

"segno" definita, componente, per componente, in $R - \{0\}$ come:

$$\text{sgn}(x) = \begin{cases} +1 & \text{per } x > 0 \\ -1 & \text{per } x < 0 \end{cases} \quad (2)$$

2) Il problema "individuale" della sintesi di una rete neurale può essere espresso nella seguente maniera:

- data una sequenza $\vec{\sigma}_1, \dots, \vec{\sigma}_m$ di stati, determinare una matrice A tale che

$$\vec{\sigma}_{k+1} = \text{sgn}(A \vec{\sigma}_k) \quad (3)$$

L'attributo "individuale", che contraddistingue il problema enunciato, è dovuto al fatto che esso si riferisce alla sintesi di reti per le quali una "singola" sequenza di stati è preassegnata. Se $\vec{\sigma}_{k+1} = \vec{\sigma}_1$, la sequenza sarà chiamata "riverbero"; se $\vec{\sigma}_k \neq \vec{0}_m$, per ogni k ed m tali che $1 \leq k < m \leq R+1$ noi diremo che la sequenza appartiene ad un transiente. In entrambi i casi, il problema individuale della sintesi di una rete neurale si riduce alla ricerca di una matrice A , i cui N^2 elementi siano soluzioni delle $N \times R$ disequazioni lineari, implicitamente contenute nelle (3). Le attuali tecniche di risoluzione sono facilitate applicarsi quando $R \leq N$, ma per i casi più interessanti, dal punto di vista delle applicazioni, cioè quando $R \gg N$, esse richiedono capacità di memoria e tempi di calcolo che crescono molto rapidamente con R . Da ciò deriva l'importanza di trovare un qualche

semplice criterio che permetta di stabilire a priori l'esistenza o meno di una soluzione, ogni qualvolta un problema concreto di sintesi deve essere affrontato. Guidati da queste considerazioni, ci è apparsa evidente l'utilità di formulare l'equivalente, per il modello in esame, di un ben noto risultato della teoria delle funzioni a soglia (Elgot 1960).

Denotando con $\tilde{\sigma}_i$ la i -esima riga della matrice A , le equazioni vettoriali (3) sono equivalenti alle seguenti $N \times R$ equazioni scalari

$$\sigma_{i,t+1} = \text{sgn} \cdot \tilde{\sigma}_i \cdot \tilde{\sigma}_t^+ \quad 1 \leq t \leq R, \quad 1 \leq i \leq N \quad (4)$$

dove $\tilde{\sigma}_i \cdot \tilde{\sigma}_t^+ = \sum_{j=1}^N a_{ij} \sigma_{j,t}$

Definiti gli insiemi

$$I_i^+ = \left\{ \tilde{\sigma}_t^+ \mid 1 \leq t \leq R; \sigma_{i,t+1} = +1 \right\}$$

$$I_i^- = \left\{ \tilde{\sigma}_t^- \mid 1 \leq t \leq R; \sigma_{i,t+1} = -1 \right\} \quad (5)$$

$$-I_i^- = \left\{ \tilde{\sigma}_t^+ \mid -\sigma_t^- \in I_i^- \right\}$$

$$I_i = I_i^+ \cup (-I_i^-)$$

e tenuto conto della (2), le (4) si trasformano nel seguente sistema di disequazioni lineari

$$\tilde{\sigma}_i \cdot \tilde{\sigma}_t^+ > 0 \quad \forall \sigma_t^+ \in I_i^+, \quad 1 \leq i \leq N \quad (6)$$

Un ben noto teorema di analisi convessa (Rockafeller, 1969) assicura che le (6) sono consistenti se e solo se:

*) il vettore nullo non appartiene all'involuppo convesso dell'insieme I , per $1 \leq i \leq N$, ovvero, se e solo se:

xy) per nessun i esistono R numeri $(g_i^v)_{v \in I_i}$ > 0 tali che

$$\begin{cases} \sum_{v \in I_i} g_i^v \vec{\sigma}^v = 0 \\ \sum_{v \in I_i} g_i^v = 1 \end{cases} \quad (7)$$

L'esame delle condizioni xy), alla luce dei teoremi classici dell'algebra lineare, porta alla determinazione di diversi semplici criteri, alcuni solo necessari, altri solo sufficienti, di realizzabilità della sintesi. Attualmente siamo impegnati nella messa a punto di un criterio necessario e sufficiente di realizzabilità che deriva dalle xy) ma, a differenza di queste fornirebbe un algoritmo esauribile in un numero finito e relativamente breve di passi. Supponiamo adesso che la sequenza $(\sigma_t^v)_{1 \leq t \leq R+1}$ sia realizzabile, cioè, esista una rete che contenga questa sequenza nella sua evoluzione. Se $\vec{\sigma}^v$ è un arbitrario stato della rete, dalle xy) segue che la sequenza $(\vec{\sigma}_t^v)_{1 \leq t \leq R+1}$ con $\vec{\sigma}_{R+2}^v = \vec{\sigma}^v$, sarà realizzabile se e solo se non esistono, per alcun i , R numeri non negativi

$$(g_i^v)_{v \in I_i} \quad \text{tali che}$$

$$\sum_{v \in I_i} g_i^v \vec{\sigma}^v = \vec{\sigma}_{R+1}^{v \neq i} \quad (8)$$

dove

$$\vec{\sigma}_{k+1}^* = -\sigma_{1,k+2}^* \vec{\sigma}_{k+1}^*$$

Il significato geometrico della condizione (8) è che la sequenza prolungata $(\sigma_i)_{1 \leq i \leq k+2}$ può essere realizzata se, essendo realizzabile la sequenza $(\sigma_i)_{1 \leq i \leq k+1}$ il vettore $\vec{\sigma}_{k+1}^*$ non appartiene al cono convesso generato dall'insieme I_1 .

Moltiplicando scalarmente il primo ed il secondo membro della (8) per $\vec{\sigma}_{k+1}^*$ si ottiene:

$$\sum_{\sigma \in I_1} \chi_{\sigma} \vec{\sigma} \cdot \vec{\sigma}_{k+1}^* = N \quad (9)$$

Quindi, una condizione sufficiente per la realizzabilità della sequenza prolungata è che, per ogni i , si abbia

$$\vec{\sigma} \cdot \vec{\sigma}_{k+1}^* \leq 0, \quad \forall \vec{\sigma} \in I_1 \quad (10)$$

2) il problema globale della sintesi. L'utilità pratica della soluzione di un problema individuale di sintesi dipende dallo stato iniziale che si assegna alla rete: cioè, una volta trovata una matrice A che realizza la sequenza preassegnata di stati $(\sigma_i)_{1 \leq i \leq k+1}$, nulla si sa, in generale, sulla evoluzione della rete quando si parte da uno stato iniziale $\vec{\sigma}_0$ diverso da ogni

$\vec{\sigma}_z$ della sequenza. Vi sono, d'altra parte, alcuni importanti problemi in cui si è interessati principalmente nella configurazione stabile finale della rete risultante da una arbitraria eccitazione iniziale: ci possiamo riferire, ad esempio, a Caiandello et al. (1967), per una impostazione teorica del problema, e a Burrattini e Liesis (1970), per un'analisi da un punto di vista sperimentale. Questo tipo di problema può chiaramente essere ridotto ad uno del tipo precedente, questa volta contenente 2^N disuguaglianze in N^2 incognite, ma le tecniche di risoluzione diventano presto di difficile applicazione, anche per piccoli valori di N . Volendolo, invece, trattare da un punto di vista intrinseco, svilupperemo questo problema in una formulazione geometrica equivalente.

Denotiamo con Q^N l'insieme dei 2^N possibili differenti stati di una rete, ovvero l'insieme dei vertici del cubo unitario simmetrico in $\mathbb{R}^N$, e definiamo l'applicazione

$$T_A : \vec{\sigma} \in Q^N \longrightarrow \operatorname{sgn}(A\vec{\sigma}) \in Q^N \quad (11)$$

L'evoluzione della rete sarà descritta da

$$T_A(\vec{\sigma}_t) = \vec{\sigma}_{t+1} \quad (12)$$

ed un generico stato $\vec{\sigma}$ apparterrà ad un riverbero di lunghezza R se e solo se R è il più piccolo intero positivo tale che:

1

752

2x-

21-

22a-

1

0-

into

ite

1-

14

a

nti

azio

$$T_A^R(\vec{\sigma}) = \vec{\sigma}$$

(13)

dove

$$T_A^0(\vec{\sigma}) = \vec{\sigma}$$

$$T_A^{n+1}(\vec{\sigma}) = T_A(T_A^n(\vec{\sigma}))$$

Se, viceversa, per ogni intero positivo R si ha $T_A^R(\vec{\sigma}) \neq \vec{\sigma}$, lo stato $\vec{\sigma}$ appartiene ad un transiente.

La funzione T_A gode delle seguenti due proprietà:

- (i) T_A è simmetrica, cioè, $T_A(-\vec{\sigma}) = -T_A(\vec{\sigma})$ per ogni $\vec{\sigma}$ in Q^N
- (ii) T_A proietta Q^N in se stesso, cioè, $T_A(Q^N) \subseteq Q^N$

Per ogni intero n si ha

$$T_A^{n+1}(Q^N) \subseteq T_A^n(Q^N)$$

(14)

e poichè Q^N è un insieme finito, l'inclusione può essere stretta solo per un numero finito di passi. Sia m il minimo intero per cui in (14) valga l'uguaglianza e poniamo

$$T_A^m(Q^N) = V$$

(15)

Il sottinsieme V di Q^N così definito è il più grande dei sottinsiemi di Q^N sul quale la funzione T_A agisce come una permutazione;

22a

che:

esso sarà detto "nucleo" della rete la cui matrice è A e sarà caratterizzato dalla seguente proprietà: uno stato appartiene ad un riverbero se e solo se esso appartiene a $\bar{V}$. In analogia con Q^n i punti di V possono essere considerati come i vertici di un poliedro simmetrico.

Le considerazioni precedenti possono essere riassunte come segue: ad ogni rete autoduale A corrisponde un sottosistema simmetrico $\bar{V}$ di Q^n tale che T_A agisce come una permutazione sugli elementi di V ; V è il più grande sottosistema di Q^n con questa proprietà.

Il problema globale della sintesi può, dunque, essere impostato come segue: dato un sottosistema simmetrico V di Q^n ed una permutazione P sui punti di V , esiste una rete autoduale A , tale che V sia il nucleo di A e la permutazione indotta da T_A su V coincida con P ?

Il nucleo di una rete è l'insieme delle sue configurazioni finali stabili in quanto, partendo da uno stato iniziale arbitrario la rete decade in uno stato appartenente al nucleo che è "stabile" rispetto alla evoluzione della rete, nel senso che, se ad un certo istante lo stato della rete appartiene al nucleo, vi apparterranno anche tutti gli stati successivi.

Il problema della realizzabilità vettoriale delle reti può essere generalizzato come segue: dato un insieme X ed un'azione su X di un sottogruppo G del gruppo di permutazioni sugli elementi di X , trovare il minimo intero positivo N tale che esista

una corrispondenza biunivoca φ di X su un poliedro Z di C^N (dove C è un qualsiasi campo) ed un omonorfismo ρ di G sul gruppo delle trasformazioni (non necessariamente lineari) di Z in se stesso, tali che se $g \in G$ si ha

$$\varphi \circ g = \rho(g) \circ \varphi \quad (16)$$

Questo approccio non è affatto ristretto al caso di reti di elementi a soglia autocoduali, ma si estende al problema generale della realizzazione vettoriale di un automa. Esso ci permette, inoltre di sviluppare una trattazione matematica del problema della realizzabilità delle reti (cioè, sistemi di funzioni booleane) completamente diversa dall'usuale. Nel caso in cui si è interessati ai comportamenti "riberberanti" della rete, bisogna ricordare che il concetto centrale nella questione è quello dell'azione del gruppo di permutazione sui vertici di un poliedro. Questa azione è usualmente espressa per mezzo della funzione di Heaviside composta con una trasformazione affine. Denotando con φ la trasformazione che porta il cubo booleano nel cubo unitario simmetrico, con π l'azione della funzione di Heaviside, con S l'azione della funzione segno si ottiene l'eguaglianza

$$\varphi \pi = S \varphi \quad (17)$$

La trasformazione φ ha due aspetti peculiari:

- 1) conserva la dimensione
- 2) porta un cubo in un cubo.

Il significato di queste proprietà è semplice: la trasformazione da un modello all'altro conserva tanto il numero di neuroni quanto la struttura geometrica del nucleo. E' evidente che nella (17) si perde ogni traccia di queste proprietà, cosicchè questa equazione può essere assunta (essendo φ biunivoca) come la definizione di

equivalenza tra due modelli vettoriali della stessa rete neuronica.

Più precisamente, se V è un arbitrario poliedro di uno spazio K^n ed S è un'azione di un sottogruppo di un gruppo di permutazione, noi diremo che la coppia (V, S) è un modello vettoriale equivalente alla coppia (Q, H) se esiste una trasformazione tale che $\varphi H = S\varphi$.

REFERENCES

- L. Accardi
Rank and Reverberations in Neural Networks
(1971)
- E. R. Caianiello
Outline of a theory of thought processes and
thinking machines
Journ. of Theoret. Biol. 2, 204-235 (1961)
- E. R. Caianiello
Decision equations and reverberations
Kybernetik, 3 Band, 2 Heft (1966) a)
A study of neural networks and reverberations
Bionics Symposium, Dayton, Ohio (1966) b)
- E. R. Caianiello
Reverberations and control of neural networks
Kybernetik, 4 Band, 2 Heft (1967)
- T. Arimoto
Periodical sequences realizable by an auton-
omous net of threshold elements, Papers of the
Committee on Automation and Automatic Control,
Inst. Elec. Commun. of Japan, March 1963
- Rockefeller
Convex Analysis
Princeton 1969
- Gantmacher
Chelsea 1960
- Rigot
Switching Circuit Theory and Logic Design
pp. 225-245, September 1961
- E. Burattini
No appear
V. Lises