Complexity of Neural Networks

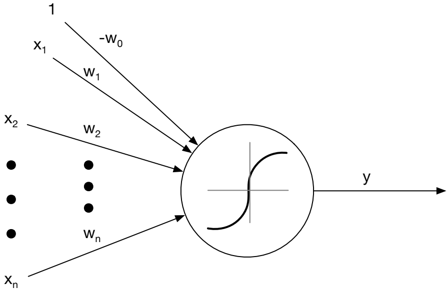
Vasquez Alessandro

Université Nice-Sophia-Antipolis

28 février 2018

Neuron's representation :
$y = \sigma(\sum_{j=1}^{n} w_j x_j - w_0)$
Transfer functions :

- $sign(t) = \begin{cases} 0 & \text{if } t \leq 0 \\ 1 & \text{if } t > 0 \end{cases}$
- $\sigma(t) = (1 + e^{-t})^{-1}$



Preliminaries

Neural nets can be

- cyclic or acyclic
- symmetric or asymmetric w.r.t. interconnections
- continuous or discrete w.r.t. updates
- synchronous or asynchronous w.r.t. updates in discrete neural nets

Type of neurons

- perceptron (with transfer function)
- threshold gate (with binary valued-input)
- majority gate (all $w = 1$, $b = \frac{|X|}{2}$)

Definition (Threshold function)

A boolean function $t : \{0,1\} \to \{0,1\}$ is a *threshold function* if it is computable by a linear threshold unit.

Preliminaries

Parameters for evaluating neural nets' complexity

- size
- depth
- weight

We're going to consider discrete-weighted

- threshold gates-circuits
- Hopfield nets

Theorem

Any threshold function on n variables can be computed by a threshold gate with integer weights w_i st :

$$|w_i| \leq \frac{(n+1)^{(n+1)/2}}{2^n}, \forall i = 0, ..., n.$$

Theorem

For infinitely many n, there are threshold functions on n variables whose computation by a single threshold gate requires weights as large as $\frac{n^{n/2}}{2^n}$.

These theorems imply that there is a class of threshold gates of polynomially bounded weights. But even in the case of non-polynomially bounded weights, these can be represented in $O(s \log s)$ bits.

Let's consider acyclic networks which compute a boolean function $f : \{0,1\}^n \to \{0,1\}$.

Definition (Complexity classes)

We define five class of functions computable by different classes of circuits.

- NC^k : polynomial-size circuits of depth $O(\log^k(n))$, using bounded fan-in AND, OR, and NOT gates,
- AC^k : polynomial-size circuits of depth $O(\log^k(n))$, using unbounded AND, OR, and NOT gates,
- TC^k : threshold circuits of polynomial size and depth $O(log^k n)$,
- TC^0_d threshold circuits of polynomial size and depth d,
- $\widehat{TC}^0_d$ majority circuits of polynomial size and depth d.

Theorem (Complexity classes' hierarchy)

$$AC^k \subseteq TC^k \subseteq NC^{k+1}.$$

$$\widehat{TC}^0_d \subseteq TC^0_d \subseteq \widehat{TC}^0_{d+1}.$$

We only know that $AC^0 \subset TC^0$, e.g. because of the majority function, and $TC^0_1 \subset \widehat{TC^0_2}$ for the parity function.

The following problems are proved to be NP-complete.

Definition (Threshold circuit loading)

Given $\{(\vec{x_1}, b_1), ..., (\vec{x_m}, b_m)\}$, where each $\vec{x_i} \in \{0,1\}^n$ and each $b_i \in \{0,1\}$, and a directed acyclic graph, can we find a weight assignment st given f the function computed by the resulting threshold circuit : $f(\vec{x_i}) = b_i, \forall i = 1, ..., m$?

Definition (Threshold circuit minimization)

Given $\{(\vec{x_1}, b_1), ..., (\vec{x_m}, b_m)\}$, where each $\vec{x_i} \in \{0,1\}^n$ and each $b_i \in \{0,1\}$, and $K \in \mathbb{N}$, does exists a threshold circuit of at most K neurones st given f the function computed by the resulting threshold circuit : $f(\vec{x_i}) = b_i, \forall i = 1, ..., m$?

Definition (Hopfield network)

Neural net in which

- the weights w_{ij} are symmetric
- the net has a vector I which represents its global state
- each neuron i has a local state I_i, where I is the states' vector
- each neuron's state is updated asynchronously
- there's an "energy function" assigned to the network

The energy function is

$$E = -\frac{1}{2}\sum_{i,j} w_{ij}s_i s_j - \sum_i I_i s_i$$

If we derivate the energy function w.r.t. each s_i and changing its sign we get

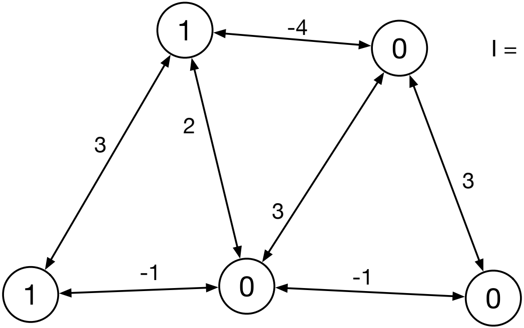
$$\frac{\partial E}{\partial s_i} = I_i + \sum_j s_j w_{ij},$$

for each neuron i.

If we consider the state of a neuron as its bias, we get

$$\frac{\partial E}{\partial s_i} = b_i + \sum_j s_j w_{ij},$$

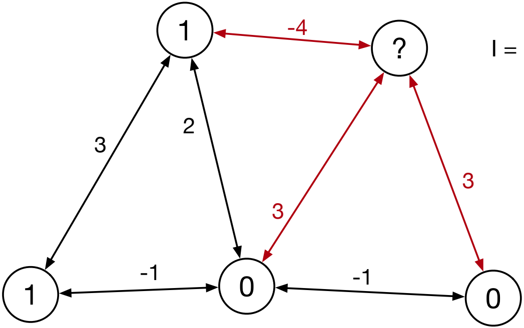
which is the binary threshold decision rule computed locally by each neuron. In 1982 Hopfield proved that E is a Lyapunov function.



I = [1, 1, 0, 0, 0]

E = -3

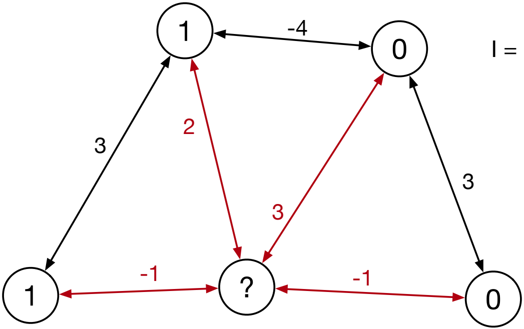
Hopfield Network - Example



I = [1, 1, ?, 0, 0]

E = -3

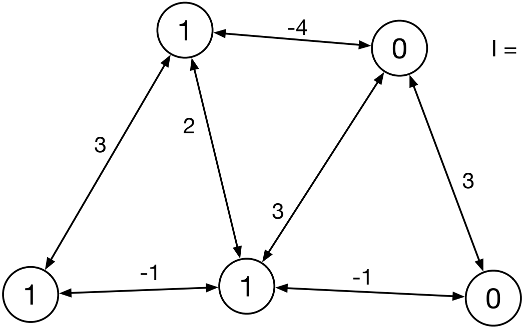
Hopfield Network - Example



I = [1, 1, 0, 0, 0]

E = -3

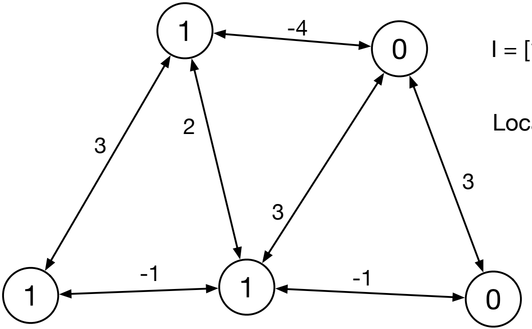
Hopfield Network - Example



I = [1, 1, 0, ?, 0]

E = -3

Hopfield Network - Example

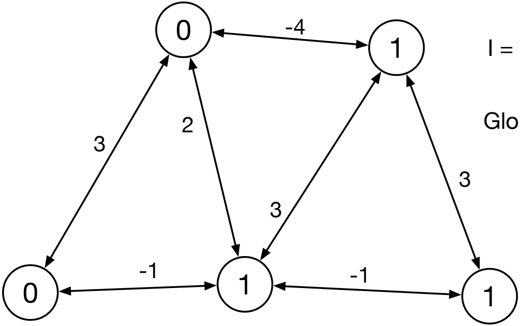


I = [1, 1, 0, ?, 0]

E = -4

Local minimum

Hopfield Network - Example



I = [1, 1, 0, 1, 0]

E = -5

Global minimum

Storage Rule

Definition (Storage Rule)

An algorithm which starts by a set of vectors $\vec{X} = \{\vec{x_1}, .., \vec{x_n}\}$ and builds an Hopfield net s.t. its stable state are those vector its called a *storage rule.*

Hopfield himself thought of its nets as an implementation of the *associative memory*.

Theorem

In a symmetric Hopfield net with w_{ij} integer weights and n neurons, the convergence is obtain in

$$O(n^2 \cdot \max_{i,j} |w_{ij}|)$$

asynchronous changes of neuron states.

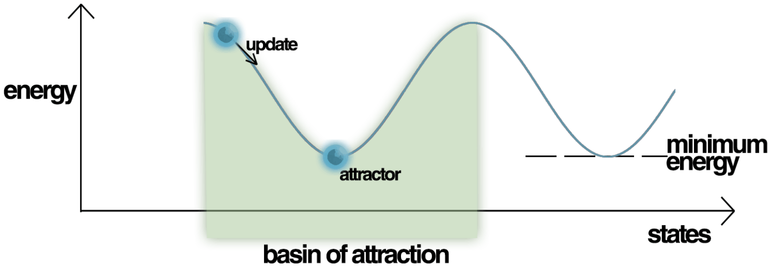
Definition (Attraction domain)

Let $\vec{x_i}$ be a stable global state for a given Hopfield net. The *attraction domain* of $\vec{x_i}$ is the set of all vectors which are guaranteed to converge to $\vec{x_i}$.

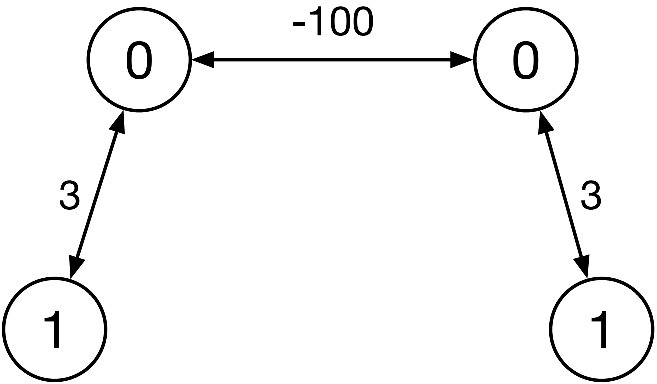
Definition (Attraction radius)

The *attraction radius* of x_i is the largest Hamming distance from which all other vectors are guaranteed to converge to $\vec{x_i}$. The attraction radius determines the basin of attraction.

Local vs global minimum



Oscillating synchronous update



- Decide if a given symmetric, simple network has more than one stable state.
- Determine if a symmetric simple network converges to a given stable state from any other initial state.

- Decide if a given asymmetric net under synchronous updates will converge from any (or all) initial state.
- Computing the attraction radius of a given stable state in a symmetric simple net.

Open problems

Definition (Hopfield Net Loading Problem)

Given a set of vectors $X = \{\vec{x_1}, .., \vec{x_m}\}$, where each $\vec{x_i} \in \{-1, 1\}^n$, and a constant ρ, is there a symmetric net of n neurons that has X as stable states with attraction radii $\geq \rho$?

Open problems

The following questions are yet to be answered.

- Is the Hopfield Net Loading Problem in P or in NP?
- Are the separations $TC_d^0 \subseteq TC_{d+1}^0, \forall d \geq 2$ proper?
- Does exist a d s.t. $AC^0 \subseteq TC_d^0$, i.e. does exists a class of $d-$depth majority circuits which can compute all functions computed by a depth 1, unbounded, AND-OR-NOT circuit?

Thank you for your attention.