CS623: Introduction to Computing with Neural Nets
(lecture-12)

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay

Training of Hopfield Net

- Early Training Rule proposed by Hopfield
- Rule inspired by the concept of electron spin
- Hebb's rule of learning
  - If two neurons *i* and *j* have activation x_i and x_j respectively, then the weight w_{ij} between the two neurons is directly proportional to the product $x_i \cdot x_j$ *i.e.*

$$w_{ij} \propto x_i \cdot x_j$$

Hopfield Rule

- To store a pattern

  $<x_n, x_{n-1}, \ldots., x_3, x_2, x_1>$

  make

$$w_{ij} = \frac{1}{(n-1)} \cdot x_i \cdot x_j$$

- Storing pattern is equivalent to '*Making that pattern the stable state*'

Training of Hopfield Net

- Establish that

  $$<x_n, x_{n-1}, \ldots., x_3, x_2, x_1>$$

  is a stable state of the net

- To show the stability of

  $$<x_n, x_{n-1}, \ldots., x_3, x_2, x_1>$$

  impress at $t=0$

  $$<x^t_n, x^t_{n-1}, \ldots., x^t_3, x^t_2, x^t_1>$$

Training of Hopfield Net

- Consider neuron *i* at *t=1*

$$a_i(t=1) = \mathrm{sgn}(net_i(t=0))$$

$$net_i(t=0) = \sum_{j \neq i,\, j=1}^{n} w_{ij} \cdot (x_j(t=0))$$

Establishing stability

$$\sum_{j \neq i,\, j=1}^{n} w_{ij} x_j(t=0)$$

$$= \frac{1}{(n-1)} \sum_j (x_i(t=0)) \cdot (x_j(t=0)) \cdot (x_j(t=0))$$

$$= \frac{1}{(n-1)} \sum_j (x_i(t=0)) \cdot [x_j(t=0)]^2$$

$$= \frac{1}{(n-1)} (x_i(t=0)) \cdot \sum_{j=1,\, j \neq i} 1$$

$$= \frac{1}{(n-1)} \cdot (x_i(t=0)) \cdot (n-1)$$

$$= x_i(t=0)$$

Thus,

$$x_i(t=1) = \mathrm{sgn}(x_i(t=0))$$

Observations

- How much deviation can the net tolerate?
- What if more than one pattern is to be stored?

Storing *k* patterns

- Let the patterns be:

  $P_1 : \langle x_n, x_{n-1}, \ldots., x_3, x_2, x_1 \rangle^1$

  $P_2 : \langle x_n, x_{n-1}, \ldots., x_3, x_2, x_1 \rangle^2$

  .

  .

  .

  $P_k : \langle x_n, x_{n-1}, \ldots., x_3, x_2, x_1 \rangle^k$

- Generalized Hopfield Rule is:

P^{th} pattern

$$wij = \frac{1}{(n-1)} \sum_{p=1}^{k} x_i \cdot x_j \,|_p$$

Storing *k* patterns

- Study the stability of

  $<x_n, x_{n-1}, \ldots., x_3, x_2, x_1>$

- Impress the vector at t=0 and observer network dynamics
- Looking at neuron *i* at *t=1*, we have

Examining stability of the qth pattern

$$
\begin{aligned}
& x_i(1)|_q \\
&= \mathrm{sgn}(net_i(1)|_q); net_i(1)|_q = \sum_{j=1, j\neq i}^{n} w_{ij} \cdot x_j(0)|_q \\
& w_{ij} \cdot x_j(0)|_q \\
&= \frac{1}{(n-1)}[\sum_{p=1}^{k} x_i(0)|_p \cdot x_j(0)|_p] \cdot x_j(0)|_q \\
&= \frac{1}{(n-1)}[\sum_{p=1, p\neq q}^{k} x_i(0)|_p \cdot x_j(0)|_p] \cdot x_j(0)|_q + \frac{1}{(n-1)} x_i(0)|_q \cdot x_j(0)|_q \cdot x_j(0)|_q \\
&= \frac{1}{(n-1)}[\sum_{p=1, p\neq q}^{k} x_i(0)|_q \cdot x_j(0)|_p] + \frac{1}{(n-1)} (x_i(0)|_q) \cdot [x_j(0)|_q]^2 \\
&= Q + \frac{1}{(n-1)} \cdot x_i(0)|_q \cdot 1 \\
&= Q + \frac{x_i(0)|_q}{(n-1)}
\end{aligned}
$$

Examining stability of the q^{th} pattern

$Thus,$

$$x_i(1) = \operatorname{sgn}[\sum_{j=1, j\neq i}^{n} Q + \sum_{j=1, j\neq i}^{n} \frac{x_i(0)}{(n-1)}]$$

$$= \operatorname{sgn}[\sum_{j=1, j\neq i}^{n} Q + x_i(0)]$$

$$= \operatorname{sgn}[\underbrace{\sum_{j=1, j\neq i}^{n} \sum_{p=1, p\neq q}^{k} (x_i(0)|_q \cdot x_j(0)|_p)}_{\text{Small when } k << n} + x_i(0)]$$

Small when $k << n$

Stability for *k memory elements*

- Condition for patterns to be stable on a Hopfield net with *n* neurons is:

$$k << n$$

- The storage capacity of Hopfield net is very small
- Hence it is not a practical memory element

Hopfield Net: Computational Complexity

- Hopfield net is an $O(n^2)$ algorithm, since
- It has to reach stability in $O(n^2)$ steps

Hopfield Net

- Consider the energy expression

$$E = -[w_{12}x_1x_2 + w_{13}x_1x_3 + \dots + w_{1n}x_1x_n$$

$$+ w_{23}x_2x_3 + w_{24}x_2x_4 + \dots + w_{2n}x_2x_n$$

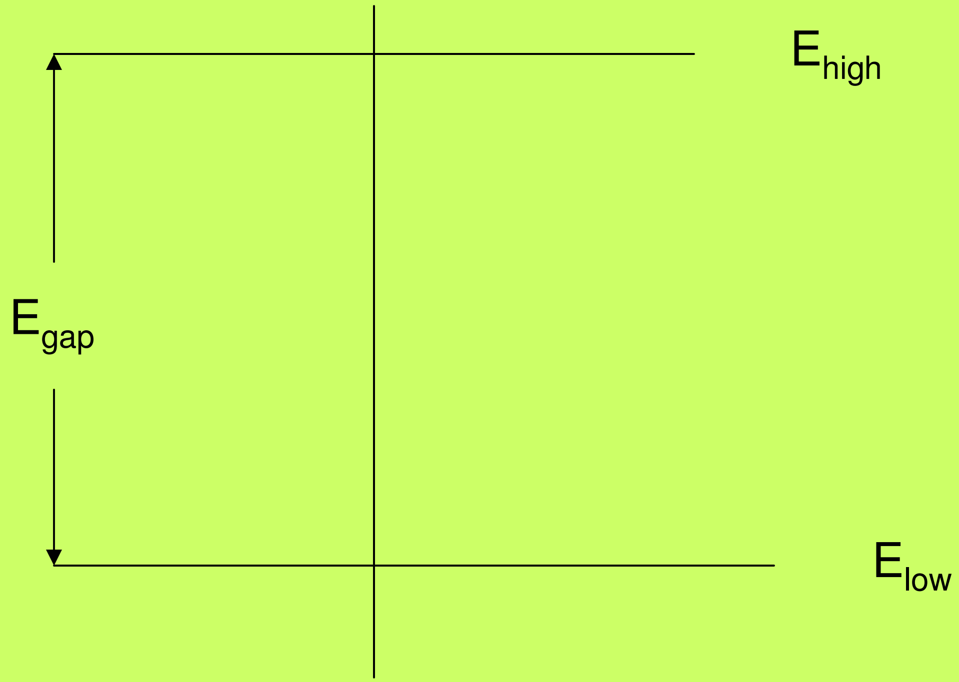
$$\vdots$$

$$+ w_{(n-1)n}x_{(n-1)}x_n]$$

- *E* has *[n(n-1)]/2* terms
- Nature of each term
  - w_{ij} is a real number
  - x_i and x_j are each +1 or -1

No. of steps taken to reach stability

- $E_{gap} = E_{high} - E_{low}$



Analysis of the weights and consequent E_{high} and E_{low}

- W_{il} is any weight with upper and lower bounds as W_{max} *and* W_{min} respectively. Suppose

  $$w_{min} \leq w_{ij} \leq w_{max}$$

  $$w_{max} > w_{min}$$

- Case 1: $w_{max} > 0, w_{min} > 0$

  $$E_{high} = (1/2) * w_{max} * n * (n-1)$$

  $$E_{low} = -(1/2) * w_{max} * n * (n-1)$$

Continuation of analysis of E_{high} and E_{low}

- Case 2: $w_{min} < 0, w_{max} < 0$

  $E_{high} = (1/2) * w_{min} * n * (n-1)$

  $E_{low} = -(1/2) * w_{min} * n * (n-1)$

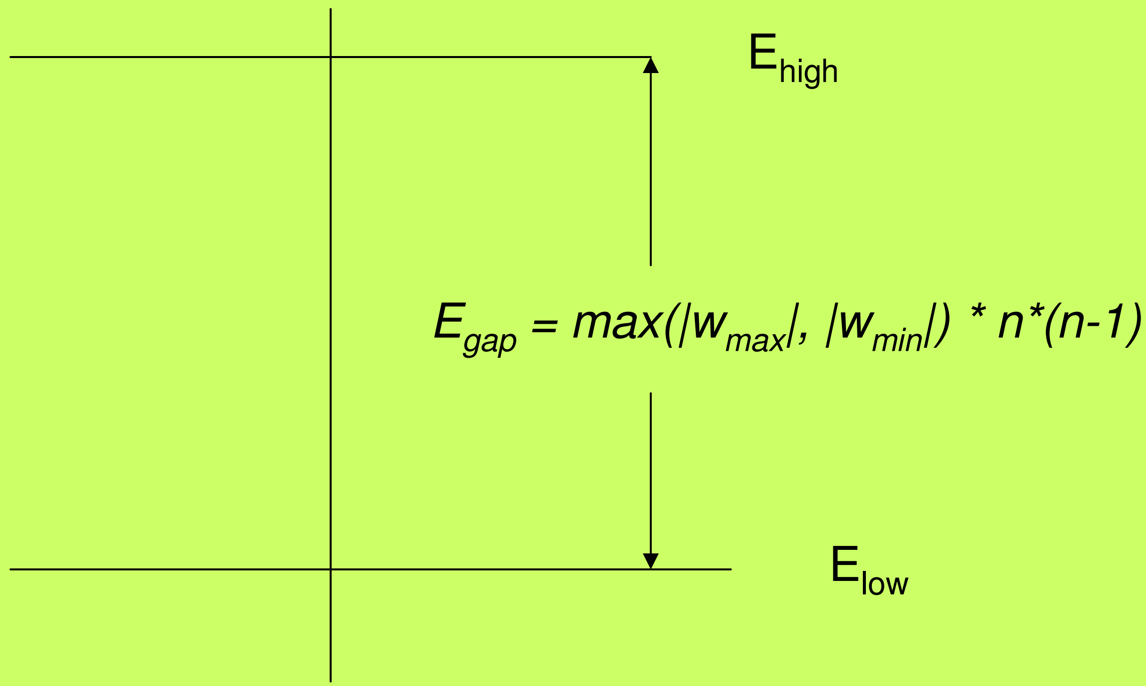
- Case 3: $w_{max} > 0, w_{min} < 0$

  $E_{high} = (1/2) * max(|w_{max}|,|w_{min}|) * n*(n-1)$

  $E_{low} = -(1/2) * max(|w_{max}|,|w_{min}|) * n*(n-1)$

The energy gap

- In general,



To find ΔE_{min}

$\Delta E_p = (x_p^{initial} - x_p^{final}) * net_p$

where ΔE_p is the change in energy due to the p^{th} neuron changing activation.

$$
\begin{aligned}
|\,\Delta E_p\,| &= |\,(x_p^{initial} - x_p^{final}) * net_p| \\
&= 2 * |\,net_p|
\end{aligned}
$$

where $net_p = \Sigma^n_{j=1,\, j\neq p}\, w_{pj} x_j$

To find ΔE_{min}

- $[\Sigma^{n}_{j=1,\, j\neq p}\, w_{pj} x_j]_{min}$ is determined by the precision of the machine computing it.
- For example, assuming 7 places after the decimal point, the value cannot be lower than *0.0000001 [it can be 0, but that is not of concern to us, since the net will continue in the same state]*
- Thus ΔE_{min} is a constant independent of n, determined by the precision of the machine.

Final observation: $o(n^2)$

- It is possible to reach the minimum independent of n.
- Hence in the worst case, the number of steps taken to cover the energy gap is less than or equal to

  $[max(|w_{max}|,|w_{min}|) * n * (n-1)] / constant$
- Thus stability has to be attained in $O(n^2)$ steps

Hopfield Net for Optimization

- Optimization problem
  - Maximizes or minimizes a quantity
- Hopfield net used for optimization
  - Hopfield net and Traveling Salesman Problem
  - Hopfield net and Job Scheduling Problem

The essential idea of the correspondence

- In optimization problems, we have to *minimize* a quantity.
- Hopfield net minimizes the energy
- THIS IS THE CORRESPONDENCE