatiche di una lingua come l'italiano possono emergere dinamicamente su una scala temporale definita da sessioni cicliche di apprendimento.

Nella serie di sperimentazioni abbiamo avuto modo di evidenziare il processo di auto-organizzazione dinamica attraverso il quale alcune forme verbali, cui la mappa è esposta in fase di addestramento, si addensano in aree topologicamente circoscritte, corrispondenti a classi morfo-sintattiche naturali. In particolare abbiamo notato che tutte le forme flesse associate ad una stessa cella del paradigma (e quindi definite dagli stessi attributi morfo-sintattici) si dispongono intorno ad un attrattore spaziale comune che, allo stesso tempo, determina la rilevanza (morfologica) a cui le varie forme sono soggette.

Ripercorrendo a ritroso i passi procedurali che hanno portato all'organizzazione macro-paradigmatica di 1560 forme flesse, si ritrova che forse l'aspetto che maggiormente enfatizza i risultati ottenuti risiede all'inizio, e non alla fine del lavoro tutto, e coincide con il criterio di codifica dei dati. Come può, infatti, una mera trascrizione ortografica, parziale o totale che sia, pervenire ad una modellizzazione del linguaggio così granulare e strutturalmente elegante?

Di certo (buona) parte della riuscita dei propositi posti all'inizio di queste pagine è da rintracciare nella forza organizzatrice che un sistema connessionistico come la SOM riserva ai dati di input. La questione del raggiungimento di un ordine di natura spaziale, nel caso di dati morfologici, è, però, una conseguenza del fatto che i dati stessi presentino comunanze e analogie strutturali soggiacenti. Un processo di auto-organizzazione, infatti, non porta all'insorgenza di alcunché se la natura dei dati non presenta una sistematica predisposizione a legami o relazioni di similarità. E' in tale accezione che il processo di auto-organizzazione morfologica si definisce "emergente".

Di necessità, l'evidenza che qui riportiamo è solo indiretta e si basa su alcune astrazioni forti sui dati, che in alcuni lettori possono, con una qualche ragione, ingenerare il sospetto di un'eccessiva semplificazione del problema. Siamo consapevoli, ad esempio, che rappresentare le forme flesse di una lingua come la semplice codifica numerica della sequenza di caratteri che le compone rappresenta una radicale semplificazione dell'evidenza che si rende disponibile al bambino nella sua fase di maturazione linguistica. Tuttavia è importante a nostro avviso tener conto del fatto che un input impoverito può, nel nostro caso, rendere più arduo l'obiettivo che intendiamo conseguire, piuttosto che semplificarlo. Aggiungere ad ogni forma flessa l'informazione relativa alla cella paradigmatica che quella forma occupa consentirebbe alla rete di individuare raggruppamenti morfo-sintattici naturali in modo relativamente rapido. Lo stesso risultato, nel nostro caso, è stato raggiunto con un input impoverito lasciando alla mappa il compito di indurre una struttura morfo-sintattica sulla semplice base delle similarità tra forme.

Allo stesso modo, abbiamo cercato per quanto possibile di non porre vincoli a priori sul tipo di similarità che la rete è in grado di estrarre dai dati. Se avessimo rappresentato una forma flessa come una semplice sequenza di caratteri da sinistra a destra avremmo indotto la rete a concentrarsi su similarità di naturale lessicale (relative cioè alla radice di una parola) piuttosto che morfo-sintattica (relative alle terminazioni flessionali). Offrendo alla rete la possibilità di cogliere entrambi i tipi di similarità allo stesso tempo, abbiamo reso più difficile il processo di induzione, dal momento che un tipo di similarità rappresenta del rumore dal punto di vista dell'altro tipo di similarità. La scelta operata infine dalla rete, e cioè quella di dare maggior enfasi agli aspetti morfo-sintattici rispetto a quelli lessicali, riflette a nostro avviso una proprietà strutturale dei dati, emergente rispetto alla topologia di neuroni sulla quale la mappa converge.

I possibili sviluppi di questo lavoro sono molteplici.

In primo luogo, intendiamo studiare in modo più analitico il tipo di dinamica che qui abbiamo evidenziato in modo ancora alquanto impressionistico. A questo proposito, una selezione più accurata del corpus di addestramento, con particolare riferimento al rapporto numerico tra forme flesse ed esponenti lessicali, ci dovrebbe consentire di definire una vera e propria curva di apprendimento, espressa nei termini di misure entropiche del caos residuo presente nella mappa ad ogni ciclo di apprendimento. Inoltre, la stessa nozione di ciclo di apprendimento dovrebbe mirare a riprodurre, in modo più realistico, il dato morfologico al quale è esposto un bambino nella sua maturazione linguistica, privilegiando

verbi molto frequenti (e molto irregolari) nelle primissime fasi di apprendimento, per poi includere gradualmente forme più regolari. Alcuni risultati preliminari (Pirrelli, Calderone, Herreros, Virgilio 2004) sono già estremamente indicativi a questo riguardo.

In secondo luogo, è nostro intendimento raffinare la codifica numerica delle forme in input. Abbiamo già avuto modo di osservare che alcuni raggruppamenti morfologici innaturali possono essere ricondotti a un effetto di falsa analogia indotta dalla casuale similarità (in termini numerici) tra caratteri distinti. E' ragionevole supporre che una codifica numerica più motivata sul piano linguistico (con le vocali che risultino più vicine tra loro nella codifica di quanto non lo siano rispetto alle consonanti) possa aiutare a risolvere questo problema.

Le strutture paradigmatiche non rispondono solo alla funzione descrittiva di esprimere generalizzazioni emergenti dai dati. Uno dei loro usi primari è quello di consentire a un parlante di generare nuove forme a partire da forme già conosciute.

In quest'ottica di ricerca vanno ricondotti alcuni tentativi di realizzazione di SOMs di secondo livello al fine di pervenire non solo alla nozione di paradigma come si è appena visto, ma di macro-paradigma o schema flessionale. Seguendo la terminologia di Carstairs (1987), un macro-paradigma è una classe di paradigmi che selezionano lo stesso insieme di terminazioni flessionali. E' chiaro che pervenire ad una corretta individuazione dei macro-paradigmi (o schemi flessionali) della coniugazione italiana è un prerequisito cruciale per poter estendere la conoscenza acquisita a casi (parole) cui non si è mai stati esposti.

E' plausibile, in questi termini, ipotizzare (come propositi per un lavoro futuro) che più SOMs possano interagire a cascata, in modo tale che l'output di una SOM possa essere codificato come input per un'altra SOM che sia in grado di generalizzare per categorie grammaticali più astratte. L'organizzazione di una mappa paradigmatica (Figure 6, 7, 8, 9, 10 e 11) diventa, così, oggetto di un nuovo processo di apprendimento dell'insieme di forme flesse appartenenti allo stesso paradigma, operato da una seconda SOM. Nella mappa di secondo livello così ottenuta, verbi con un comportamento flessionale simile dovrebbero occupare posizioni adiacenti.

L'idea di fondo, infatti, è che se due esponenti lessicali, come *venire* e *tenere*, condividono un numero critico di terminazioni, allora è ragionevole supporre che essi condividano tutto lo schema flessionale. E' questo tipo di estensione analogica che dovrebbe essere catturata attraverso l'uso di mappe del secondo livello. Ed è esattamente questo tipo di generalizzazione a svolgere un ruolo cruciale nei processi di generazione di forme sconosciute. E' pertanto ragionevole supporre che l'architettura di mappe SOMs di doppio livello (una prima SOM con organizzazione paradigmatica intesa come nuovo input ad una SOM di dimensionalità ridotte) possa essere utilmente sfruttata per simulare processi *bottom-up* di inferenza linguistica attraverso i quali sia possibile interpolare paradigmi incompleti riempiendone le celle rimaste vuote con forme indotte a partire da altri paradigmi simili. Così facendo, saremmo in grado di dare un fondamento emergentista a quei processi di estensione analogica che la letteratura linguistica ha sempre considerato come cruciali nell'apprendimento e nell'evoluzione di una lingua e che hanno ricevuto di recente notevole attenzione da parte di studiosi di sistemi automatici di apprendimento, sia simbolici che sub-simbolici.

I risultati di questo lavoro potrebbero infine avere un impatto anche su problemi di architettura della grammatica interiorizzata dal parlante e sulla natura delle interfacce che governano lo scambio di informazioni tra i vari livelli di analisi linguistica. Caratterizzare su base neurale quei processi di "composizionalità debole" del linguaggio che sono assai

ricorrenti in lingue la cui morfologia è strutturata per macro-paradigmi ci consentirebbe infatti di fare un passo in avanti nella nostra comprensione dei principi e delle strutture che governano il lessicale mentale di un parlante e di capire meglio come lessico e grammatica interagiscono al livello della competenza linguistica.

E' probabile che principi di massima economia descrittiva svolgano un ruolo relativamente marginale nella dinamica delle prime fasi di apprendimento. Resta aperta tuttavia la possibilità che una volta raggiunta una configurazione ragionevolmente stabile e poco caotica della mappa responsabile di un certo livello di analisi, principi fino ad allora rimasti relativamente inerti quali il *minimum description length* possano acquisire un ruolo importante nell'ottimizzare la topologia della mappa stessa.

Indirizzo dell'autore:

Scuola Normale Superiore, Laboratorio di Linguistica

Piazza dei Cavalieri, 7 56126 Pisa, Italia

<b.calderone@sns.it>

Bibliografia

- Anderson, J. A. & Rosenfeld, E., eds. 1988 *Neurocomputing. Foundations of Research*. Cambridge: M.I.T. Press.
- Aronoff, M. 1976 *Word Formation in Generative Grammar*. Cambridge: M.I.T. Press.
- 1994 *Morphology by Itself*. Cambridge: M.I.T. Press.
- Bertinetto, P. M.
1995 "Compositionality and Non-Compositionality in Morphology". In Dressler e Burani, eds., 1995: 9-33.
- Boden, M. A., ed.
1990 *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press.
- Bybee, J.
1988 "Morphology as Lexical Organization". In Hammond e Noonan, eds., 1988: 119-141.
- 1995 "Regular Morphology and the Lexicon". *Language and Cognitive Processes*, 10 (5): 425-455.
- Bybee, J. & Hopper, P., eds.
2000 *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
- Bybee, J. & Moder, C.
1983 "Morphological Classes as Natural Categories". *Language*, 59: 251-270.
- Bybee, J. & Newman, J.
1995 "Are Stem Changes as Natural as Affixes?". *Linguistics*, 13: 633-654.
- Bybee, J. & Pardo, E.
1981 "Morphological and Lexical Conditioning of Rules: Experimental Evidence From Spanish". *Linguistics*, 19: 937-968.

- Bybee, J. & Slobin, D. I.,
 1982 "Rules and Schemas in the Development and Use of the English Past Tense. *Language*, 58: 265-289.
- Carstairs (McCarthy), A.
 1987 *Allomorphy in Inflection*. London: Croom Helm.
- Chalmers, D. J.
 1992 "Subsymbolic Computation and the Chinese Room". In Dinsmore, ed., 1992: 25-48).
- Dinsmore, J., ed.
 1992 *The Symbolic and Connectionist Paradigms: Closing the Gap*. London: Lawrence Erlbaum Associates.
- Dressler, U. & Burani, C., eds.
 1995 *Crossdisciplinary Approaches to Morphology*. Austria: Akademie der Wissenschaft.
- Dyer, M. G.
 1990 "Distributed symbol formation and processing in connectionist networks". *Journal of Experimental and Theoretical Artificial Intelligence*, 2: 215-239.
- Feldman, J. A. & Ballard D. H.
 1982 "Connectionist Models and Their Properties". *Cognitive Science*, vol. 6: 205-254.
- Floreato, D.
 1996 *Manuale sulle reti neurali*. Bologna: Il Mulino.
- Fodor, J. A. & Pylyshyn, Z. W.
 1988 "Connectionism and Cognitive Architecture: a Critical Analysis". In Pinker e Mehler, eds., 1988: 3-71.
- Gallant, S. I.
 1995 *Neural Network Learning and Expert Systems*. Cambridge: M.I.T. Press.
- Gaussier, E.
 1999 "Unsupervised Learning of Derivational Morphology from Inflectional Lexicons". In *Proceedings of the Workshop on Unsupervised Methods in Natural Language Processing*, University of Maryland: 24-30.
- Goldsmith, J.
 2001 "Unsupervised Learning of the Morphology of a Natural Language". *Computational Linguistics* 27(2):153-198.
- Hafer, M. & Weiss, S.
 1974 "Word segmentation by letter successor varieties". *Information Storage and Retrieval*, 10: 371-85.
- Harris, J. W.
 1968 *Spanish Morphology*. Cambridge: M.I.T. Press.
- Hebb, D. O.
 1975 *L'organizzazione del comportamento*. Milano: Angeli.
- Hockett, F.
 1954 "Two models of grammatical description". *Word*, 10: 210-231.
- Honkela, T.
 1997 *Self-Organizing Maps in Natural Language Processing*. Tesi di dottorato, Helsinki University of Technology.
- Jensen, J. T.

- 1990 *Morphology: Word Structure in Generative Grammar*. Amsterdam: John Benjamins.
- Kazakov, D.
1997 "Unsupervised Learning of Naive Morphology with Genetic Algorithms". In Daelemans, van den Bosch e Weijters, eds., 1997: 105-112.
- Kohonen, T.
1984 *Self-organization and Associative Memory*. Berlin: Springer.
- 1990 "The Self-Organizing Map". In *Proceedings of the Institute of Electrical and Electronics Engineers*, 78 (9): 1464-1480 (ora in Mehra e Wah, eds., 1992: 359-375).
- 1995 *Self-Organizing Maps*. Berlin: Springer.
- Ling, C. X.
1994 "Learning the past tense of English: the symbolic pattern associator versus connectionist models". *Journal of Artificial Intelligence Research*, 1: 209-229.
- Mehra, P. & Wah, B. W., eds.
1992 *Artificial Neural Networks. Concepts and Theory*. Los Alamitos: IEEE Computer Society Press.
- Miikkulainen, R.
1997 "Natural Language Processing with Subsymbolic Neural Networks". In Browne, ed., 1997: 120-139.
- Newell, A. & Simon, H. A.
1976 "Computer Science as Empirical Inquiry: Symbols and Search". *Communication of the Association for Computing Machinery*, 19: 113-126 (ora in Boden, ed., 1990: 105-135).
- Parisi, D.
1989 *Intervista sulle reti neurali*. Bologna: Il Mulino.
1999 *Mente*. Bologna: Il Mulino.
- Penna, M. P. & Pessa, E.
1993 *Introduzione alla psicologia connessionistica*. Roma: Di Renzo.
- Pessa, E.
1992 *Intelligenza artificiale. Teorie e sistemi*. Torino: Bollati Boringhieri.
1993 *Reti neurali e processi cognitivi*. Roma: Di Renzo.
- Pessa, E. & Penna, M. P.
1993 *La rappresentazione della conoscenza*. Roma: Armando.
2000 *Manuale di scienza cognitiva*. Bari: Laterza.
- Pinker S. & Mehler, J.
1988 *Connections and Symbols*. Cambridge: M.I.T. Press.
- Pinker, S. & Prince, A.
1988 "On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition". *Cognition*, 28 (2): 2-193.
- Pirrelli, V.
2000 *Paradigmi in Morfologia. Un approccio interdisciplinare alla flessione verbale dell'italiano*. Pisa: Istituti editoriali e poligrafici internazionali.
- Pirrelli, V. & Battista, M.
2000a "The Paradigmatic Dimension of Stem Allomorphy in Italian Inflection". *Italian Journal of Linguistics*, vol. 12 (2): 307-380.
2000b "On The Interaction of Paradigmatic and Syntagmatic Stem Alternation in Italian Conjugation". *Acta Linguistica Hungarica*, vol. 47(1-4): 289-2314.

- Pirrelli V., Calderone B., Herreros I. & Virgilio M.
 2004 "The Emergence of Paradigm-based Relations in Self-Organizing Maps of the Morphological Lexicon" (in preparation).
- Pirrelli V. & Yvon, F.
 1999 "Analogy in the lexicon: a probe into analogy based machine learning of language". In *Proceedings of the 6th International Symposium on Human Communication*, Santiago de Cuba: 576-581.
- Plunkett, K. & Elman J.E.
 1998 *Exercises in Rethinking Innateness*. Cambridge: M.I.T. Press.
- Rumelhart, D. E. & McClelland, J. L.
 1986 "On learning the past tenses of english verbs". In Rumelhart, D. E. e McClelland, J. L., eds., 1986, *Parallel Distributed Processing*, vol. 2: 216-271, Cambridge: M.I.T. Press.
- Shannon, C. & Weaver, W.
 1971 *La teoria matematica delle comunicazioni*. Milano: Etas Kompass.
- Skousen, R.
 1989 *Analogical Modelling of Language*. Dordrecht: Kluwer.
- Smolensky, P.
 1992 *Connessionismo: tra Simboli e Neuroni*. Genova: Marietti.
- Taft, M..
 1988 "A morphological-decomposition model of lexical representation". *Linguistics*, 26:657-667.
- van den Bosch, A.
 1997 *Learning to pronounce written words: A study in inductive language learning*. Tesi di dottorato, University of Maastricht.
- Vennemann, T.
 1972 "Rule Inversion". *Lingua*, 29: 209-242.

