

Verona - 10 Giugno 2008

(Seminario integrativo del corso di MMB,
titolare M. Squassina, A.A. 2007/2008)

*Memoria associativa
e dinamica del riconoscimento:
introduzione al modello di Hopfield*

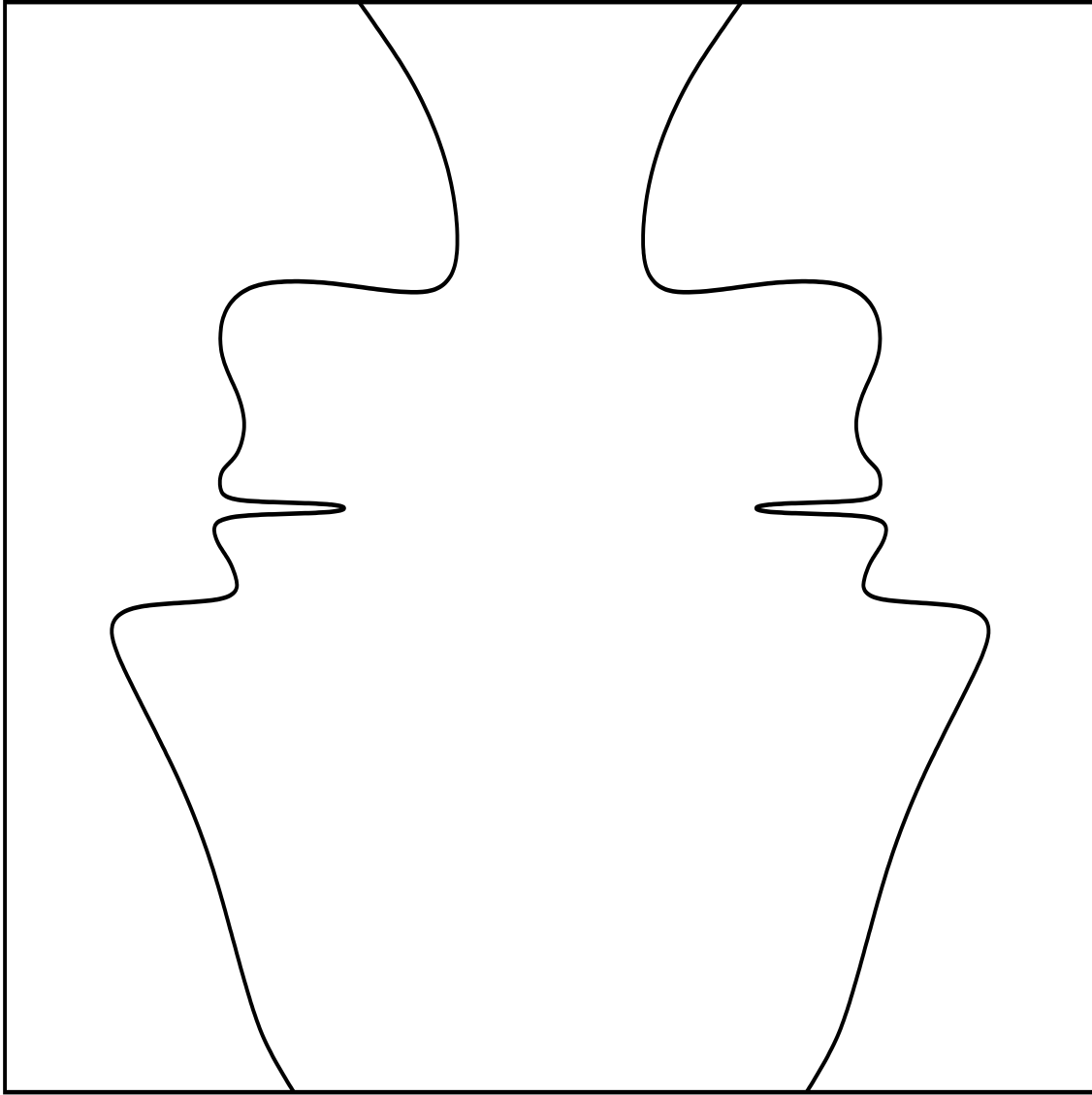
Antonio Ponno* †

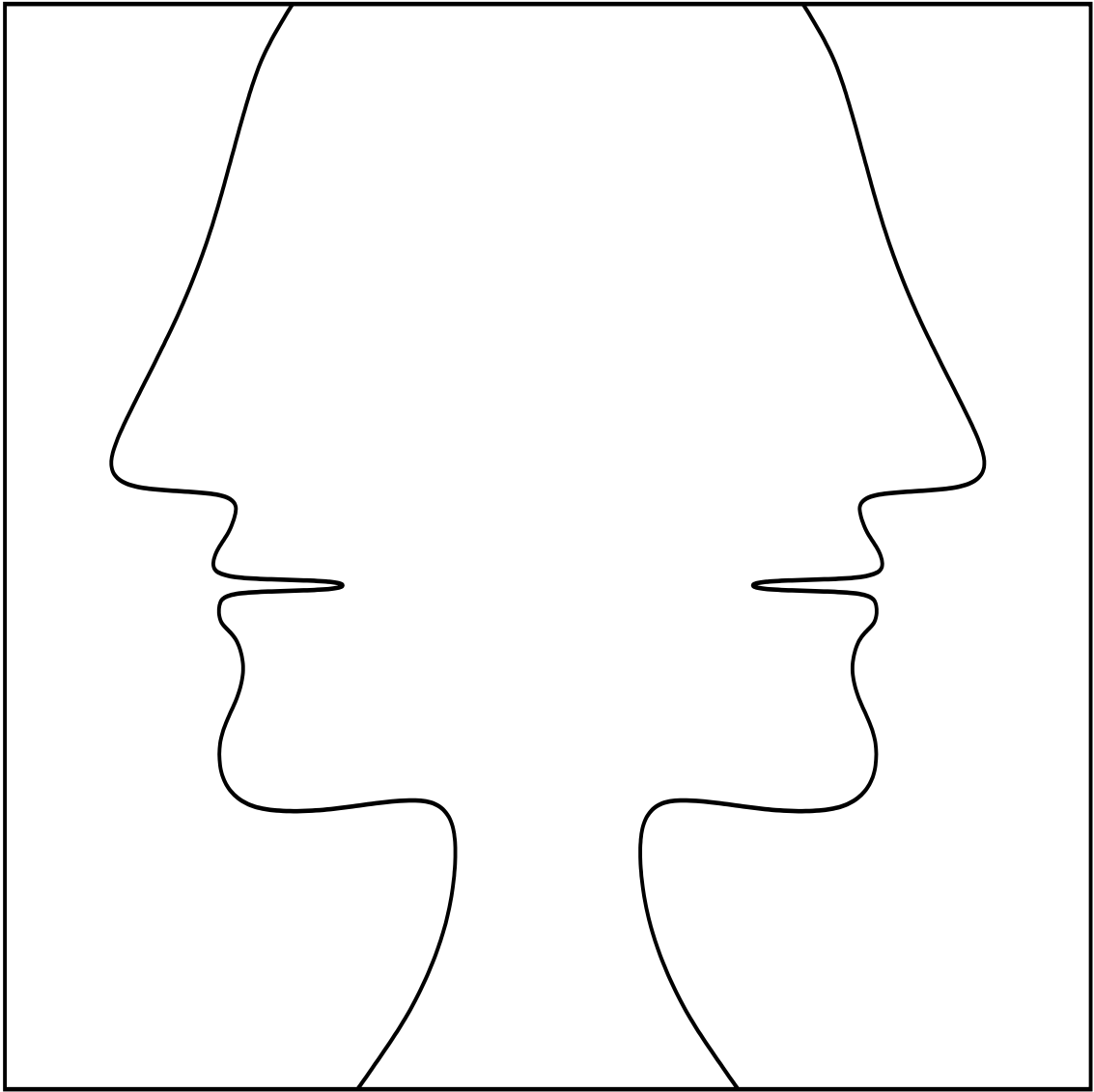
Università degli Studi di Padova
Dipartimento di Matematica Pura ed Applicata

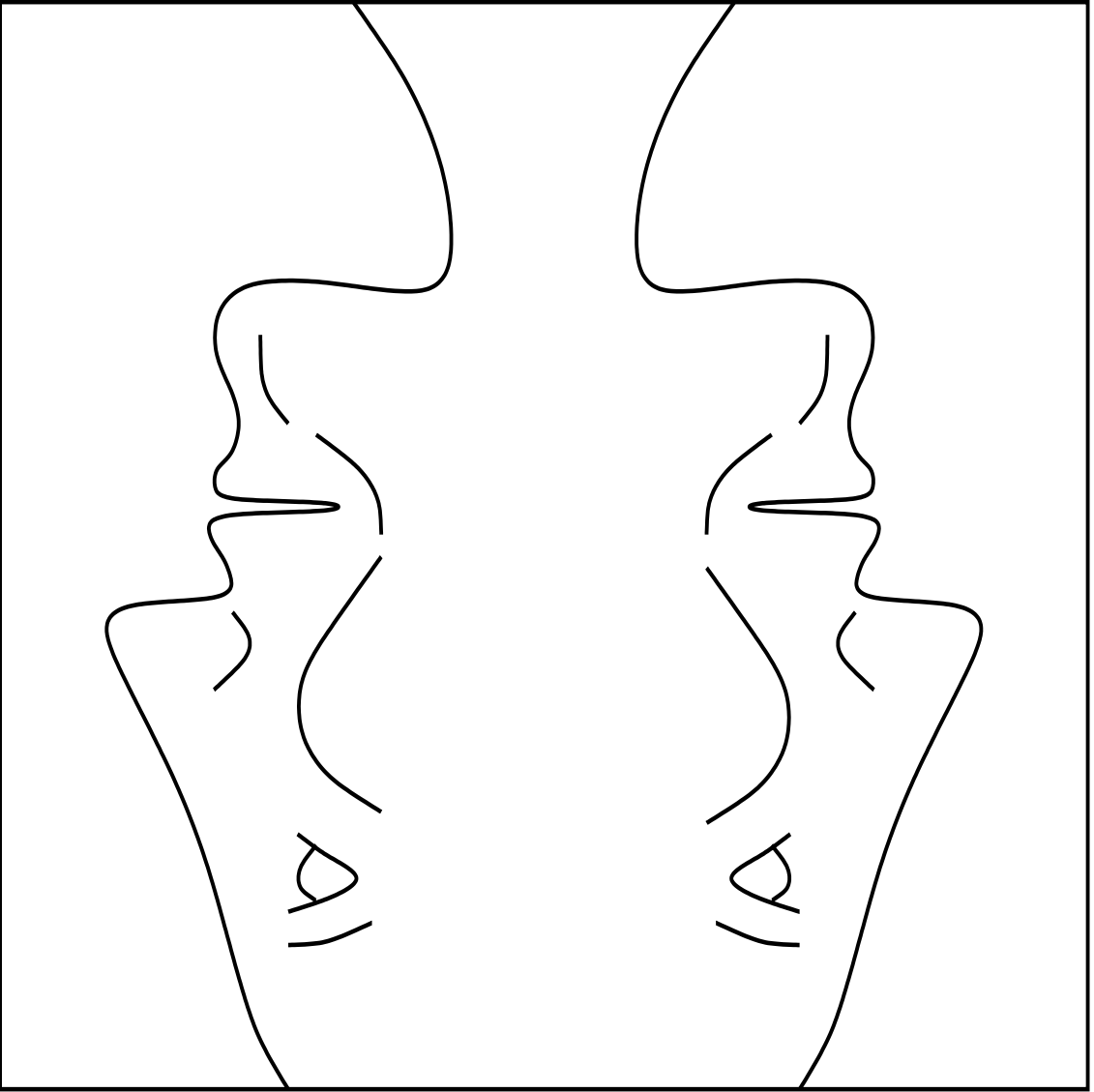
*E-mail: ponno@math.unipd.it

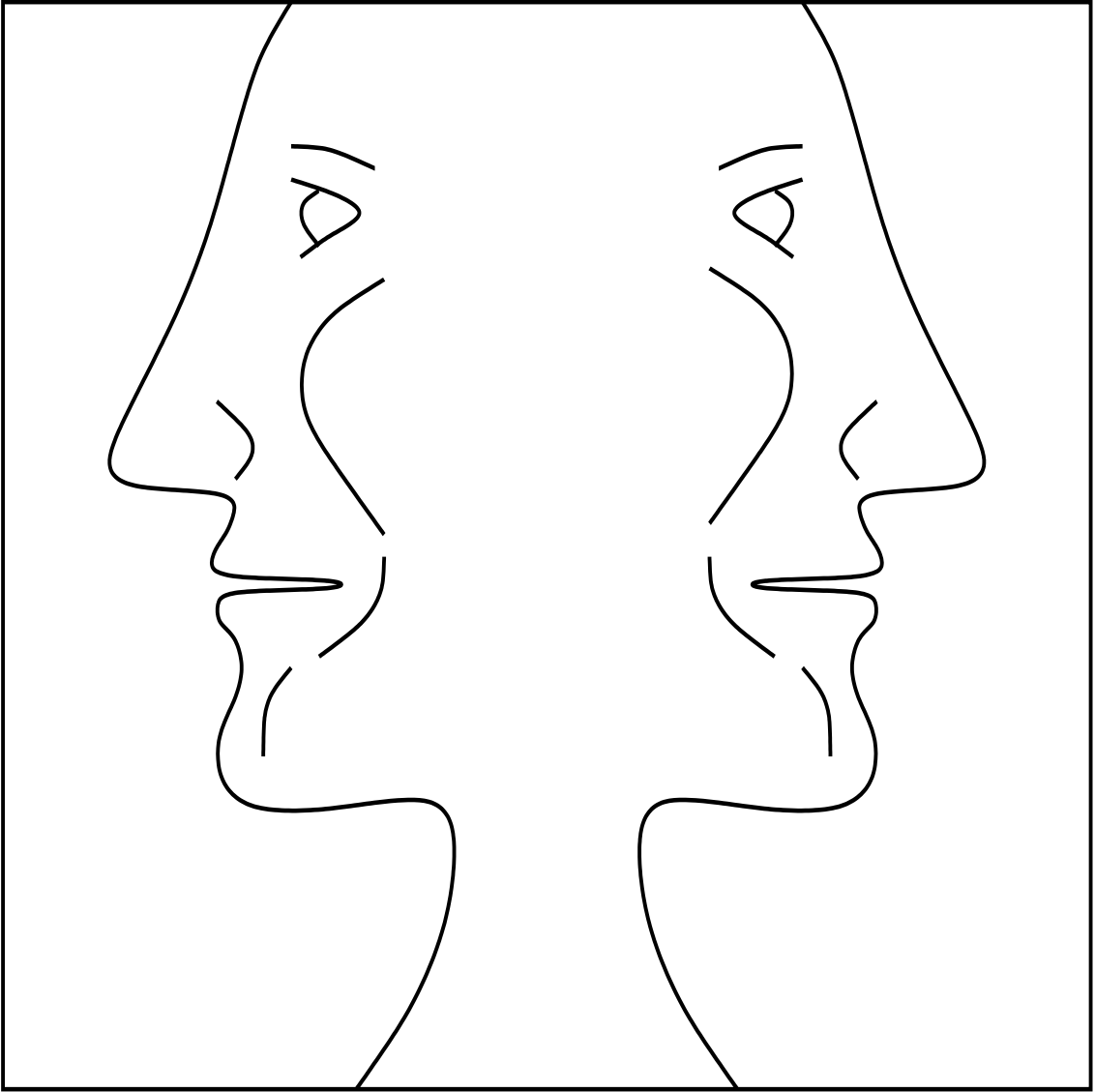
†Homepage: <http://www.math.unipd.it/~ponno>

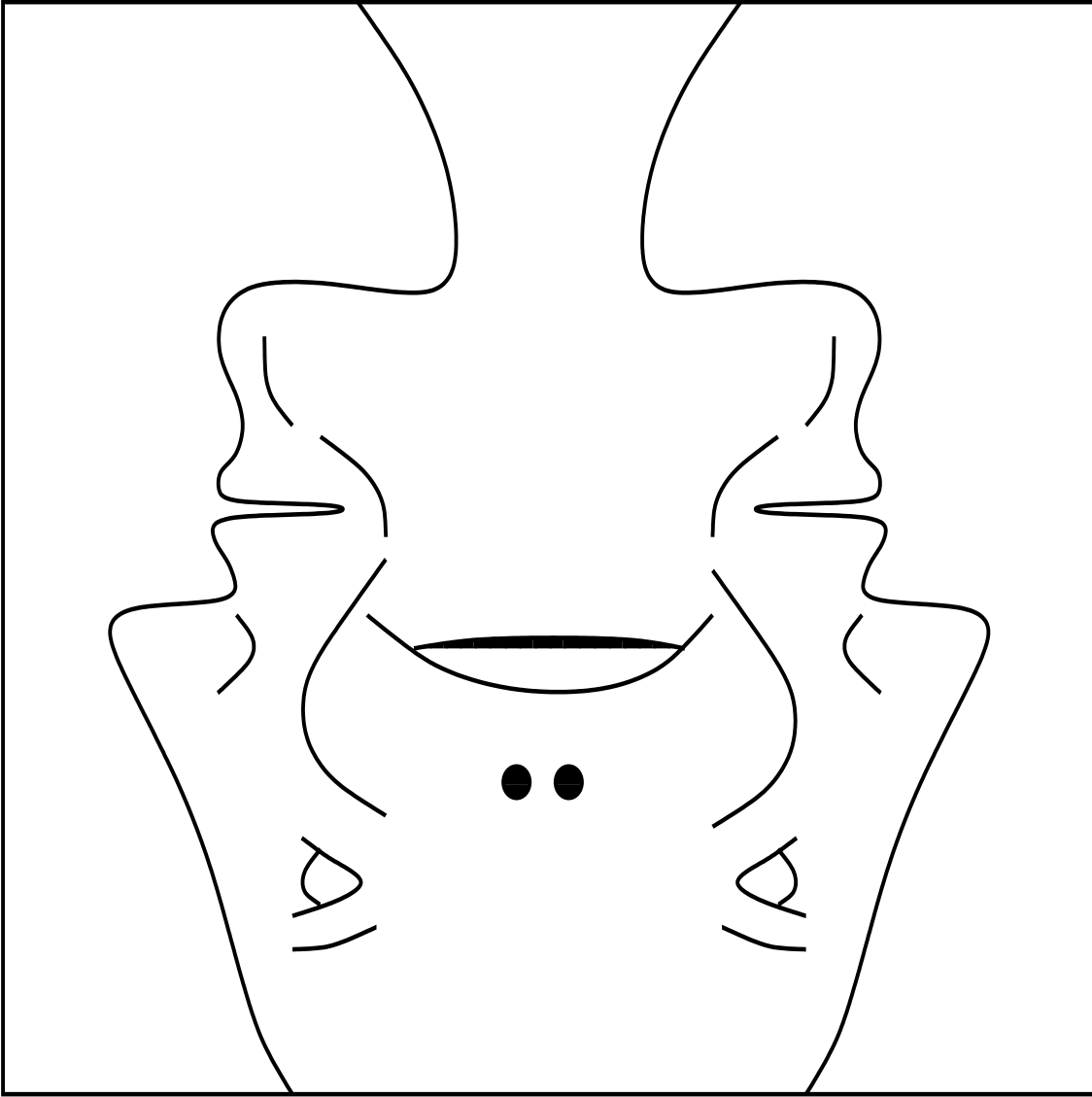
**MINI - ESPERIMENTO
SULLA DINAMICA
DEL RICONOSCIMENTO**

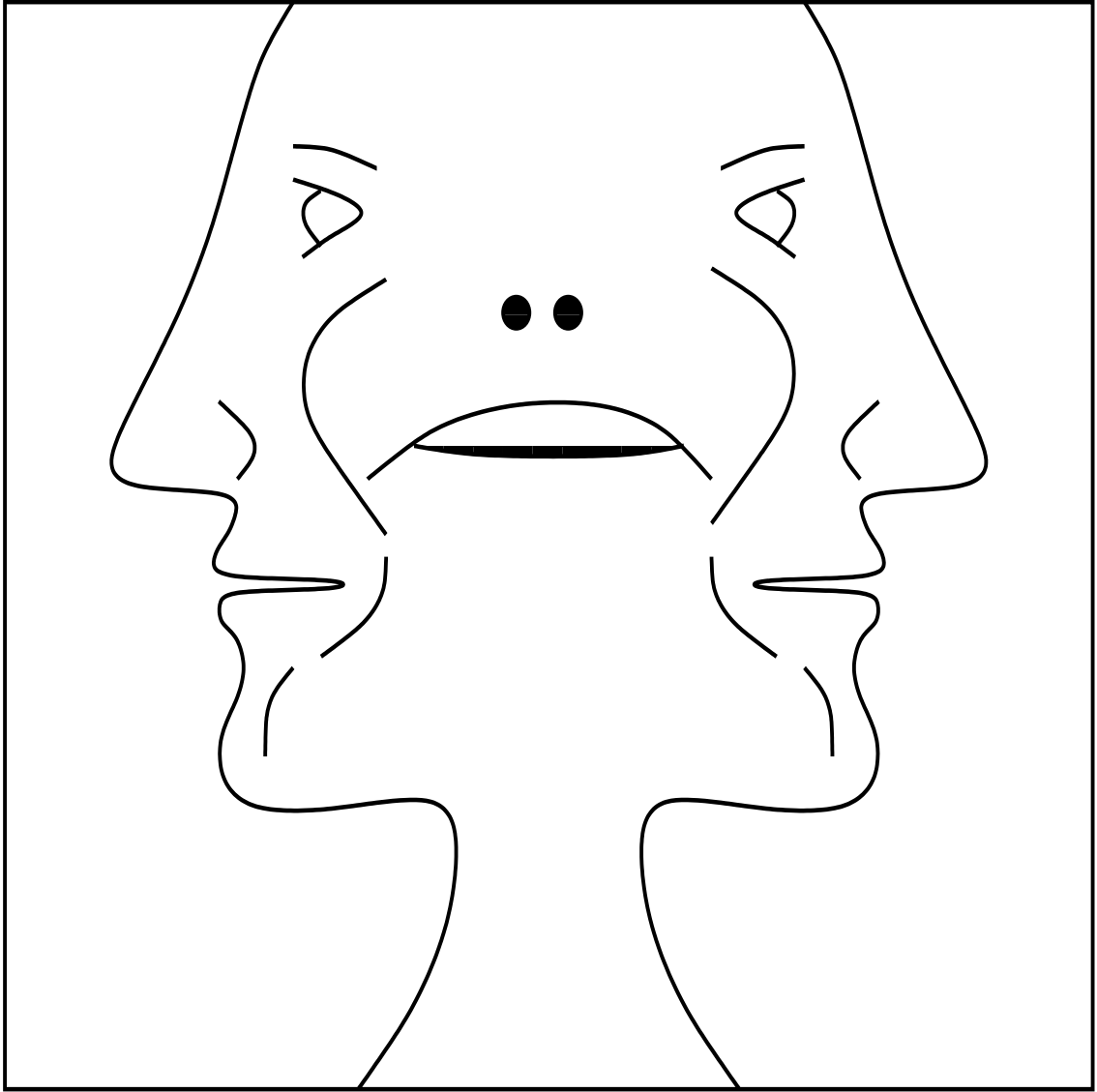












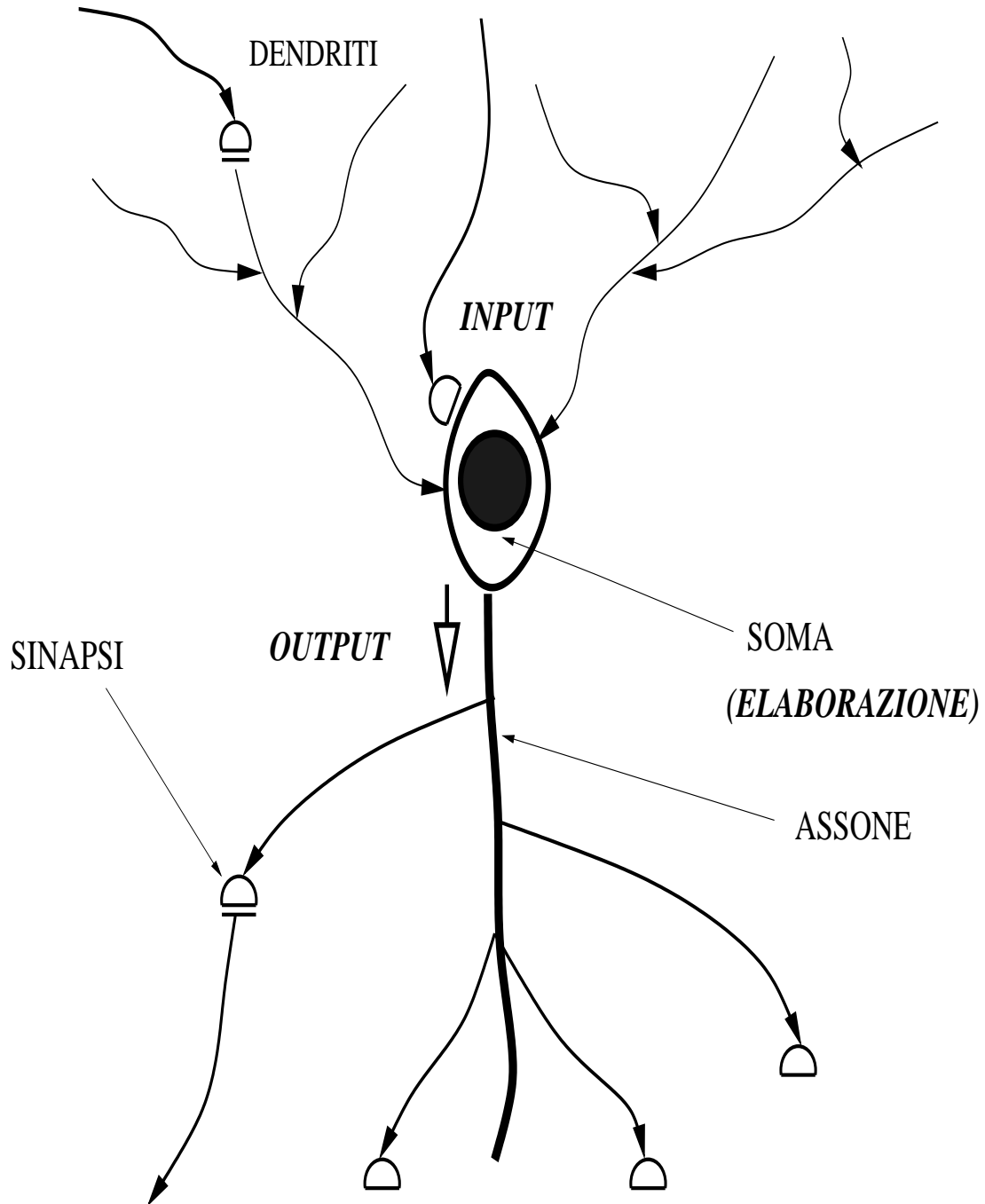
- **Neuroni:** struttura e funzionamento
- **Un modello di rete** neurale
- **Memoria associativa**

il problema degli attrattori

Il modello di Hopfield

- Dinamica del **riconoscimento**

Il Neurone



La d.d.p. di membrana di un neurone a riposo è

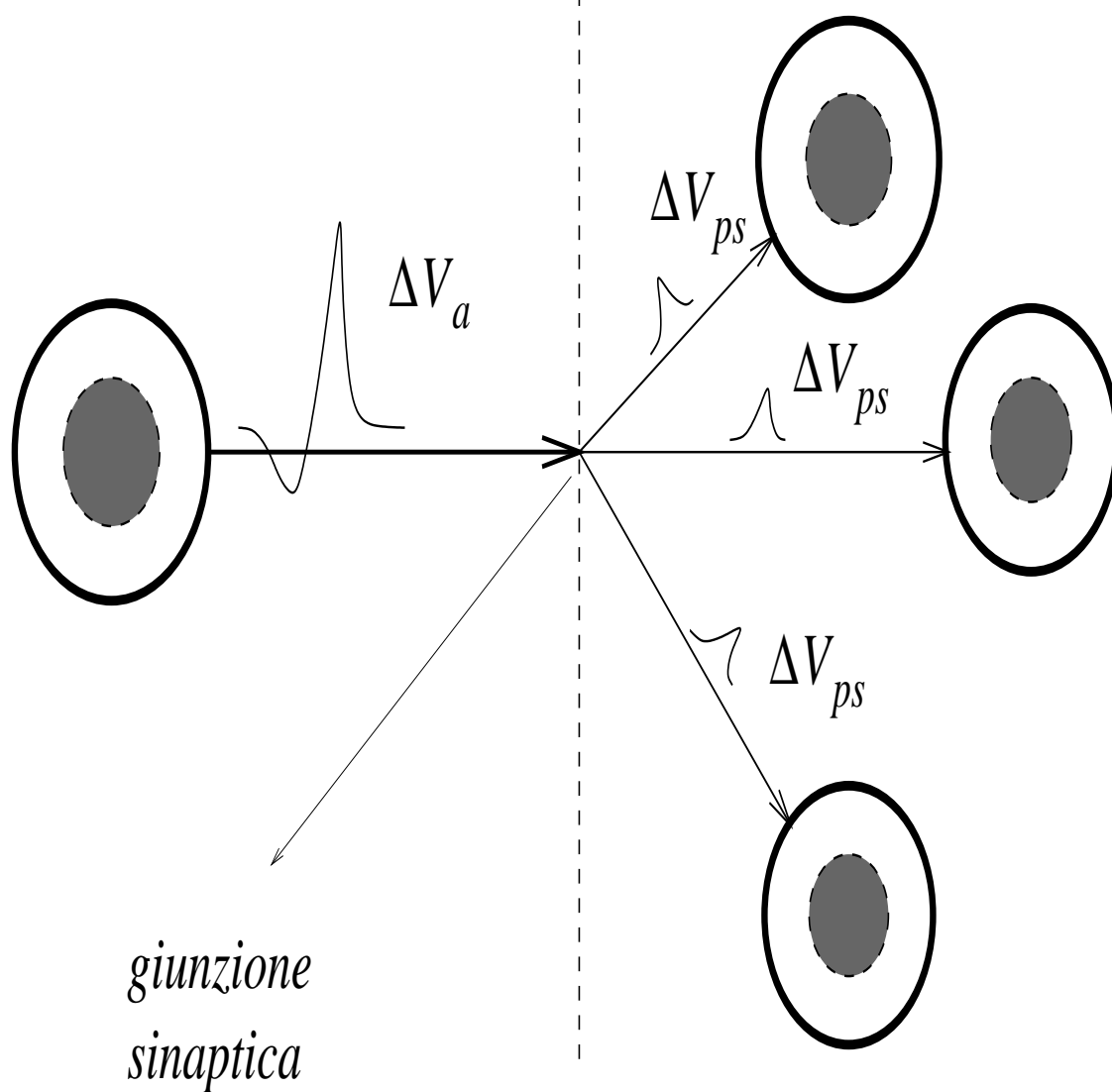
$$\Delta V_r = V_r^{int} - V_r^{est} \simeq -70mV ;$$

pompa sodio/potassio: eccesso di ioni Na^+ all'esterno e ioni K^+ all'interno. Quando è stimolato il neurone può depolarizzarsi e il potenziale di membrana raggiunge una intensità di picco $\Delta V_a \simeq +40mV$ (**potenziale d'azione**). Si genera così un'onda di depolarizzazione (impulso elettrico) che si propaga (1-100 m/s) lungo l'assone e si ramifica attraverso le sinapsi (giunzioni chimiche) raggiungendo i dendriti o il soma di altri neuroni.

Il neurone si depolarizza se una opportuna **somma pesata dei potenziali postsinaptici** ($\Delta V_{ps} \simeq 1mV$ da ogni neurone presinaptico) che lo raggiungono supera una soglia specifica (circa 10 mV), altrimenti non risponde allo stimolo.

**Regione
pre-sinaptica**

**Regione
post-sinaptica**



Corteccia cerebrale: circa 10^{11} neuroni suddivisi in circa 10^7 “microreti” fortemente connesse, ognuna contenente circa 10^4 neuroni ed avente dimensione lineare di 1mm. Ogni neurone della microrete riceve da molti dei neuroni componenti. La connettività tra microreti è minore.

Tra lo “sparo” di un neurone e la ricezione del segnale da parte dei neuroni ad esso connessi in uscita passa un intervallo di tempo τ di qualche millisecondo, durante il quale il neurone non può operare di nuovo: è il **periodo refrattario (assoluto)**. Dunque l'attività cerebrale ha frequenza limitata (meno di 1 KHz) e la dinamica neuronale è a tempo discreto (multipli di τ).

Topologia delle connessioni

- È conveniente introdurre la così detta **matrice di adiacenza del grafo di rete**:
 $\varepsilon_{ij} = 1$ se j afferisce ad i , altrimenti $\varepsilon_{ij} = 0$; in particolare $\varepsilon_{ii} = 0$.

- La matrice ε definisce completamente la topologia delle connessioni. Ad esempio essa individua le due componenti pre- e post-sinaptica:

$$\mathcal{C}_{pre}(i) = \{j \in \{1, \dots, N\} : \varepsilon_{ij} = 1\}$$

e

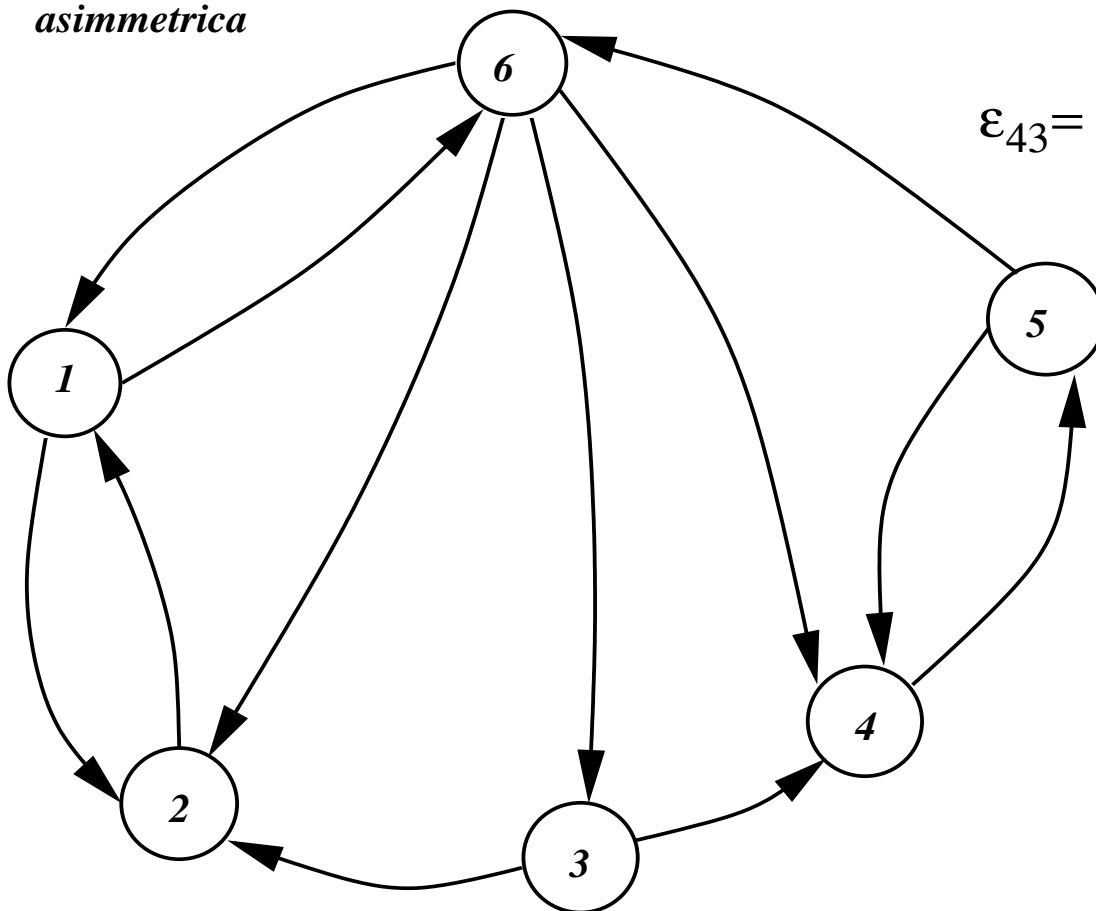
$$\mathcal{C}_{post}(i) = \{j \in \{1, \dots, N\} : \varepsilon_{ji} = 1\} .$$

- Il neurone i riceve segnali in ingresso dai neuroni in $\mathcal{C}_{pre}(i)$ e quando è attivo invia un segnale di uscita ai neuroni di $\mathcal{C}_{post}(i)$.

Esempio di rete

*rete a 6 neuroni
asimmetrica*

$2^6=64$ configurazioni
possibili



$$\begin{aligned}\epsilon_{12} &= \epsilon_{16} = 1 \\ \epsilon_{21} &= \epsilon_{23} = \epsilon_{26} = 1 \\ \epsilon_{36} &= 1 \\ \epsilon_{43} &= \epsilon_{45} = \epsilon_{46} = 1 \\ \epsilon_{54} &= 1 \\ \epsilon_{61} &= \epsilon_{65} = 1\end{aligned}$$

Modellizzazione

Introduciamo le seguenti definizioni

- $\Delta V_i(t)$: la d.d.p. (tra interno ed esterno) del neurone i -esimo al tempo t . ΔV_i può assumere solo due valori: ΔV_a (depolarizzazione) o ΔV_r (riposo).
- $\Delta v_i(t)$ d.d.p. post-sinaptica del neurone i al tempo t ; Δv_i può assumere due valori: 0 (i a riposo) oppure $\Delta V_{ps} \simeq 1mV$ (i attivo).
- $h_i(\Delta v_1, \dots, \Delta v_N)$: funzione di integrazione sinaptica del neurone i (o “campo locale” i -esimo), differenziabile vicino a $(0, \dots, 0)$ e t.c. $h_i(0, \dots, 0) = -S_i < 0$.

Per $i = 1, \dots, N$ si ha

$$\Delta V_i(t + \tau) - \Delta V_r = (\Delta V_a - \Delta V_r) \Theta(h_i(t))$$

dove $\Theta(x) = 1$ se $x > 0$ e $\Theta(x) = 0$ se $x \leq 0$. Se $h_i(t) \leq 0$ $\Delta V_i(t + \tau) = \Delta V_r$ e il neurone non risponde allo stimolo, mentre se $h_i(t) > 0$ allora $\Delta V_i(t + \tau) = \Delta V_a$ e il neurone “spara”.

Le d.d.p. post-sinaptiche Δv_i sono piccole,
quindi

$$h_i(\Delta v_1, \dots, \Delta v_N) \simeq -S_i + \sum_j K_{ij} \Delta v_j$$

a meno di un resto $O(\sum_j \Delta v_j^2)$ (Taylor).
Le K_{ij} sono le componenti della **matrice sinaptica**. Dalle ipotesi fatte sulle connessioni

$$K_{ij} \propto \varepsilon_{ij} \quad :$$

h_i dipende da Δv_j solo se $j \in \mathcal{C}_{pre}(i)$.

Introduciamo **variabili dicotomiche simmetriche**, definite per ogni $i = 1, \dots, N$ da

$$\Delta V_i(t) - \Delta V_r \equiv (\Delta V_a - \Delta V_r) \frac{\sigma_i(t) + 1}{2} ,$$

$$\Delta v_i(t) \equiv \Delta V_{ps} \frac{\sigma_i(t) + 1}{2} ,$$

$$\begin{aligned} \sigma_i(t) &= 2 \frac{\Delta V_i(t) - \Delta V_r}{\Delta V_a - \Delta V_r} - 1 = \\ &= 2 \frac{\Delta v_i(t)}{\Delta V_{ps}} - 1 . \end{aligned}$$

- $\sigma_i(t) = +1$ se i è attivo ($\Delta V_i = \Delta V_a$, $\Delta v_i = \Delta V_{ps}$);
- $\sigma_i(t) = -1$ se i è a riposo ($\Delta V_i = \Delta V_r$, $\Delta v_i = 0$).

In termini di variabili di spin $\sigma_i = \pm 1$ le equazioni del modello diventano

$$\sigma_i(t + \tau) = \text{sign}(h_i(t))$$

$$h_i(t) = \sum_j J_{ij} \sigma_j(t) - s_i$$

per $i = 1, \dots, N$, dove $\text{sign}(x) = +1$ se $x > 0$ e $\text{sign}(x) = -1$ se $x \leq 0$.

$$J_{ij} \equiv \frac{1}{2} K_{ij} \Delta V_{ps}$$

$$s_i \equiv S_i - \sum_j J_{ij}$$

ridefiniscono matrice sinaptica e soglia rispettivamente.

Una **semplificazione ragionevole (?)**:

$$s_i = 0 \quad \text{cioè} \quad S_i = \frac{1}{2} \sum_j K_{ij} \Delta V_{ps} :$$

la soglia è metà di quello che il neurone riceve se i suoi afferenti (pre-sinaptici) sono tutti attivi.

Il modello dinamico di rete è descritto da

$$\sigma_i(t + \tau) = \text{sign} \left(\sum_j J_{ij} \sigma_j(t) \right)$$

per $i = 1, \dots, N$. Se è nota la matrice sinaptica J_{ij} ($N \times N$) di ogni neurone i , resta definita una mappa

$$\mathbf{S} : \{-1, +1\}^N \longrightarrow \{-1, +1\}^N : \boldsymbol{\sigma} \mapsto \mathbf{S}[\boldsymbol{\sigma}]$$

Lo stato della rete al tempo $n\tau$ è dato dalla iterata n -esima di \mathbf{S} calcolata nello stato iniziale $\boldsymbol{\sigma}(0)$:

$$\boldsymbol{\sigma}(n\tau) = \mathbf{S}^n[\boldsymbol{\sigma}(0)] \equiv \overbrace{(\mathbf{S} \circ \dots \circ \mathbf{S})}^{n \text{ volte}}[\boldsymbol{\sigma}(0)]$$

Lo spazio degli stati $\{-1, +1\}^N$ può essere visualizzato come l'insieme dei vertici di un cubo N -dimensionale di spigolo lungo 2. Ad ogni iterata, ovvero ad ogni passo temporale (τ) si salta da un vertice ad un altro.

Per l'esempio di rete a 6 neuroni (v. figura):

$$\sigma_1(t + \tau) = \text{sign}[J_{12}\sigma_2(t) + J_{16}\sigma_6(t)]$$

$$\sigma_2(t + \tau) = \text{sign}[J_{21}\sigma_1(t) + J_{23}\sigma_3(t) + J_{26}\sigma_6(t)]$$

$$\sigma_3(t + \tau) = \text{sign}[J_{36}\sigma_6(t)]$$

$$\sigma_4(t + \tau) = \text{sign}[J_{43}\sigma_3(t) + J_{45}\sigma_5(t) + J_{46}\sigma_6(t)]$$

$$\sigma_5(t + \tau) = \text{sign}[J_{54}\sigma_4(t)]$$

$$\sigma_6(t + \tau) = \text{sign}[J_{61}\sigma_1(t) + J_{65}\sigma_5(t)]$$

Memoria ed attrattori

Problema Il modello deve includere una memoria contenente oggetti (parole, immagini, suoni...), ogni oggetto corrispondente ad un determinato stato di attività della rete:

un oggetto = stringa di N bit = ξ_1, \dots, ξ_N
= una configurazione della rete;

$\xi_i = \pm 1$: neurone i -esimo attivo o a riposo

Evidentemente matrice sinaptica e memoria devono essere legate. **Regola di Hebb:**

$$J_{ij} = \frac{\varepsilon_{ij}}{N} \sum_{\mu=1}^p \xi_i^{(\mu)} \xi_j^{(\mu)} = \frac{\varepsilon_{ij}}{N} \sum_{\mu=1}^p \left[\xi^{(\mu)} \left(\xi^{(\mu)} \right)^T \right]_{ij}$$

Dunque p stringhe $\xi^{(\mu)} = \xi_1^{(\mu)}, \dots, \xi_N^{(\mu)}$ ($\mu = 1, \dots, p$) di N bit ciascuna (oggetti in memoria) e una assegnata matrice di adiacenza (struttura delle connessioni) determinano la matrice sinaptica: **questo è il modello di Hopfield.**

Le stringhe $\xi^{(\mu)}$ che determinano la matrice sinaptica sono attrattori per il sistema dinamico se

$$\sigma(n\tau) \xrightarrow{n \rightarrow \infty} \xi^{(\nu)}$$

per qualche $\nu = 1, \dots, p$ che dipende dallo stato iniziale $\sigma(0)$ della rete. In questo caso parliamo di memoria associativa:

ad una classe di stati iniziali di attività viene associato (dalla dinamica) uno stesso stato di attività asintotico della rete, stato che quindi rappresenta la classe di dati di partenza.

Il modello matematicamente piú semplice da trattare è quello in cui $\xi_i^{(\mu)}$ e $\xi_j^{(\nu)}$ sono variabili aleatorie indipendenti (per $i \neq j$ e/o $\mu \neq \nu$) e suscettibili di assumere i valori ± 1 con la stessa probabilità a priori. In questo caso il riconoscimento ha ovvio carattere probabilistico.

Perchè la regola di Hebb?

Vogliamo che un dato pattern ξ risulti punto fisso di \mathbf{S} : $\mathbf{S}[\xi] = \xi$, ovvero

$$\xi_i = \text{sign} \left(\sum_j J_{ij} \xi_j \right) . \quad (\star)$$

Se $J_{ij} = c \varepsilon_{ij} \xi_i \xi_j$ (un solo pattern memorizzato), con $c > 0$, allora

$$\begin{aligned} \sum_j J_{ij} \xi_j &= \sum_j [c \varepsilon_{ij} \xi_i \xi_j] \xi_j = \left(\sum_j \varepsilon_{ij} \right) c \xi_i \\ &= \text{card} [\mathcal{C}_{pre}(i)] c \xi_i , \end{aligned}$$

di segno ξ_i , cioè la (\star) , se esiste almeno un neurone presinaptico.

Quando si memorizzano più pattern e la matrice J è somma di proiettori, ci si aspetta l'**esistenza di punti fissi in senso probabilistico**.

Esistenza di punti fissi

Proposizione 1

Per il modello di Hopfield con p patterns di memoria aleatori $\xi^{(\mu)}$, $\mu = 1, \dots, p$, e matrice di adiacenza assegnata, vale, per ogni $i = 1, \dots, N$ e ogni $\mu = 1, \dots, p$, la stima

$$P \left[\xi_i^{(\mu)} = \text{sign} \left(\sum_j J_{ij} \xi_j^{(\mu)} \right) \right] \geq 1 - \frac{p-1}{A_i} ,$$

dove $A_i = \text{card}[C_{pre}(i)]$.

Dimostrazione

$$\sum_j J_{ij} \xi_j^{(\mu)} = \frac{A_i}{N} \xi_i^{(\mu)} + R_i ,$$

in cui il “resto” R_i ha valore di aspettazione nullo:

$$E(R_i) = 0 .$$

Osserviamo inoltre che se $|R_i| < A_i/N$, allora $\text{sign} \left(\sum_j J_{ij} \xi_j^{(\mu)} \right) = \xi_i^{(\mu)}$.

Quindi:

$$\begin{aligned}
& \mathbb{P} \left[\xi_i^{(\mu)} = \text{sign} \left(\sum_j J_{ij} \xi_j^{(\mu)} \right) \right] = \\
& = \mathbb{P} \left[\xi_i^{(\mu)} = \text{sign} \left(\frac{A_i}{N} \xi_i^{(\mu)} + R_i \right) \right] \geq \\
& \geq \mathbb{P} \left[|R_i| < \frac{A_i}{N} \right] = \\
& = 1 - \mathbb{P} \left[|R_i| \geq \frac{A_i}{N} \right] \geq \\
& \geq 1 - \frac{\mathbb{E}(R_i^2)}{(A_i/N)^2} = 1 - \frac{p-1}{A_i} .
\end{aligned}$$

Osservazione: se $A_i \propto N$ e $p/N \rightarrow 0$ quando $N \rightarrow +\infty$, si dimostra che

$$\mathbb{P} \left[\boldsymbol{\xi}^{(\mu)} = \mathbf{S} \left[\boldsymbol{\xi}^{(\mu)} \right] \right] \rightarrow 1$$

per $N \rightarrow +\infty$, per ogni pattern μ .

Dinamica I

Proposizione 2

Se $J_{ij} = J_{ji}$ (cioè $\varepsilon_{ij} = \varepsilon_{ji}$) allora $\sigma(n\tau) = \mathbf{S}^n[\sigma(0)]$ tende, per $n \rightarrow \infty$, ad un punto fisso di \mathbf{S} oppure ad un ciclo binario (orbita periodica di periodo 2τ).

Dimostrazione Definiamo, per $t \geq \tau$,

$$E(t) \equiv - \sum_{i,j=1}^N J_{ij} \sigma_i(t) \sigma_j(t - \tau)$$

$E(t) \geq -N(N-1) \sup_{i,j} \{|J_{ij}|\}$. Inoltre

$$\begin{aligned} E(t + \tau) - E(t) &= \\ &= - \sum_i [\sigma_i(t + \tau) - \sigma_i(t - \tau)] \sum_{j(\neq i)} J_{ij} \sigma_j(t) = \\ &= - \sum_i [\sigma_i(t + \tau) - \sigma_i(t - \tau)] h_i(t) \leq 0 \end{aligned}$$

e il segno di uguaglianza vale se e solo se $\sigma_i(t + \tau) = \sigma_i(t - \tau)$ per ogni i .

Dinamica II

(Riconoscimento ad un passo)

Proposizione 3

Sia $\sigma_i(0) = \eta_i \xi_i^{(\mu)}$ ($i = 1, \dots, N$), dove $\eta_i = -1$ con probabilità $q < 0.5$ ed $\eta_i = 1$ con probabilità $1 - q$ (meno del 50% dei bit di partenza sbagliati rispetto alla stringa μ). Se $A_i \propto N$ per ogni i e $p/N \rightarrow 0$ per $N \rightarrow +\infty$, allora

$$P \left[\mathbf{S} [\boldsymbol{\sigma}(0)] = \boldsymbol{\xi}^{(\mu)} \right] \rightarrow 1$$

per $N \rightarrow +\infty$.

- D. J. Amit, *Modellizzare le funzioni del cervello*, Cedam, Padova, 1995.
- B. Tirozzi, *Modelli matematici di reti neurali*, Cedam, Padova, 1995.
- J. J. Hopfield, Brain, neural networks, and computation, *Review of Modern Physics* **71** S431 (1999).