Manipulation of Attractors as a Recurrent Network Paradigm

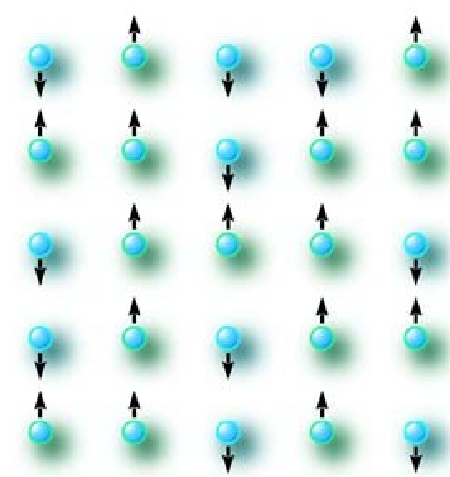
11.11.15.(Tue)

Joon Shik Kim

Ising Model

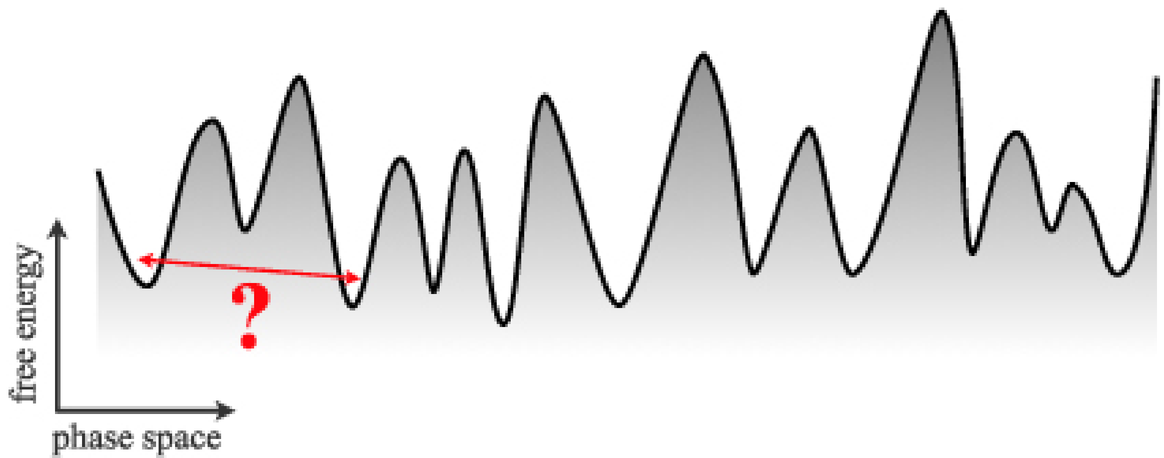
- Spin is a microscopic magnetic moment in an atom.
- Ferromagnetism of Iron was explained with the Ising model.
- $H(S) = -\sum_{i \neq j} J S_i S_j - h \sum_i S_i$, S_i is a spin and h is an external magnetic field.
- J>0 in ferromagnet and

  J<0 in antiferromagnet



Spin Glass

- $H = -\sum_{<ij>} J_{ij} S_i S_j - h \sum_i S_i$
- J_{ij} has a value which depends on i and j.
- There exist many minima.



Additive Model of a Neuron

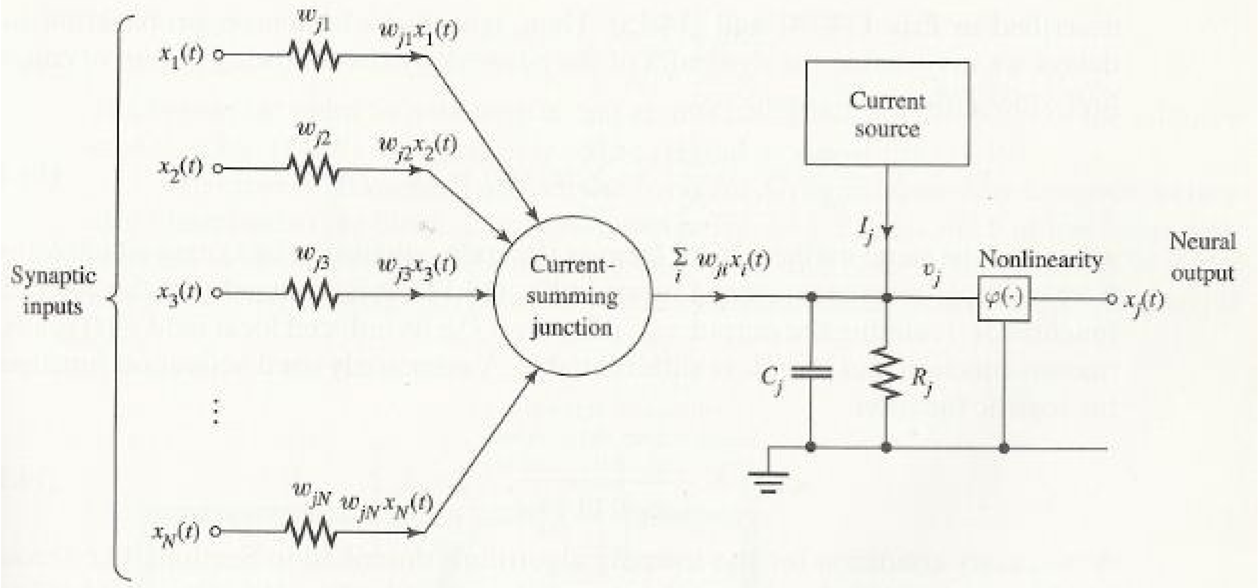


FIGURE 14.7 Additive model of a neuron.

$$C_j \frac{dv_j(t)}{dt} + \frac{v_j(t)}{R_j} = \sum_{i=1}^{N} w_{ji} x_i(t) + I_j \qquad (1)$$

$$x_j(t) = \varphi(v_j(t))$$

Architectural Graph of a Hopfield Network

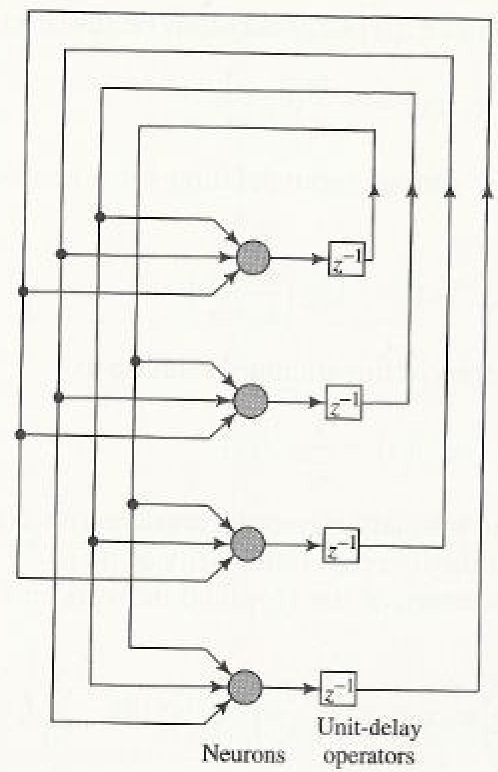


FIGURE 14.9 Architectural graph of a Hopfield network consisting of $N = 4$ neurons.

Recurrent network allows a feedback loop.

The Energy (Lyapunov) function (1/4)

- $E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ji} x_i x_j + \sum_{j=1}^{N} \frac{1}{R_j}\int_0^{x_j} \varphi_j^{-1}(x)dx - \sum_{j=1}^{N} I_j x_j$
- Differentiating *E* with respect to time, we get

$$\frac{dE}{dt} = -\sum_{j=1}^{N}\left(\sum_{i=1}^{N} w_{ji} x_i - \frac{v_j}{R_j} + I_j\right)\frac{dx_j}{dt}$$

$$\frac{dE}{dt} = -\sum_{j=1}^{N} C_j \left(\frac{dv_j}{dt}\right)\frac{dx_j}{dt}$$

$$\frac{dE}{dt} = -\sum_{j=1}^{N} C_j \left[\frac{d}{dt}\varphi_j^{-1}(x_j)\right]\frac{dx_j}{dt}$$

$$= -\sum_{j=1}^{N} C_j \left(\frac{dx_j}{dt}\right)^2 \left[\frac{d}{dx_j}\varphi_j^{-1}(x_j)\right]$$

The Energy (Lyapunov) function (2/4)

- 1. The energy function E is a Lyapunov function of the continuous Hopfield model.
- 2. The model is stable in accordance with Lyapunov's Theorem 1.
- Theorem 1. The equilibrium state $\bar{x}$ is stable if in a small neighborhood of $\bar{x}$ there exists a positive definite function $V(x)$

  such that its derivative with respect to time is semidefinite in that regions.

The Energy (Lyapunov) function (3/4)

- V(x) is positive definite in the state space L if, for all x in L, it satisfies the following requirements:
- 1. V(x) has continuous partial derivatives with respect to the elements of the state vector x.
- 2. $V(\bar{x}) = 0$.
- 3. $V(x) > 0$ if $x \neq \bar{x}$
- According to Theorem 1, if V(x) is a Lyapunov function, the equilibrium state $\bar{x}$ is stable if $\frac{d}{dt} V(x) < 0$ for $x \in U - \bar{x}$

The Energy (Lyapunov) function (4/4)

- The time evolution of the continuous Hopfield model described by the nonlinear first-order differential equation (1) represents a trajectory in state space, which seeks out the minima of the energy (Lyapunov) function E and comes to a stop at such fixed points.

The Discrete Hopfield (1/5)

- Content-Addressable Memory
- Storage Phase. We wish to store a set of *N*-dimensional vectors, denoted by $\{\xi_\mu | \mu = 1,2,\ldots,M\}$.
- Synaptic weight from neuron *i* to neuron *j* is defined by

$$w_{ji} = \frac{1}{N} \sum_{\mu=1}^{M} \xi_{\mu,j} \xi_{\mu,i}$$

The Discrete Hopfield (2/5)

- We set $w_{ii} = 0$ for all i, which means that the neuron has no self-feedback.
- Then the synaptic weight matrix is
$$W = \frac{1}{N}\sum_{\mu=1}^{M} \xi_\mu \xi_\mu^T - MI.$$
- Retrieval Phase. The state updating from one iteration to the next is deterministic, but the selection of a neuron to perform the updating is done randomly, i.e., asynchronous learning.

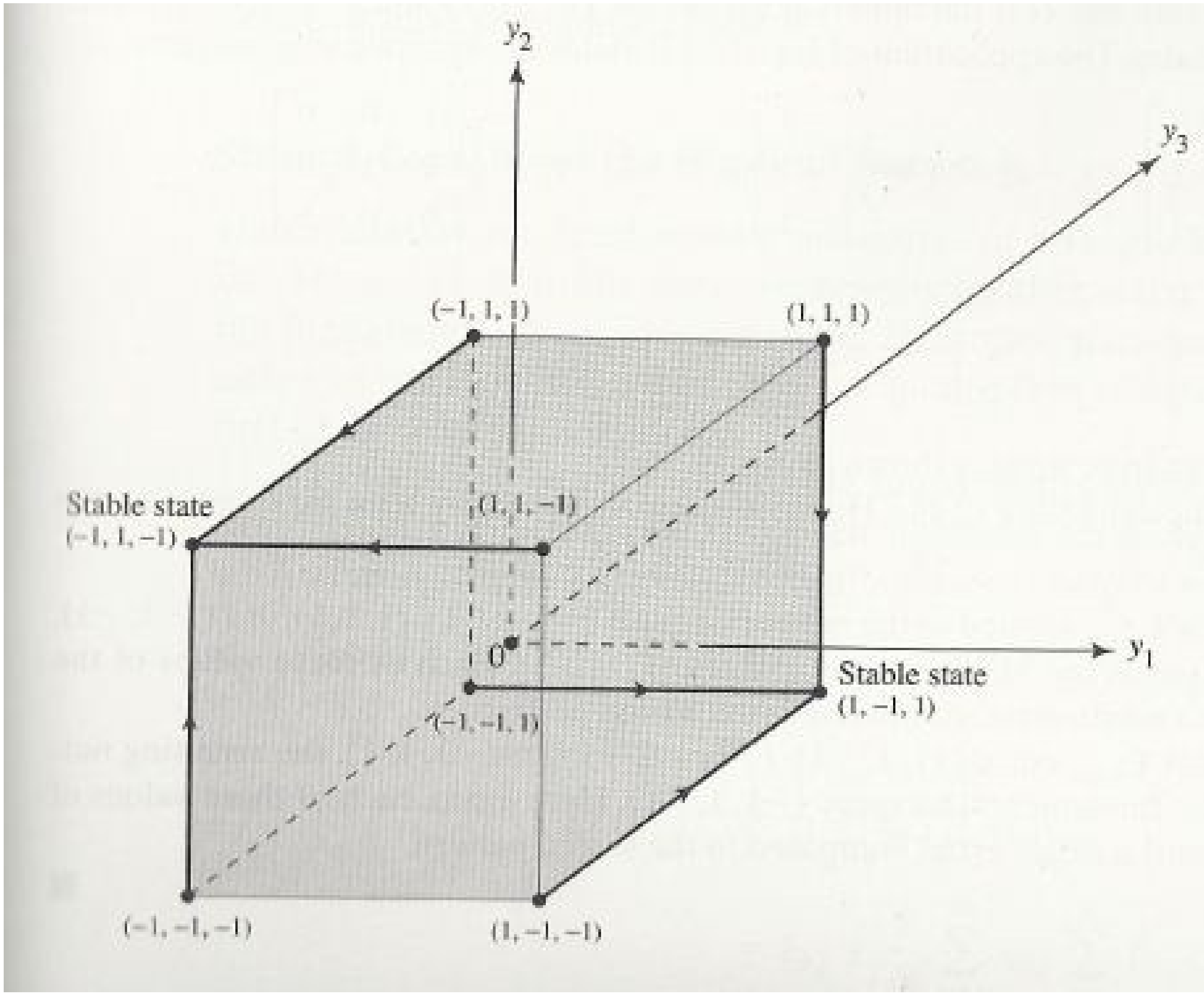
The Discrete Hopfield (3/5)

- $\mathbf{y}(t+1) = \text{sgn}(\mathbf{W}\mathbf{y}(t) + \mathbf{b})$, where sgn(x) is a signum function.
- sgn(x) = x/|x|, if x ≠ 0 and
  sgn(x) = 0, if x = 0.
- Example.

$$\mathbf{W} = \frac{1}{3}\begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix}[+1, -1, +1] + \frac{1}{3}\begin{bmatrix} -1 \\ +1 \\ -1 \end{bmatrix}[-1, +1, -1] - \frac{2}{3}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -2 & +2 \\ -2 & 0 & -2 \\ +2 & -2 & 0 \end{bmatrix}.$$

The Discrete Hopfield (4/5)



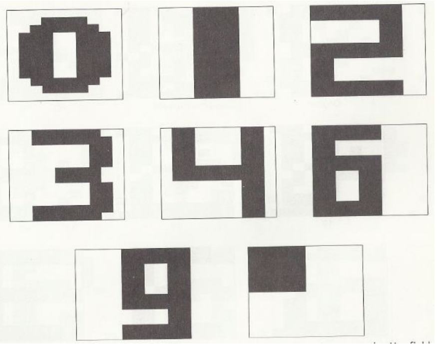
The Discrete Hopfield (5/5)

- $$\mathbf{W}\mathbf{y} = \frac{1}{3}\begin{bmatrix} 0 & -2 & +2 \\ -2 & 0 & -2 \\ +2 & -2 & 0 \end{bmatrix}\begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} +4 \\ -4 \\ +4 \end{bmatrix}$$, limiting this result yields

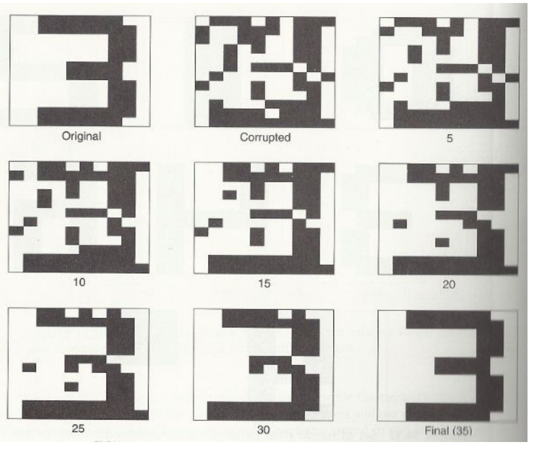
$$\operatorname{sgn}[\mathbf{W}\mathbf{y}] = \begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix} = \mathbf{y}.$$

Computer Experiment (1/2)

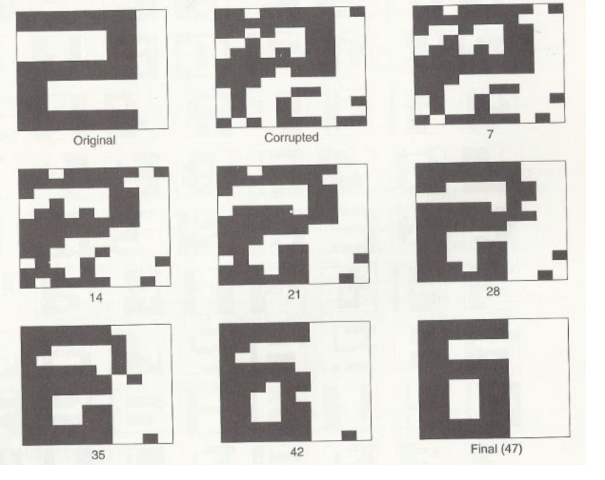
- Set of handcrafted patterns



Correct recollection of corrupted 3

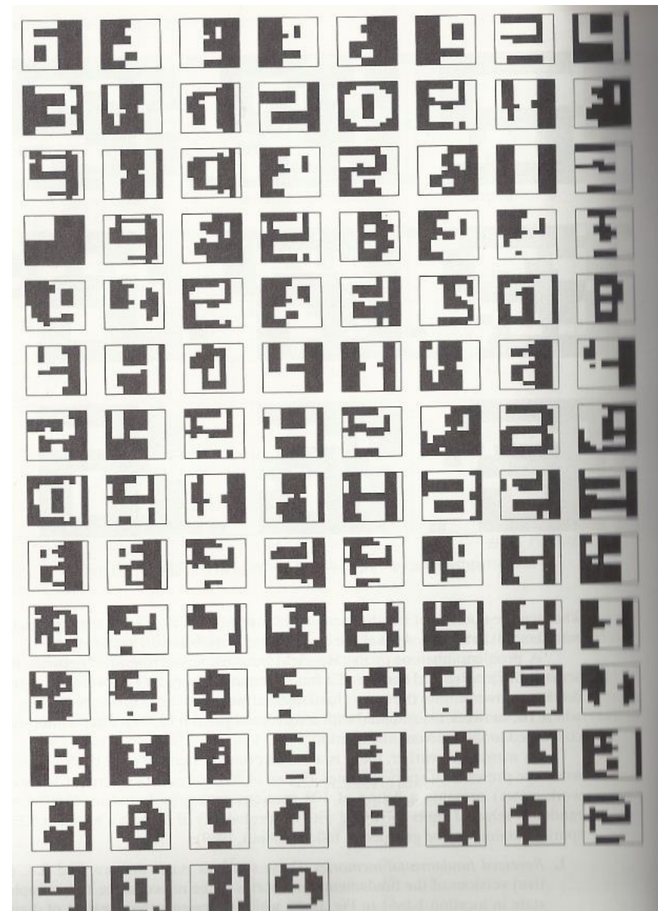


Incorrect recollection of corrupted 2



Computer Experiment (2/2)

- Load parameter $\alpha = \frac{M}{N}$
- Critical value $\alpha_c = 0.14$
- Mixture states: a state formed out of three fundamental memories.
  row 6, column 4: negative of digit 1, digit 4, and digit 9.
- Spin-glass states: local minima of the energy landscape that are not correlated with any of the fundamental memories of the network.
  Row 7, column 6.



Reference

- S. Haykin, Neural networks, New Jersey: Prentice Hall, 1999.