

Self-Organizing Maps in Natural Language Processing



Giovanni Da San Martino

Corso di

Trattamento del Linguaggio Naturale

A.A. 2002/2003

Sommario:

- *Un nuovo approccio al Natural Language Processing : le self organizing maps(SOM)*
- *Matematica + Statistica + informatica + linguistica=SOM*
- *Motivazioni per il loro uso*
- *SOM : l'algoritmo*
- *Principi ed elenco degli utilizzi in NLP*
- *Alcune Applicazioni :*
 - *Word Category Maps*
 - *Information retrieval con le SOM : WEBSOM*

Reti Neurali (feed-forward) ^[10]

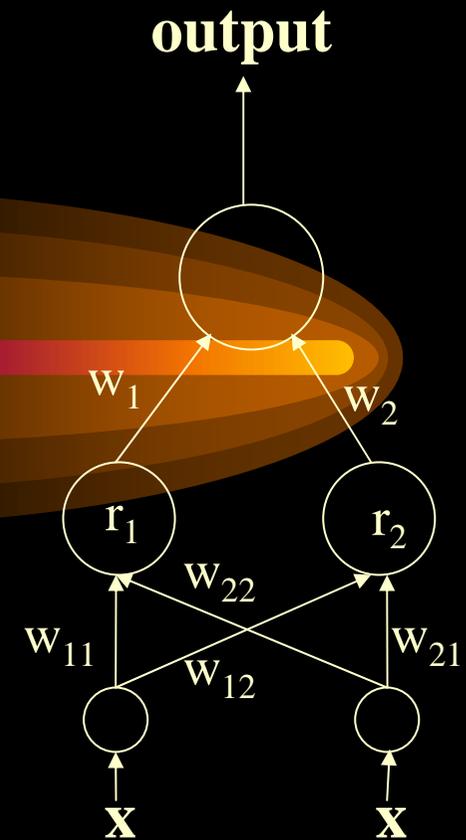
Cercano di modellare il neurone biologico :

- Progettazione dell'architettura della rete
- Apprendimento (supervisionato):

a. \forall esempio x in T (training set)

b. Calcola $y = \mathbf{output}(x)$

c. Se $y \neq d(x)$ output aspettato per x allora modificano i pesi w_{ij} in modo da avvicinare y a $d(x)$



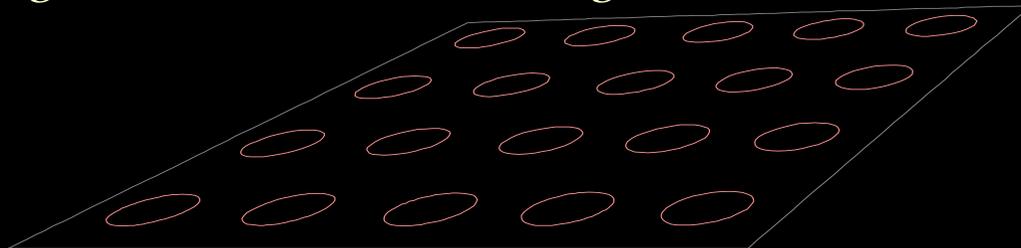
la rete è in grado di generalizzare dai pochi esempi e classificare correttamente anche esempi mai visti prima (non memorizza soltanto le istanze dell'input)

Reti Neurali : Caratteristiche

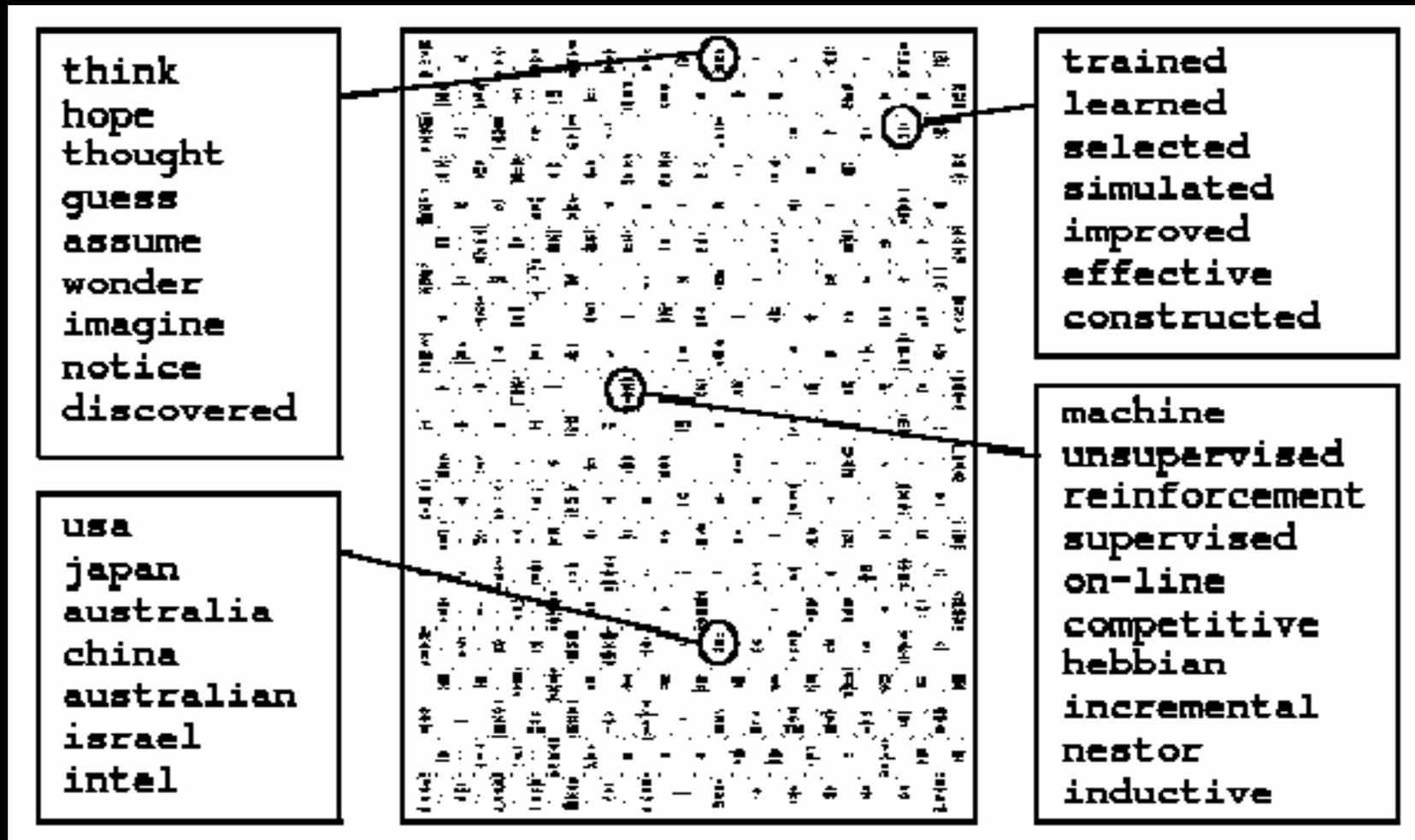
- *Sono “Black-box” : non si riesce cioè ad estrarre regole da una rete neurale (si conosce il risultato ma non come lo calcola)*
 - *Scarsa capacità di sintesi e trasmissione della conoscenza*
- *Riescono a trattare agevolmente (senza modifiche) dati con rumore o incompleti*

Introduzione : le SOM

- *Organizzano automaticamente insiemi di dati n -dimensionali su una mappa k -dimensionale ($k < n$) in modo da rispettare la distribuzione statistica dell'input*
 - *elementi simili tendono a posizionarsi vicino nella mappa*
- *Rete neurale ad Apprendimento NON Supervisionato*
 - *la formazione della mappa avviene senza che nessun esperto debba definire quale sia l'output corretto per ogni istanza del training set*

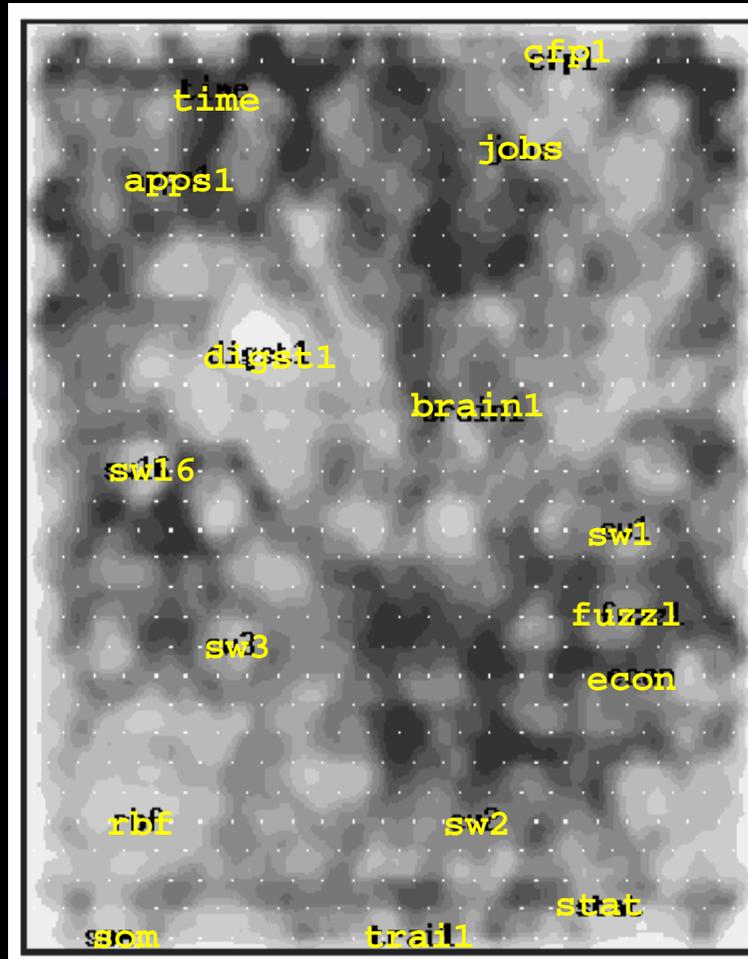


Esempio 1 : Word Categorization



Self organized semantic map formata su articoli tratti dal newsgroup comp.ai.neural-nets [21]

Esempio2:document Categorization



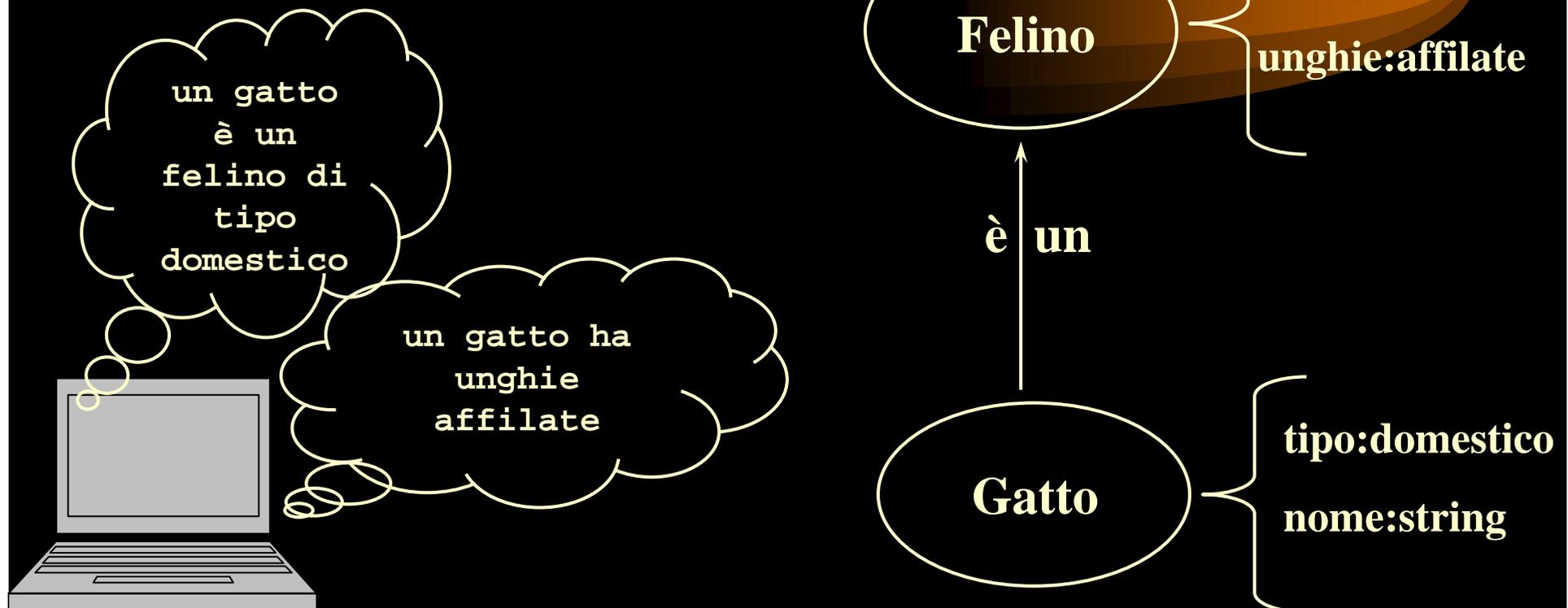
Explanation of the symbols on the map

cfp1 - conferences
time - time series
jobs - vacancies
apps1 - applications: face, speech
digst1 - Neuron Digest, CFPs
brain1 - brain sized NN
sw16 - software
sw1 - implementations
fuzz1 - fuzzy logic
sw3 - source code
econ - finance
rbf - Radial Basis Function networks
sw2 - software
stat - NN vs statistics
som - Self-Organizing Maps
trail1 - training, testing

Gli stessi articoli tratti dal newsgroup comp.ai.neural-nets sono ora raggruppati per argomento

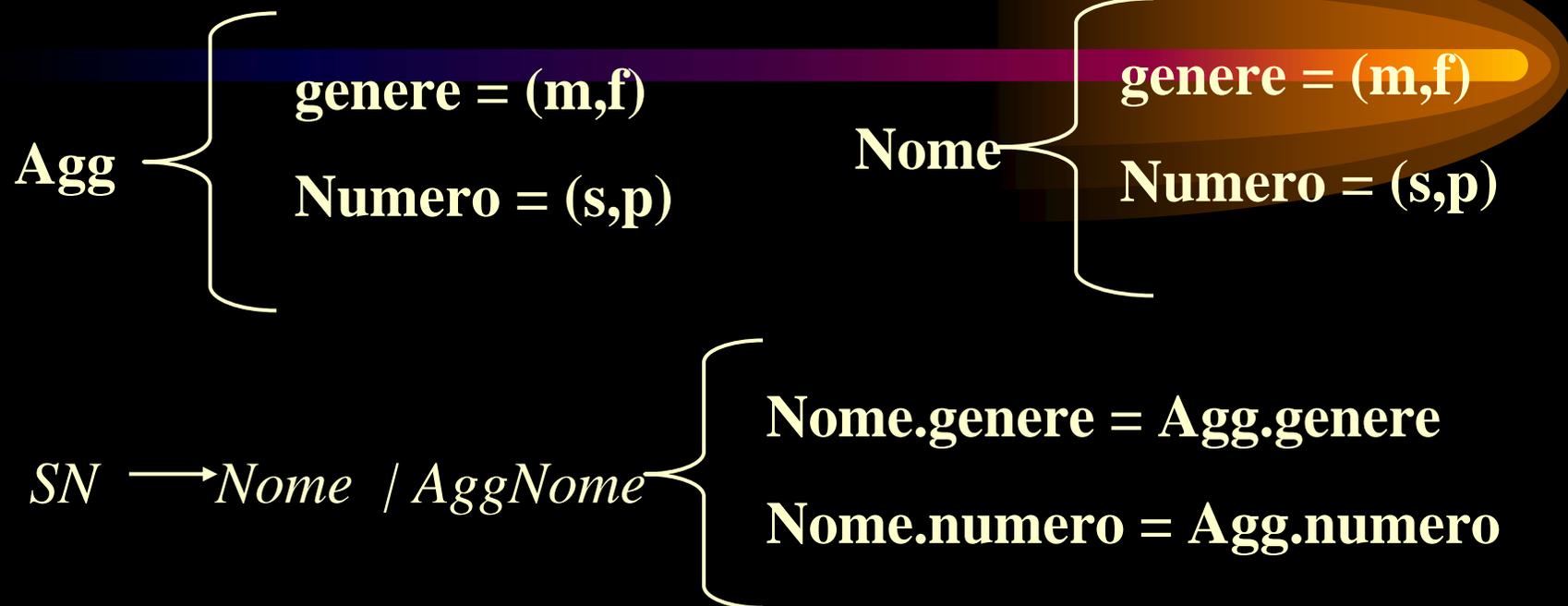
Traditional AI Knowledge Representation

- *Classi di Oggetti (gatto, felino)*
- *Relazioni (appartenenza: il gatto è un felino)*
- *Attributi (domestico, 4 zampe)*



Semantics in Traditional AI

- *Logica del primordine*
- *Attribute Grammars :*



un sintagma nominale è valido solo se aggettivo e nome concidono in genere e numero

Classic NLP VS Reti Neurali

il mondo è costituito da oggetti e relazioni tra questi - le parole sono etichette per gli oggetti

Il linguaggio naturale è uno strumento tramite il quale gli individui si costruiscono un modello del mondo

approccio deduttivo, uso di conoscenza instillata a priori dal progettatore (regole)

linguaggio è dinamico^[5], riflette la percezione del mondo del soggetto → no conoscenza assoluta a priori ma approccio induttivo, si costruisce un modello da esempi

•Ultimamente alcuni tentativi di unire i due approcci ^[9]

Need for Adaptation ^[11]

Nello studio di un linguaggio tradizionalmente si usano categorie prefissate e statiche ($SN \rightarrow Nome \mid AggNome$)

Un sistema a regole non permette di tenere di conto agevolmente di tutte le eccezioni e i particolari usi di una struttura sintattica



Il numero di regole tende a crescere esponenzialmente



È difficile costruire e maneggiare le Attribute Grammars



*In più il linguaggio è dinamico, "rumoroso" e context-sensitive :
é quindi necessario (oltre che auspicabile) un metodo automatico per la determinazione di sintassi e semantica*

In Sostanza...Perchè le SOM?

I formalismi presenti fin'ora permettono di modellare alcuni aspetti del linguaggio, le SOM intendono allargare queste possibilità rappresentative per includere :

- *Contesto : il significato di termini ambigui si ricava dal contesto*
- *Soggettività : l'uso della lingua varia da persona a persona perchè ognuno ha una sua diversa concezione del mondo*

A Cosa Ci Serve Il Contesto?

- *Durante la fase di interpretazione semantica si assiste ad una esplosione combinatoria data dall'ambiguità dei termini*
- *L'ambiguità è molto più frequente di quanto si pensi (l'uomo riduce il numero di interpretazioni grazie al contesto e ad un sistema anticipatorio)*
- *Noise Handling : certe parole possono essere usate erroneamente ma il contesto può rendere chiaro il significato*
- *Fra la grafia delle parole e la loro semantica non c'è correlazione → se non si usano regole (conoscenza a priori) il contesto è necessario per interpretare una parola*

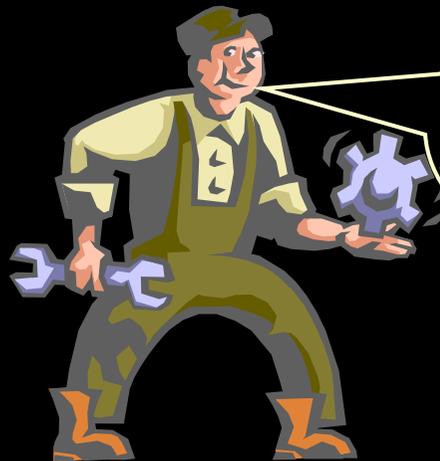
Soggettività

- *Per l'uomo la comprensione del linguaggio naturale è basata sulla sua esperienza personale*
 - *Persone con storie diverse hanno un diverso modo di interpretare una stessa espressione in linguaggio naturale*
 - *La logica tradizionale difficilmente riesce a trattare questo aspetto*
 - *Fin'ora si è semplificato scegliendo una interpretazione a priori su cui si potesse operare con la logica matematica*

Soggettività : un Esempio

- *Se ammettiamo diverse concezioni del mondo*
- *es. diverse concezioni del viola rendono impossibile la comunicazione se non si ammettono limiti più “laschi” al concetto di viola (fuzzy)*
- *SOM : misura di vicinanza della mappa = fuzzyness*

Viola=[21.7,250.8,158.1]



*La chiave
inglese è
viola*

Viola=[21.7,250.8,158.2]

NO!

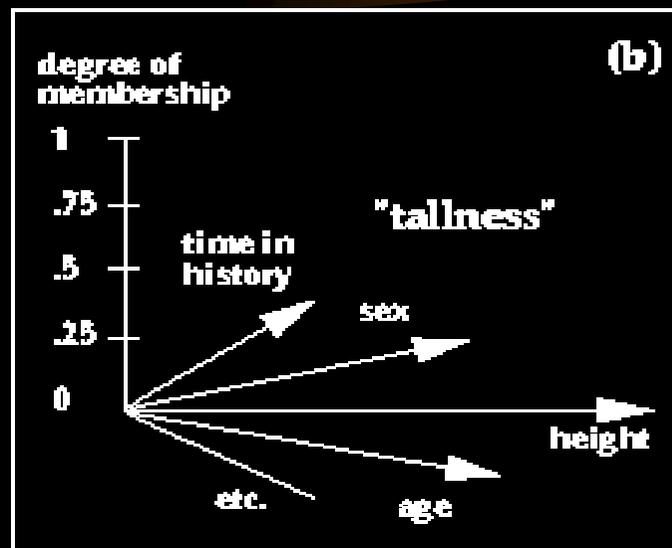
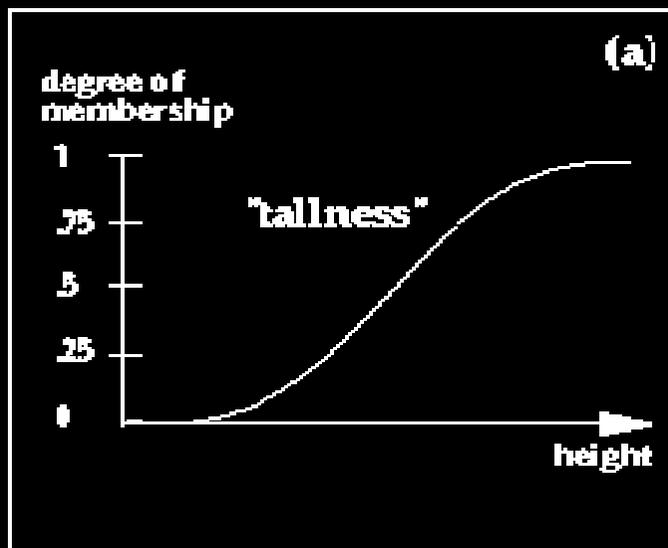


Soggettività

- *Nel tempo anche la concezione del colore viola che ha una stessa persona può cambiare*
 - *es. Cos'è un algoritmo al 1° e 5° anno di studi?*
- *Portata all'estremo questa considerazione ci dice che non esistono due persone con la stessa concezione del mondo*
- *La comunicazione nasce da un processo di apprendimento per far avvicinare quanto basta le due diverse rappresentazioni*
- *Le SOM che vedremo possono in principio essere estese per trattare questo aspetto!*

Logica Sfumata e Soggettività

- Anche la logica fuzzy non è del tutto adatta per la soggettività.
- La scelta di ogni funzione di membership è del tutto arbitraria e diventa, come accadeva per la logica, un'interpretazione scelta a priori, a cui doversi attenere

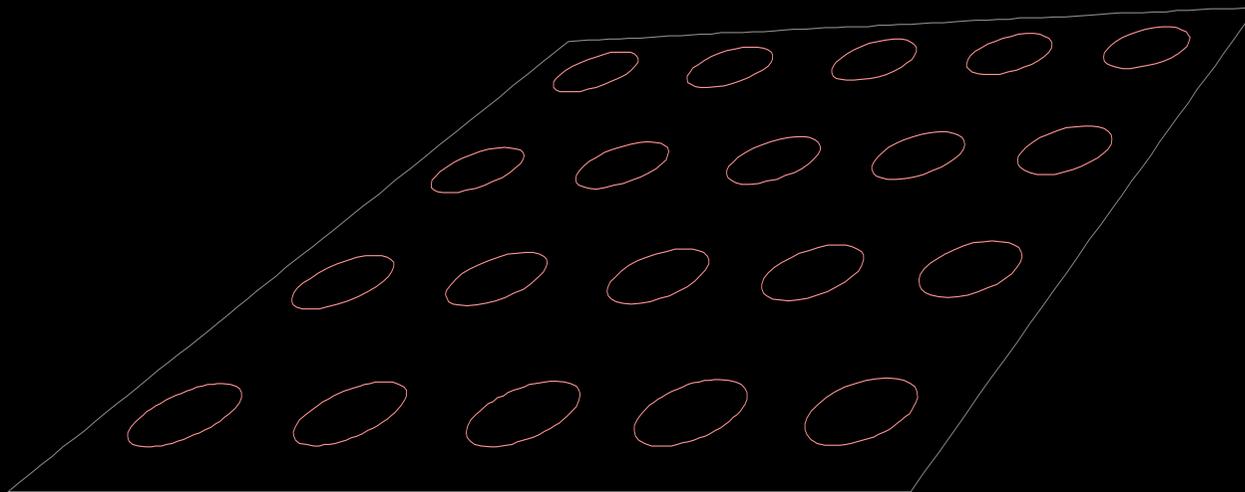


Motivazioni biologiche (e non)

- *Le SOM nascono dall'osservazione che nella corteccia cerebrale i neuroni si auto-organizzano topograficamente secondo la loro funzione.*
- *Ogni area si specializza per un insieme preciso di compiti.*
- *Queste mappe non sono interamente predeterminate geneticamente:*
 - *piuttosto sono organizzate (non si sa come) durante i primi tempi dello sviluppo .*
- *Confermano le teorie di Rosch e Lakoff sui prototipi adattivi (vedi dopo)*

SOM : Algoritmo (1)

- *Ciò che vogliamo è rappresentare in un array bidimensionale l'insieme di input in maniera significativa (neighbour preserving)*
- *Ogni istanza j dell'input è rappresentata da un vettore $v_j \in \mathcal{R}^n$*
- *Ad ogni nodo i è associato un vettore prototipo $m_i \in \mathcal{R}^n$*
- *Si definisce la topologia della mappa (n° di nodi)*



SOM : Algoritmo (2)

- 1. $\forall i$ $m_i(0)$ nodo della mappa è inizializzato casualmente*
- 2. $\forall j$ v_j istanza di input (scelta a caso):*
 - 1. Calcolare il nodo della mappa che ha Distanza minore da v_j*
winner = $\arg \min_i \{ \| m_i - v_j \| \}$
 - 2. Il nodo winner e i suoi adiacenti vengono “avvicinati” all’input (più un nodo è vicino a winner e più viene influenzato)*

$$\begin{aligned} m_i(t+1) &= m_i(t) + \alpha(t) h(t, v_j - m_i(t)) & \forall i \in N_w(t) \\ m_i(t+1) &= m_i(t) & \text{altrimenti} \end{aligned}$$

$$\alpha(t) \in [0,1] \quad N_w(t) = \text{neighborhood del nodo } w$$

Considerazioni sull'Algoritmo(1)

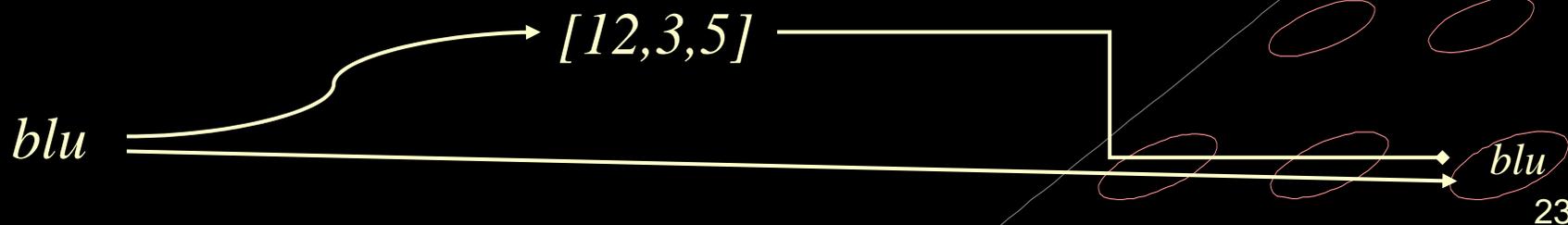
- *Inizializzare casualmente i nodi è il caso più generale : diversamente significherebbe specificare esattamente da quali cluster vogliamo partire*
- *Solitamente la distanza considerata in 2.1 è quella Euclidea (ma si potrebbe usare ad es. la edit-distance)*
- *$N_w(t)$ decresce nel tempo*
 - *all'inizio comprende tutti i nodi (organizzazione generale della mappa)*
 - *alla fine nessuno o gli immediati vicini (specializzazione dei singoli nodi)*

Considerazioni sull'Algoritmo(2)

- *Se $N_w(t) = \emptyset$ otteniamo l'algoritmo K-means*
- *Poichè l'input è rappresentato come un vettore le SOM risultano adatte (così come avviene nel cervello umano) a trattare tipi di dati diversi (ad es. testi e immagini) contemporaneamente*
- *Infatti sono molto usate nella costruzione di neurocontrollori ad es. per la coordinazione visuo-motoria*

Considerazioni sull'Algoritmo(3)

- *Per etichettare :*
 - *I nodi : è sufficiente, una volta terminato l'algoritmo far processare ancora una volta tutti i vettori input alla mappa ed etichettare ogni volta il nodo vincitore con il nome dell'oggetto rappresentato da tale vettore*
 - *Aree della mappa (feature extraction) : la questione è più complessa, vedremo alcune proposte (LABELSOM ^[46])*
 - *Esistono però tecniche per disegnare le aree sulla mappa*



Considerazioni sulla Complessità

I bottleneck dell'algoritmo sono :

- *Ricerca del vincitore*
 - *Tree Structured SOM ^[31] : albero di SOM dove a livelli più bassi corrispondono via via regioni più specifiche della mappa (si perdono le relazioni tra cluster diversi)*
 - *Hierarchical SOM ^[31] : ogni SOM tratta una ristretta parte del vettore input e poi tutti i risultati vengono combinati in una mappa più grande*
 - *Non sappiamo però se sono ottimali. Infatti sono stati proposti altri nuovi metodi ^[27] anche recentemente*
- *Dimensione dei vettori ^[26] (vedi dopo)*

Considerazioni sulla Complessità

I bottleneck dell'algoritmo sono :

- *Ricerca del vincitore*
 - *Tree Structured SOM ^[31] : a livelli più bassi corrispondono via via regioni più specifiche della mappa*
 - *si perdono le relazioni tra cluster diversi*
 - *Hierarchical SOM ^[31] : ogni SOM tratta una ristretta parte del vettore input e poi tutti i risultati vengono combinati in una mappa più grande*
 - *Il problema è ancora aperto perchè non sappiamo se i metodi sono ottimali ^[27]*
- *Dimensione dei vettori ^[26] (vedi dopo)*

Curiosità

Inventate da Teuvo Kohonen nei primi anni '80 [32], la loro prima applicazione è stata speech-to-text transformation

L'algoritmo è generale e può essere applicato in molti contesti (esistono 5384 articoli sulle SOM [22]):

- *Speech Recognition [50]*
- *Grammatical Inference*
- *Data mining e Information Retrieval [25]*
- *Bioinformatica, Robotica (ad es. neuro controllori) ecc...*
- *Topic Identification in Natural Language Dialogues [41]*
- *Associare testi a immagini...*

Word Category Maps : Un Esempio

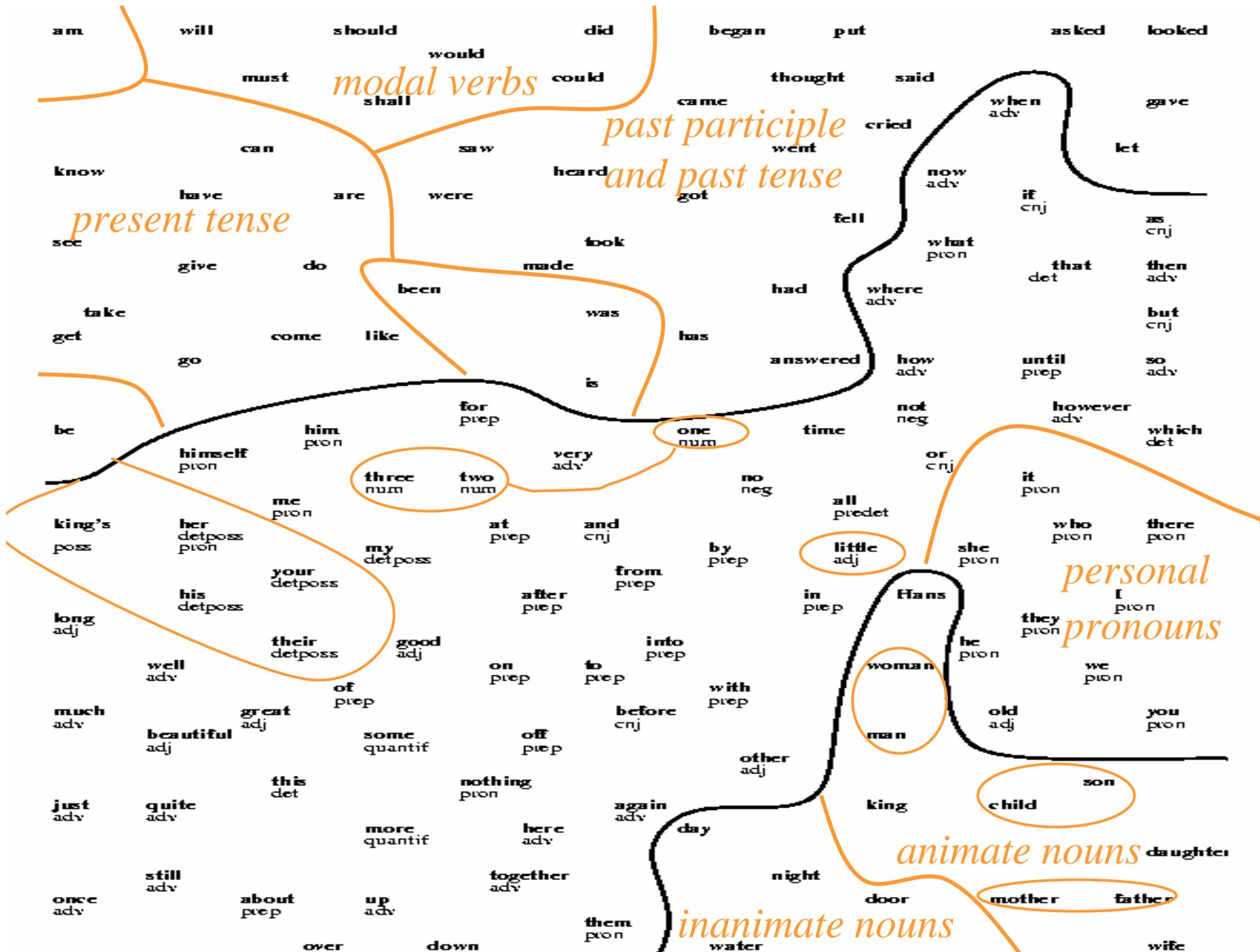
- *Dati presi da 200 fiabe dei fratelli Grimm ^[14]*
 - *Dimensione testo : 250 000 parole*
 - *Dimensione vocabolario : 7000 parole*
 - *Linguaggio non formale e vario, le fiabe hanno contesti molto diversi*
- *Preprocessing :*
 - *Eliminata punteggiatura, gli articoli*
 - *Maiuscole trasformate in minuscole*
 - *Considerate solo le 150 parole più frequenti*

Grimm Tales : Preprocessing

- *Input della rete è una serie di triple*
 - *[parola precedente , parola , parola successiva]*
- *Fra la grafia delle parole e la loro semantica non c'è correlazione → non deve esserci correlazione nemmeno nella loro rappresentazione semantica → vettori devono essere ortogonali tra loro.*
- *Per ogni parola si usa un vettore di 90 numeri reali → ogni vettore d'input ha dimensione 270*
- *Ecco perchè la dimensione dei vettori può essere un bottleneck per l'algoritmo!*

Grimm Tales : Implementazione

- *Per velocizzare l'algoritmo*
 - *In generale esistono tecniche per ridurre notevolmente la dimensione di un vettore [26]*
- *In questo caso :*
 - *Non si è considerata ogni singola tripla come input*
 - *Nemmeno la media dei contesti (solo 150 input)*
 - *Ma per ogni parola si è costruita una piccola SOM 2X2 che fornisce 4 medie di contesti per ogni parola*



Conclusioni

- *Indipendentemente dalla qualità ottenuta per la mappa la vera innovazione delle SOM è il fornire la semantica dei termini in base al solo contesto*
- *Inoltre, ci sono curiose somiglianze tra la mappa di quest'ultimo esempio e la mente umana:*
 - *ad es. Caramazza ^[4] ha mostrato che*
 - *alcune persone non riescono ad associare nomi agli oggetti corrispondenti*
 - *ma solo per oggetti animati*
 - *nel cervello le due categorie sono separate come nella mappa precedente!*

Prototipi Adattivi

- *Prototype Theory of concepts* ^[42] ^[47] : *invece che da una serie di attributi un concetto è rappresentato da una struttura prototipo*

- *Più un oggetto è “vicino” a un prototipo più è alto il valore della sua funzione di appartenenza a quel prototipo*

- *I nodi della SOM possono essere considerati prototipi adattivi*

- *Alcuni lavori di altri autori (MacWhinney ^[45] e Gärdenfors ^[8]) si rifanno al concetto di prototipi adattivi*

SOM come Modello di Memoria

• *Difatti sono:*

• *Dinamiche*

• *Associative (con i vicini nella mappa)*

• *Gli elementi (i nodi) possono considerarsi prototipi adattivi*

• *L'input al tempo t dipende dall'organizzazione della mappa in quel momento, cioè dagli input precedenti*



cioè dall'esperienza passata

E la Soggettività?

- *Un modo per trattare la soggettività è ^[19] :*
- *Input della SOM = [Parola, Contesto, Identità Soggetto]*
 - *In questo modo la SOM impara a riconoscere l'uso di una stessa parola fatto da persone diverse*
 - *La SOM può essere usata per scegliere la parola più adatta da usare in un determinato contesto.*
- *L'apprendimento diventa più difficile*
 - *Bisogna evitare che l'elemento soggettività influenzi troppo la formazione della mappa*

Difetti delle SOM

Ovvero i parametri dell'algoritmo :

- *Numero di unità prefissato (Neural gas networks ^[6])*
- *Neighborhood function da calibrare “a mano” (GTM ^[3])*
- *Impossibilità di prevedere e sintetizzare il risultato ^[9]
(rules extraction o anche soltanto ricavare una densità di probabilità)*
- *Difficoltà nel valutare le performances tra diverse esecuzioni dell'algoritmo ^[12]*
- *Non è dimostrato matematicamente che oggetti simili finiscano in punti vicini della mappa*

Qualità di una Mappa

- *Calcolare L'errore medio nel classificare un'istanza di input : $E\{ || v_j - m_{winner}(v_j) || \}$*
- *Kaski e Lagus ^[23] : si confronta la media delle distanze(come numero di nodi)tra il nodo dove è mappato un vettore su una mappa e sull'altra*
 - *Il metodo si applica più che altro per stimare le differenze prodotte dalle diverse scelte di un parametro*
- *Confrontare le prestazioni su alcuni insiemi di oggetti dei quali sappiamo a priori la relazione di somiglianza (dove dovrebbero finire sulla mappa)*

Qualità di una Mappa 2

- *Valutazione di un esperto*
- *Se consideriamo le SOM come un modello di memoria come proposto da Honkela ^[20] valutare con un giudizio assoluto la qualità di una mappa in certi contesti può non essere possibile*

LABELSOM

- *Permette di scegliere le parole che meglio caratterizzano un nodo* ^[46]
- *In pratica ciò che si fa è selezionare le parole che in media risultano essere più usate da tutti i documenti che appaiono in quel nodo*
- *Esistono altri metodi, tutti basati su simili analisi statistiche per la mappatura di aree più vaste*
 - *Comunque il problema, dato il proliferare di articoli sull'argomento, sembra essere ancora non risolto in maniera ottima*

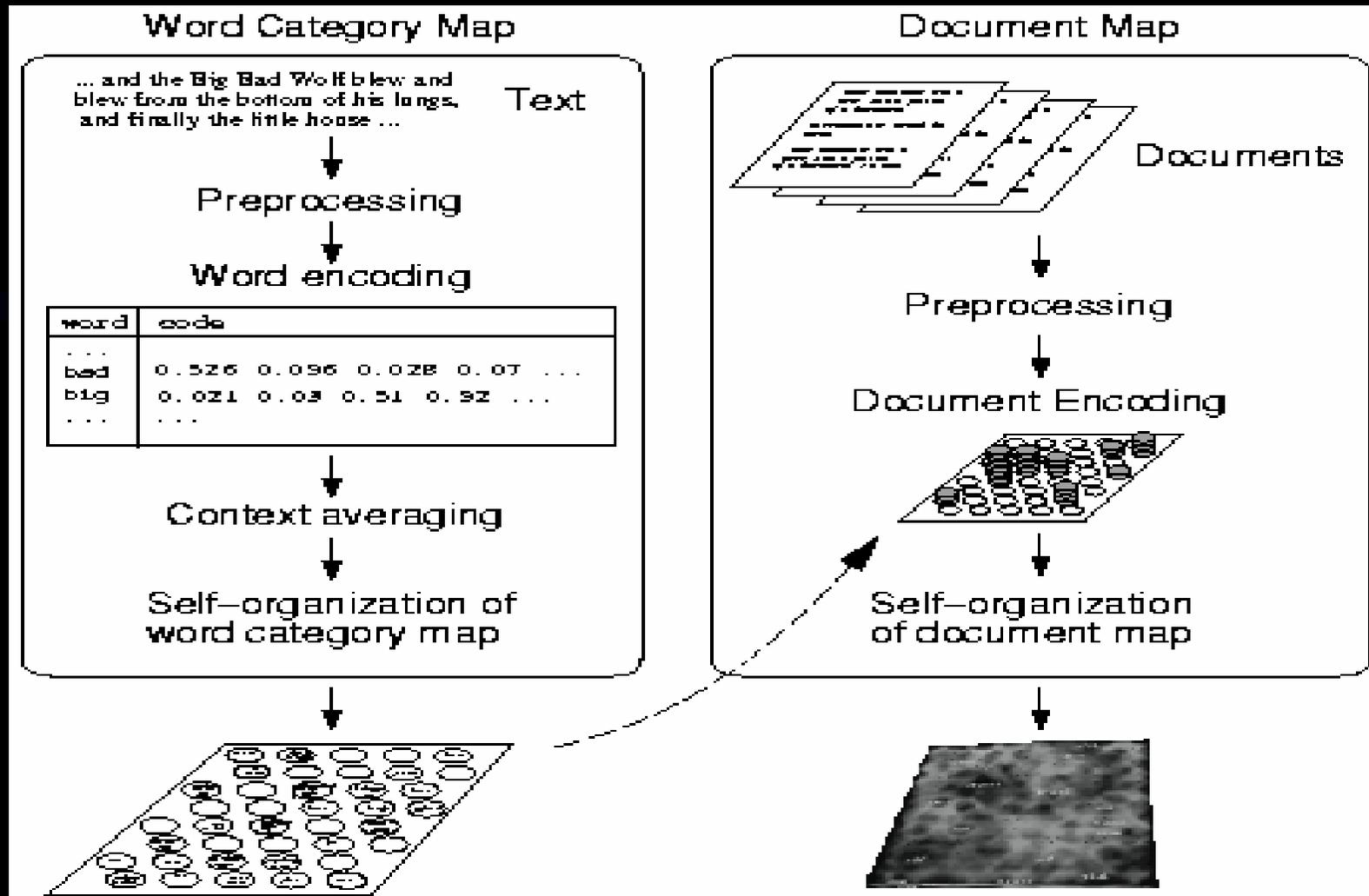
Information Retrieval tradizionale

- *Keyword Search* : vengono recuperati tutti quei documenti che contengono le parole della query (vengono usati i classici connettivi booleani)
 - *I risultati non sono classificati in ordine di importanza*
 - *L'uso di sinonimi comporta risultati completamente differenti*
- *Vector Space Model* : la query e i documenti immagazzinati sono rappresentati da un vettore $v \in \mathbb{R}^n$:
 - *n = numero di parole del documento o della query*
 - *$v[i]$ = funzione della frequenza del termine nel documento*

Information Retrieval tradizionale

- *Varianti del Vector Space Model sono usate in pratica da moltissimi motori di ricerca, ma i risultati non sono molto soddisfacenti perché:*
 - *Non viene in genere tenuta in considerazione la semantica delle parole → documenti sbagliati o omessi*
 - *Se una persona non conosce bene l'argomento non sarà in grado di interrogare in maniera efficace il motore di ricerca*
 - *Non permettono di fare browsing tra i risultati*
- *Google : fornisce buoni risultati ma si basa su un fenomeno di auto-organizzazione dei link del web molto simile a quello delle SOM! (<http://webselforganization.com>)*

Struttura di WEBSOM



Il Problema del Vocabolario

- *Le persone tendono ad usare parole differenti per descrivere lo stesso concetto*
 - *Furnas ^[7] ha mostrato in un esperimento dove si dovevano scegliere parole per descrivere concetti, come lo stesso termine venga usato da persone differenti con una probabilità minore del 20%*

Websom : Perché questa struttura?

- *Uno dei ruoli della Word Category Map è quindi quello di “uniformare l’input” evitando le fluttuazioni nei risultati classiche della ricerca per keywords ed eludendo il problema del vocabolario*



- *Inoltre usando una Word Category Map (calcolata off-line) si riduce la dimensione del vettore-documento (parole simili sono rappresentate dallo stesso nodo → da un solo elemento nel vettore-documento)*

Preprocessing : codifica documenti

- *Un vettore con la rappresentazione di tutte le parole*
- *Vector Space Model :*
 - *Creazione del vocabolario*
 - *Eliminazione delle stop-words (il,e,del...)*
 - *Eliminazione delle parole poco o troppo frequenti*
 - *Assegnamento di pesi ad ogni componente del vettore documento:*
 - *0-1 a seconda che la parola appaia o meno nel documento*
 - *Frequenza del termine nel documento*
 - *Inverso del n° di documenti in cui il termine occorre*

Analisi delle Fasi dell'Algoritmo

- *Due tipi di Websom :*

- *Unsupervised : la mappa dei documenti viene organizzata usando l'algoritmo SOM usuale*

- *Supervised : ad ogni input della SOM si associa un'ulteriore informazione : a quale classe dovrebbe appartenere*

Complessità

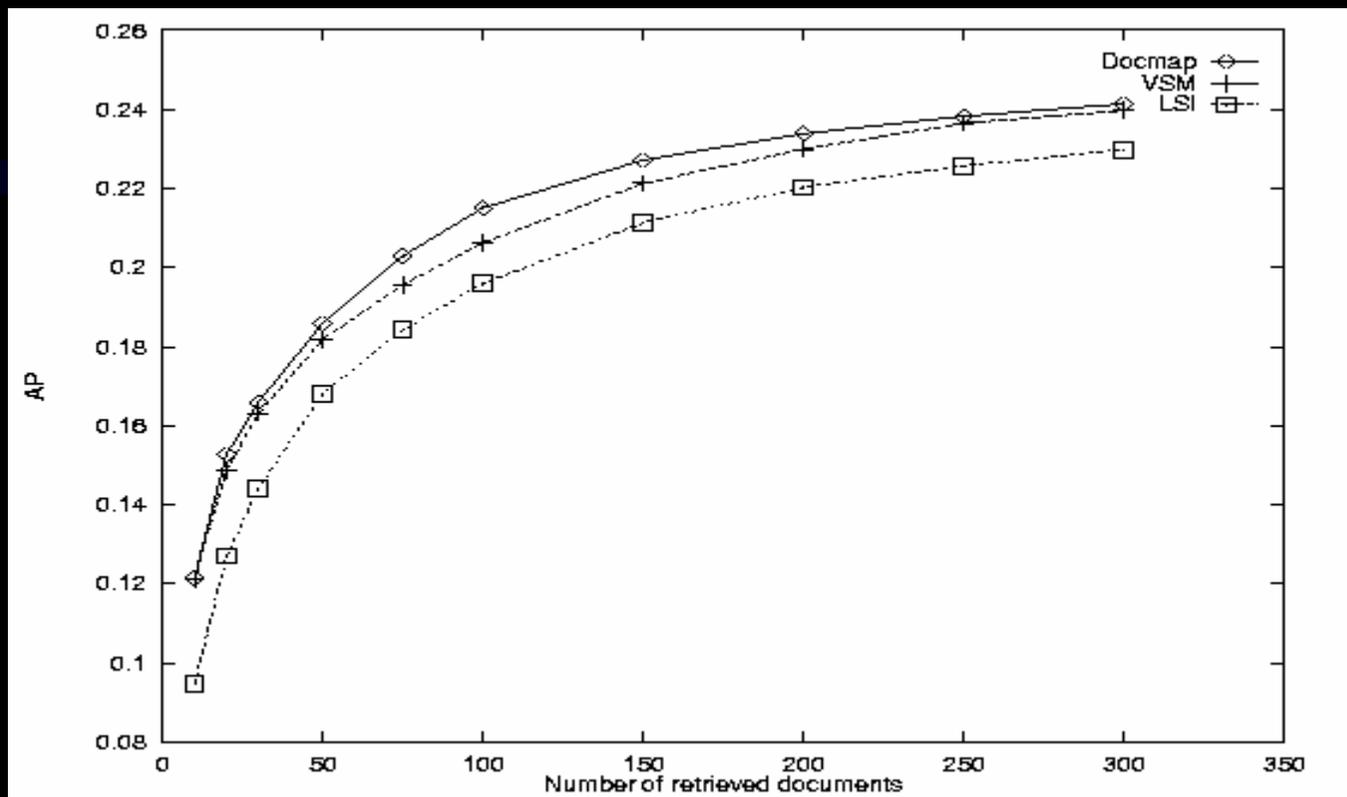
- *OFF-LINE* : il calcolo delle due mappe è un compito assai gravoso ma, con alcuni “trucchi” a cui abbiamo accennato precedentemente, è possibile ridurre notevolmente il carico di lavoro
- *ON-LINE* : all’arrivo di una query tutto ciò che rimane da fare è
 - *Calcolare l’istogramma dell’input (la serie di nodi winner, uno per ogni parola differente della query) :*
 - *può essere fatto in maniera estremamente veloce costruendo una look-up table con chiave l’input e valore il nodo corrispondente*
 - *Trovare i nodi della document map più simili (vedi dopo)*₄₇

Problemi dell'Algoritmo

- *I bottleneck dell'algoritmo rimangono gli stessi (ricerca del vincitore e dimensione dell'input)*
- *Le dimensioni della mappa rendono ancora più evidenti questi problemi, per cui sono stati introdotti altri trucchi:*
 - *Si allena prima una mappa più piccola che poi si usa per definire un intorno dove troveremo il winner*
 - *si deve perciò eseguire soltanto una ricerca locale*
 - *Salviamo $winner(x,n)$ e all'iterazione successiva andremo a cercare $winner(x,n+1)$ in un intorno del $winner(x,n)$*

Risultati

- *CISI collection : 1460 documenti e 76 query di cui si conoscono i migliori risultati (http://local.dcs.gla.ac.uk/idom/ir_resources/text_collection/cisi)*

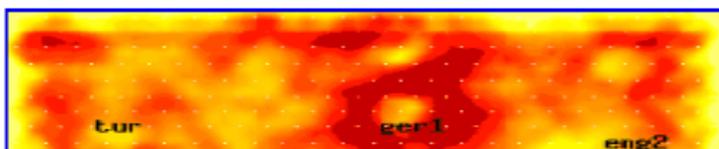


Risultati

<i>Anno</i>	<i>Numero documenti</i>	<i>Numero nodi</i>
<i>1996</i>	<i>5.000</i>	<i>768</i>
<i>1997</i>	<i>1.000.000</i>	<i>100.000</i>
<i>1999</i>	<i>7.000.000</i>	<i>1.000.000</i>

- *Il metodo risulta essere del tutto scalabile (per dimensioni e tipo)*
- *Gli autori hanno dichiarato che l'organizzazione prodotta da Websom è paragonabile a quella di YAHOO!*
- *Comunque la precisione di WEBSOM è stimata solo intorno al 67%*

WEBSOM map



Explanation of the symbols on the map

tur - Turkic, Greek, Basque etc.
ger1 - German
eng2 - English accent

WEBSOM zoomed map



Click arrows

to move to neighboring areas on the map, and to move up to the overall view.

Explanation of the symbols on the map

eng1 - English (Singapore, Welsh etc.)
eng3 - English (European)
eng4 - English (Internet, Singapore, etc.)

WEBSOM node t13



Click arrows

to move to neighboring nodes on the map.

[Instructions](#)

[Re: Dutch and English accents](#) ♦ Stewart McKenna, 21 Jun 1995, L

[Re: How language evolves-- "emails"](#) ♦ Iorn Barger, 23 Jun 1995,

[Re: What is "Online Mayan"](#) From: 73513.2350@compuserve.com (D Gary Grady)

[Re: english or](#) Newsgroups: sci.archaeology,sci.lang

[Re: english or](#) Subject: Re: Linguistic looniness (was Re: Otto Muck

[Re: english or](#) Date: Wed, 29 Nov 1995 01:17:21 GMT

[Re: english or](#) Lines: 16

[Wanted: Book](#) petrich@netcom.com (Loren Petrich) wrote:

[Re: Linguistic](#) > For example, Old English is essentially a fo

[Re: What Is S](#) >and OE texts are only 1500-1000 years old. . . .

Just to be picky, Old English goes back rather more started evolving into Middle English almost 900 year (1340-1400) wrote in full-blown Middle English 1600 couple of centuries later Shakespeare was writing in English.

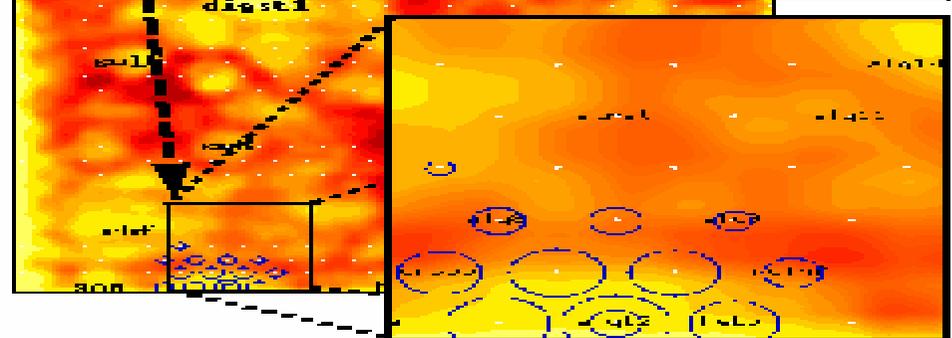
Applicazioni di WEBSOM

- *WEBSOM permette diversi tipi di ricerca :*

- *Browsing semplice*
- *Query con parole*
- *Query con frasi*
- *Query con interi documenti*

Content-directed search

Does anybody know of good books about hebbian learning? I am especially interested in describing the evidence of Hebbian type neural plasticity. Some information about the relations of synapt. strength and neural spreading in Hebbian learning would also be welcome.



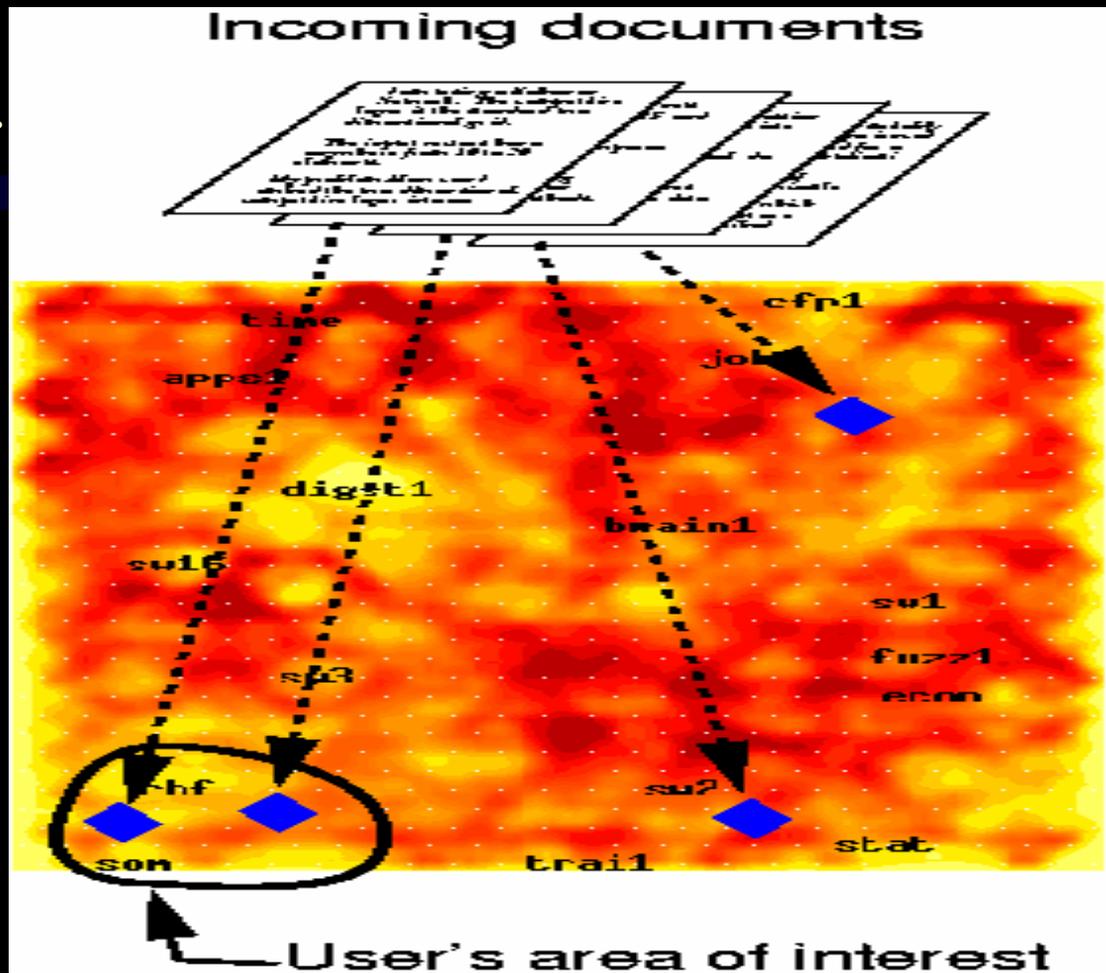
Click arrows to move to neighboring nodes on the network.
Instructions

[Competitive learning](#) • Dr. Shlomo (Sean) Engelvor
[hebbian learning](#) • tareq i. j. albahar, 13 Sep 1995 [1]
[Re: hebbian learning](#) • Kevin Spencer, 18 Sep 1995
[Re: Reinforcement learning - predicting reinforcem](#)
[Re: Reinforcement learning - predicting ret](#) • Tho
[\[Q\] Books on reinforcement learning ?](#) • Luc-Jean
[DeSieno's 'conscience' in Kohonen SOM](#) • Rudy S.

Applicazioni di WEBSOM

- Può essere applicato a diversi tipi di collezioni [36]:

- Filtri [48]:



Alcuni Esempi d'uso delle SOM

- *In tutti gli esperimenti seguenti le SOM vengono usate per organizzare collezioni di documenti :*
- *WEBSOM per l'esplorazione del NewsGroup comp.ai.neural-nets ^[21] e vari altri*
- *ET-MAP dell'università dell'Arizona : una SOM per organizzare un sottoinsieme di pagine web per argomento (<http://ai.bpa.arizona.edu/>)*
- *Lin ha costruito una SOM basata sui titoli di articoli scientifici*
- *Alcune demo si possono trovare all'indirizzo:*
 - *<http://websom.hut.fi/websom/>*

Bibliografia (1)

Tutta la documentazione da cui è stata tratta questa presentazione è disponibile in linea (in aggiunta ci sono alcuni riferimenti per argomenti integrativi) :*

- [1] **Arnulfo P. Azcarraga and Teddy Yap Jr. SOM-based Methodology for Building Large Text Archives**
- [2] **R. Basili and M. Cammisa. Self Organizing Maps e disambiguazione semantica**
- [3] **Christopher M Bishop, Markus Svensén, Christopher K.I. Williams. GTM:A Principled Alternative to the Self-Organizing Map. To appear in Advances in Neural Information Processing Systems 9, MIT Press, 1997**
- [4] **Caramazza, A., Hillis, A., Leek, E.C., and Miozzo, M. Mapping the mind: Domain specificity in cognition and culture, chapter The Organization of lexical knowledge in the brain: Evidence from category and modality-specific deficits, pages 68-84. Cambridge University Press, Cambridge.**
- [5] **Elmann, J. L. Language as a dynamical system. In Robert F. Port & T. van Gelder (Eds.) Mind as Motion: Explorations in the Dynamics of Cognition. Cambridge, MA: MIT Press, 1995. Pp. 195-223.**

Bibliografia (2)

- [6] *Fritzke, B. Growing Cell Structures - A Self-organizing Network for Unsupervised and supervised Learning. ICSI TR-93-026, 1993. Neural Networks 7(9):1441-1460, 1994a.*
- [7] *Furnas, G W., Landauer, T.K., Gomez, L.M., and Dumais, S.T. The vocabulary problem in human-system communication. Communications of the ACM, 30(11):964-971.*
- [8] *Gärdenfors, P. Mental representation, conceptual spaces and metaphors. Synthese, 106:21-47.*
- [9] *Barbara Hammer , Andreas Reichtien , Marc Strickert , Thomas Villmann. Rule Extraction from Self-Organizing Networks*
- [10] * *Haykin S. Neural Networks-A Comprehensive Foundation. Prentice Hall, 2nd ed., 1999*
- [11] *Timo Honkela. Adaptive and Holistic Knowledge Representations Using Self-Organizing Maps*
- [12] *Timo Honkela. Comparisons of Self-Organized Word Category Maps*

Bibliografia (3)

- [13] **Timo Honkela.** **Connectionist Analysis and Creation of context for Natural language understanding and Knowledge Management**
- [14] **Timo Honkela, Ville Pulkki, and Teuvo Kohonen.** **Contextual relations of words in Grimm tales, analyzed by self-organizing map.** *In F. Fogelman-Soulie and P. Gallinari, editors, Proc. ICANN'95, Int. Conf. on Artificial Neural Networks, volume II, pages 3-7, Nanterre, France, 1995. EC2.*
- [15] **Timo Honkela.** **Learning to Understand General Aspects of Using Self-Organizing Maps in Natural Language Processing.** *Proceedings of the CASYS'97, Computing Anticipatory Systems, Liège, Belgium, August, 1997, in press.*
- [16] **Honkela, T., Kaski, S., Lagus, K., and Kohonen, T. (1996b).** **Newsgroup exploration with WEBSOM method and browsing interface.** *Technical Report A32, Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo.*
- [18] **Timo Honkela.** **Self-Organizing Maps in symbol Processing.** *Hybrid Neural Systems, Stefan Wermter, Ron Sun (eds.), Springer, Heidelberg, 2000, pp. 348-362.*

Bibliografia (4)

- [19] **Timo Honkela. Self-Organizing Maps in Natural Language Processing.** *PhD thesis, Helsinki University of Technology, Espoo, Finland, 1997*
- [20] **Timo Honkela , Teemu Leinonen , Kirsti Lonka and Antti Raike. Self-Organizing Maps and Constructive Learning.** *Proceedings of ICEUT'2000, IFIP, Beijing, August 21-25, 2000, pp. 339-343.*
- [21] **Timo Honkela, Samuel Kaski, Krista Lagus and Teuvo Kohonen. WEBSOM - Self-Organizing Maps of Document Collections.** *In Proceedings of WSOM'97, Workshop on Self-Organizing Maps, Espoo, Finland, June 4-6, pages 310-315. Helsinki University of Technology, Neural Networks Research Centre, Espoo, Finland.*
- [22] **Samuel Kaski, Jari Kangasz, Teuvo Kohonen. Bibliography of Self-Organizing Map (SOM) Papers: 1981 1997**
- [23] **Samuel Kaski, Krista Lagus. Comparing Self-Organizing Maps.** *Proceedings of ICANN96, International Conference on Artificial Neural Networks, lecture notes in Computer Science vol. 1112, pag 809-814, Springer Berlin*

Bibliografia (5)

- [24] *Kaski, S., Honkela, T., Lagus, K., and Kohonen, T. (1996). Creating an order in digital libraries with self-organizing maps. In (to appear): Proc. of World Congress on Neural Networks (WCNN-96).*
- [25] *Samuel Kaski. Data Exploration with Self-Organizing Maps. PhD thesis, Helsinki University of Technology, Espoo, Finland, 1997*
- [26] *Samuel Kaski. Dimensionality Reduction by Random Mapping:Fast Similarity Computation for Clustering. Proceedings of IJCNN'98, International Joint Conference on Neural Networks, vol. 1, pp. 413-418, Piscataway, NJ: IEEE Service Center, 1998.*
- [27] *Samuel Kaski. Fast Winner Search for SOM-Based Monitoring and Retrieval of High Dimensional Data. In Proceedings of ICANN'99, Ninth International Conference on Artificial Neural Networks, Edinburgh, UK, 7-10 September, to appear*
- [28] *S. Kaski, K. Lagus, T. Honkela, and T. Kohonen. Statistical Aspects of the WEBSOM System in Organizing Document Collections Computing Science and Statistics 29:281-290, 1998 (Scott, D. W., ed.), Interface Foundation of North America, Inc.: Fairfax Station, VA*

Bibliografia (6)

- [29] **Teuvo Kohonen. Exploration of Very Large Databases by Self-Organizing Maps**
- [30] **Teuvo Kohonen, Samuel Kaski, Krista Lagus, Jarkko Salojärvi, Jukka Honkela, Vesa Paatero, and Antti Saalera. Self Organization of a massive document collection. *IEEE Transaction on Neural Networks*, vol. 11, NO. 3, May 2000**
- [31] **Teuvo Kohonen. Self-Organization of Very Large Document Collections: State of the Art. In Niklasson, L., Boden, M., and Ziemke, T., editors, *Proceedings of ICANN98, the 8th International Conference on Artificial Neural Networks*, volume 1, pages 65-74. Springer, London.**
- [32] **Kohonen T. Self-Organizing Maps. Springer, Berlin, Heidelberg**
- [33] **Teuvo Kohonen, Samuel Kaski, Krista Lagus, and Timo Honkela. Very Large Two-Level SOM for the Browsing of Newsgroups. *ICANN'96, Bochum, Germany, July 16-19, 1996. Lecture Notes in Computer Science*, vol. 1112, pages 269-274. Springer, Berlin.**

Bibliografia (7)

- [34] *Koikkalainen P. Fast Deterministic self-organizing maps.*In Fogelman Soulié, F. and Gallinari, P., editors, *Proceedings of ICANN'95, International Conference on Neural Networks, volume II, pages 279-284, IEEE Service Center, Piscataway, NJ.*
- [35] *Krista Lagus, Samuel Kaski, Timo Honkela, and Teuvo Kohonen. Browsing Digital Libraries with the Aid of Self-Organizing Maps.*In *Proceedings of the Fifth International World Wide Web Conference WWW5, May 6-10, Paris, France, volume Poster Proceedings, pages 71-79. EPGL.*
- [36] *Krista Lagus. Generalizability of the Websom Method to Document Collection of Various Types.*In *Proc. of 6th European Congress on Intelligent Techniques & Soft Computing (EUFIT'98), volume 1, pages 210-214, Aachen, Germany*
- [37] *Krista Lagus, Samuel Kaski. Keyword selection for characterizing text document maps.*

Bibliografia (8)

- [38] *Krista Lagus, Timo Honkela, Samuel Kaski, and Teuvo Kohonen. Self-Organizing Maps of Documents Collections:A New Approach to Interactive Exploration. In Simoudis, E., Han, J., and Fayyad, U., editors, Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pages 238-243. AAAI Press, Menlo Park, California.*
- [39] *Krista Lagus. Text retrieval using self-organized document maps*
- [40] *Krista Lagus. Text Mining with the Websom. PhD thesis, Helsinki University of Technology, Espoo, Finland, 2000*
- [41] *Krista Lagus and Jukka Kuusisto. Topic Identification In Natural Language Dialogues Using Neural Networks*
- [42] *Lakoff, G. Women, Fire and Dangerous Things. University of Chicago Press, Chicago.*
- [43] *Krister Lindén, Krista Lagus. Word Sense Disambiguation in Document Space.*
- [44] *Luttrell, S.P. Hierarchical Vector Quantization. IEEE proceedings, 136:405-413.*

Bibliografia (9)

- [45] *MacWhinney, B. Cognitive approaches to language learning, chapter Lexical Connectionism. MIT Press.*
- [46] *Anreas Rauber, Erich Schweighofer, Dieter Merkl. Text Classification and Labelling of Document Clusters with Self-Organizing Maps.*
- [47] *Rosch, E. Studies in cross-cultural psychology, vol. 1, chapter Human categorization, pages 3-49. Academic Press, New York*
- [48] * *Scholtes, J. C. Neural Networks in Natural Language Processing and Information Retrieval. PhD thesis, Universiteit van Amsterdam, Amsterdam, Netherlands.*
- [49] *Erich Schweighofer, Andreas Rauber, Michael Dittenbach. Improving the Quality of Labels for Self-Organising Maps Using Fine-Tuning*
- [50] *Roberto Togneri, M.D. Alder, Yianni Attikiouzel. Speech Processing using Artificial Neural Networks*

Bibliografia (10)

- [51] 5384 Works that have been Based on Self-Organizing Map(SOM) method developed by Kohonen Part I, authors from A to K
- [52] 5384 Works that have been Based on Self-Organizing Map(SOM) method developed by Kohonen Part II, authors from L to Z