

European
Research
Council



<http://ccnl.psy.unipd.it>

Dal connessionismo classico ai modelli generativi gerarchici

Marco Zorzi

**Computational Cognitive Neuroscience Lab
Università di Padova**

Connessionismo “classico”

Molti modelli connessionisti sono **implausibili** sia per l'*architettura neurale* che per le *procedure di apprendimento*

→ problematico se l'obiettivo è di simulare l'apprendimento ed il comportamento umano! Compromesso accettabile?

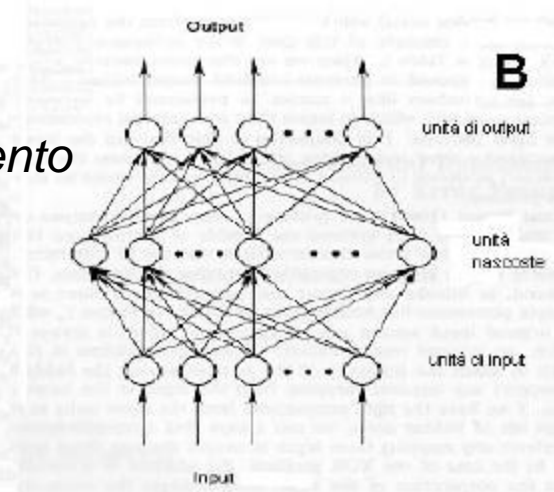
La tipica rete neurale:

- **architettura: multistrato feed-forward**

- le connessioni vanno solo da input ad output (S-R)
- non sono possibili effetti top-down non sono possibili
- perché il cervello è “pieno” di connessioni rientranti (ricorrenti) ?

- **algoritmo di apprendimento: back-propagation**

- molto potente ma biologicamente implausibile (il segnale d'errore è propagato all'indietro, dal dendrite del neurone ricevente, attraverso la sinapsi, verso l'assone del neurone trasmittente, e successivamente integrato e moltiplicato per la forza della sinapsi e la derivata della funzione di output)
- l'apprendimento supervisionato richiede sempre la presenza di un insegnante esterno (vettore di output desiderato per il calcolo del segnale di errore) che è utilizzato per aggiustare gradualmente i pesi fino a quando le prestazioni della rete raggiungono un livello considerato adeguato in base a qualche criterio esterno.



L'approccio probabilistico (Bayesiano) alla cognizione

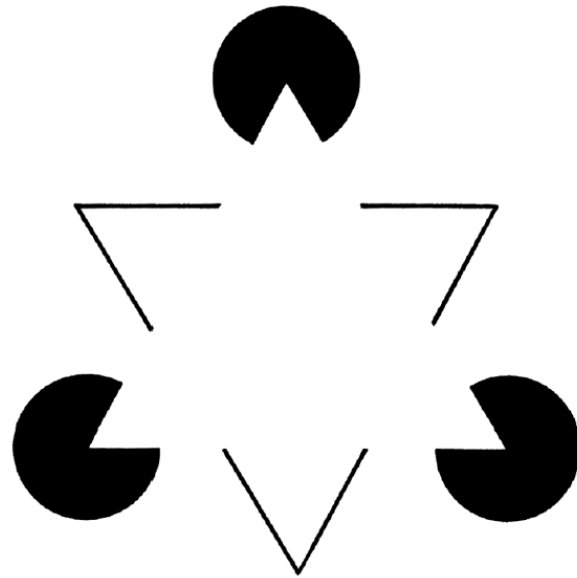
- I modelli Bayesiani della cognizione assumono che l'apprendimento umano e l'inferenza seguono approssimativamente i principi dell'inferenza probabilistica Bayesiana

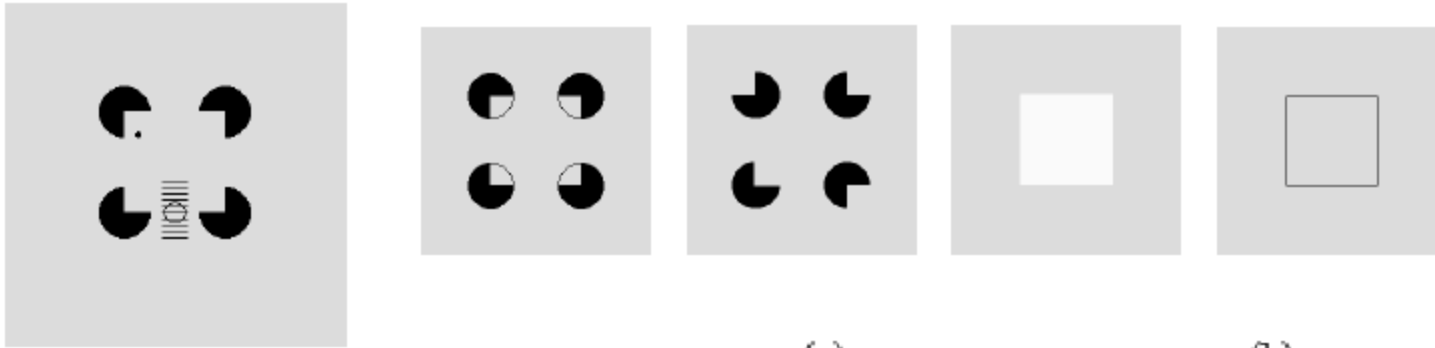
Un esempio: *La visione come problema di inferenza probabilistica* (cf. Helmholtz)

La nostra conoscenza a priori sulle scene visive viene combinata con le caratteristiche dell'immagine per inferire l'interpretazione più probabile. Data un'immagine, un modello generativo ci può dire quali scene è probabile abbiano generato l'immagine (ovvero, troviamo la scena che massimizza la probabilità a posteriori usando la regola di Bayes).

Un modello generativo probabilistico offre una spiegazione naturale di come le aspettative top-down dovrebbero essere utilizzate per disambiguare i dati sensoriali.

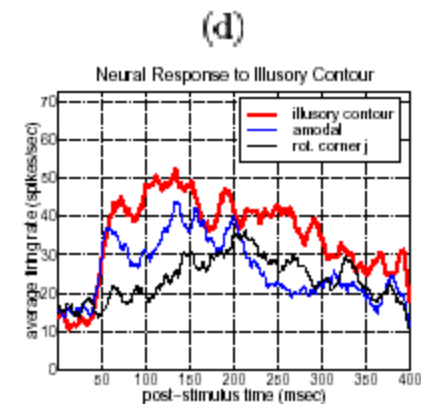
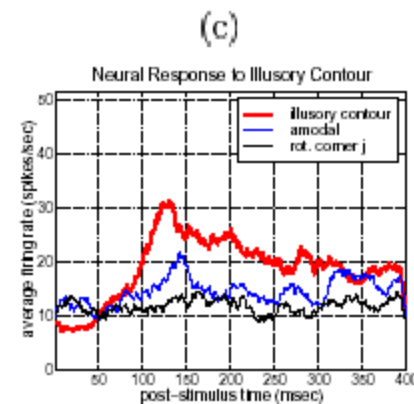
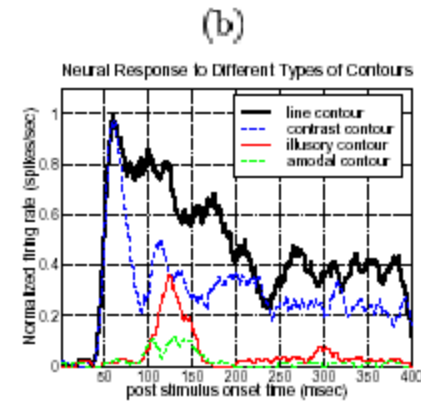
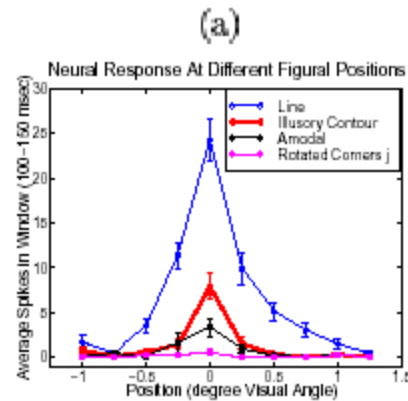
THE CAT





Correlati neurali dei contorni illusori

A. Il profilo spaziale della risposta di un neurone V1 ai contorni di quadrati reali ed illusori (finestra di 100-150 ms post-onset). **B.** Evoluzione temporale della risposta a contorni illusori in confronto ad altri tipi di contorno. **C.** Risposta media di una popolazione di 49 neuroni in V1. **D.** Risposta media di una popolazione di 39 neuroni in V2.



Lee & Nguyen, 2001, PNAS.

Modelli generativi

La formulazione dei modelli generativi proviene da un approccio di tipo statistico, che fonda le sue radici nel *teorema di Bayes*. Un modello generativo avanza delle ipotesi sui dati sensoriali (l'input), per poter modificare i parametri che lo caratterizzano.

- Un modello generativo è un modello probabilistico di come i dati sensoriali siano causati da sottostanti proprietà fisiche del “mondo”, e rappresenta una utile “funzione oggettiva” per l'apprendimento **non supervisionato**.
- L'apprendimento può essere visto come la massimizzazione della verosimiglianza dei dati osservati rispetto al modello generativo, il che equivale a scoprire modi efficienti di codificare l'informazione sensoriale
- rappresentano modelli plausibili dell'apprendimento corticale e forniscono ipotesi forti sul ruolo delle connessioni rientranti (feedback) e laterali nella corteccia

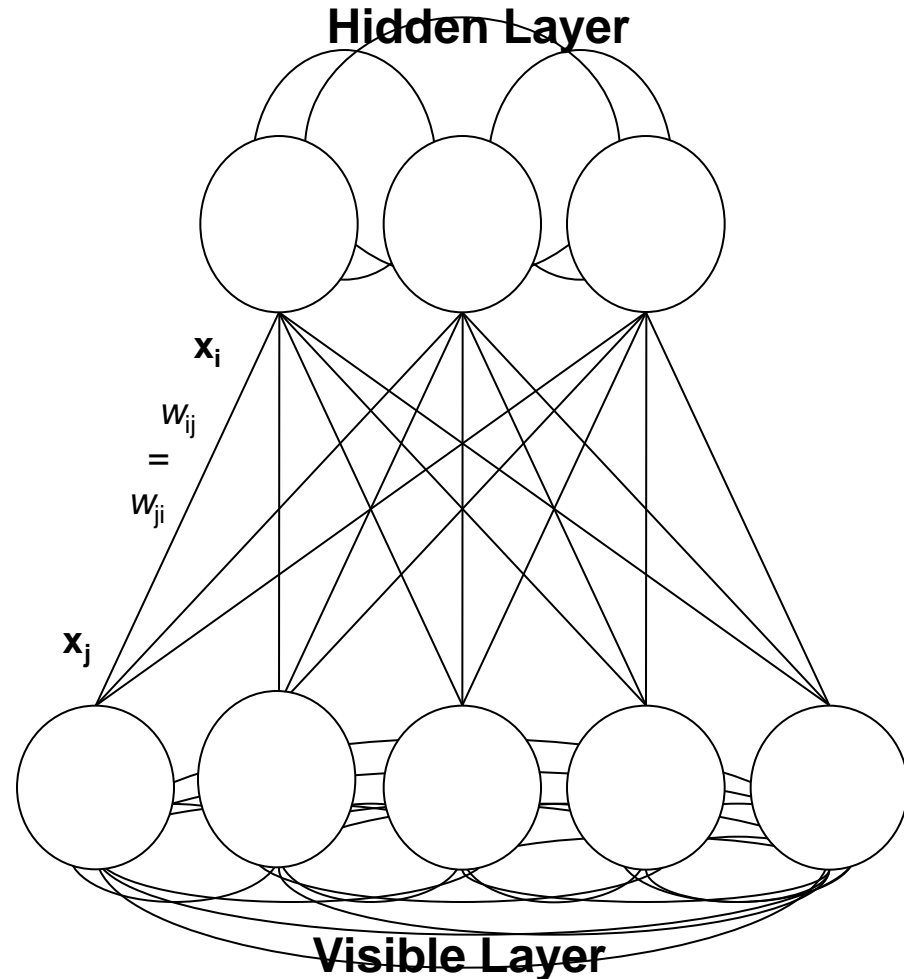
Quale formato per le rappresentazioni interne? (come rappresentare il mondo?)

- **non-lineare** (rappresentazioni lineari implicano la limitazione alla cattura di informazione nella covarianza pairwise delle variabili di input - l'informazione d'ordine superiore è invisibile alla rete).
- **distribuito** (rappresentazioni localistiche si basano sull'assunzione che ogni osservazione è generata da una singola variabile).

(NB: considerato un problema intrattabile fino a pochi anni fa!)

Boltzmann Machines (BM)

- Strato visibile: dati (può essere diviso tra input e output)
- Strato nascosto: rappresentazioni interne distribuite
- Architettura ricorrente: connessioni bidirezionali tra strati (bottom-up e top-down) e tra strati (connessioni laterali). Sono possibili molte varianti (restricted BMs).
- Apprendimento correlazionale (hebbiano) basato sulla minimizzazione della divergenza di Kullback-Liebler tra la distribuzione dei dati (fase positiva) e dopo un passo di ricostruzione dei dati (fase negativa). Approssimazione mean-field delle distribuzioni di probabilità (contrastive divergence mean-field learning; Welling & Hinton, 2001).
- L'apprendimento è senza supervisione (tutte le unità visibili sono libere durante la fase negativa). L'apprendimento consiste nel costruire un modello interno dei dati (modello generativo)
- Dinamica di attrattori per il recupero di pattern (memoria associativa). I tempi necessari per raggiungere stati stabili (equilibrio) è un'utile analogia dei tempi di reazione umani



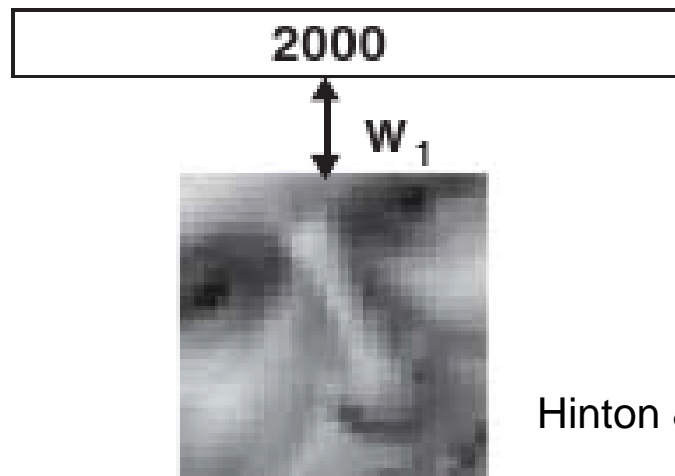
Modelli generativi

- Apprendimento non supervisionato di un modello interno dei dati (features o cause latenti).
- Funzione obiettiva: minimizzare la differenza contrastiva (Kullback-Liebler) tra i dati di input e la ricostruzione (top-down) dei dati

Es: Restricted Boltzmann Machines (RBM)

strato di input connesso in modo bidirezionale simmetrico con uno strato nascosto di feature-detectors, senza connessioni laterali.

- Principale limite: bassa complessità delle features apprese

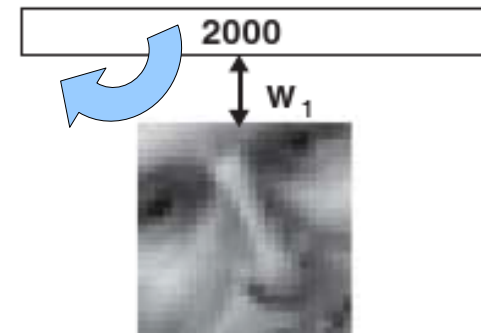
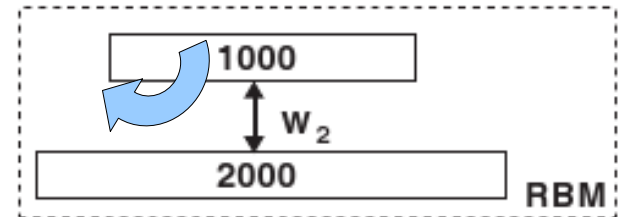
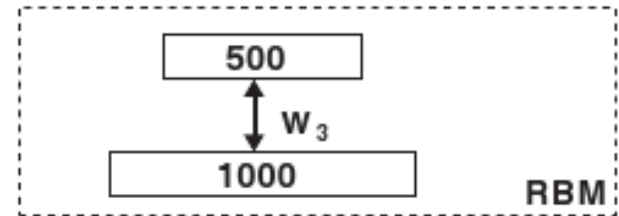


Hinton & Salakhutdinov, *Science* 2006

Modelli generativi gerarchici

(deep belief networks)

- Reti con strati nascosti organizzati in modo gerarchico (ridotti per convenienza ad uno “stack” di RBMs)
- “Deep learning” non supervisionato: apprendimento strato per strato
- Possono apprendere proprietà dell’input di ordine superiore con (feature-detectors sempre più complessi)



Esempi di applicazione alla simulazione dei processi cognitivi

- Dati comportamentali:
 - codifica ortografica delle parole (Di Bono)
 - sequenze di lettere / regole ortotattiche (Testolin)
- Dati neurofisiologici:
 - codifica dello spazio nei neuroni parietali (De Filippo, Cutini)
- Dati comportamentali e neurofisiologici
 - percezione e codifica della numerosità (Stoianov)

CONCLUSIONI

Modelli generativi distribuiti e non lineari (come la BM) :

- modelli più plausibili dell'apprendimento (non supervisionato) e della computazione corticale (indicazioni precise sul ruolo della connettività ricorrente nel cervello)
- superamento del connessionismo classico basato su mapping S-R (costruiscono modelli interni del "mondo")
- aumentare la complessità e la plausibilità neurale dei modelli generativi: sistemi organizzati in modo gerarchico
- buona corrispondenza con i dati empirici ai diversi livelli di descrizione, della psicologia alle neuroscienze

Sviluppi futuri:

- aumentare la plausibilità delle simulazioni: adottando i metodi della vita artificiale che permettono di considerare il ruolo dei fattori genetici e sociali - modelli generativi "*embodied*"