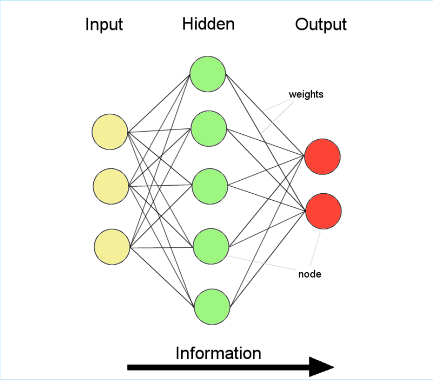
هوش محاسباتی:
شبکه های عصبی مصنوعی

دانشکده مهندسي کامپيوتر
دانشگاه علم و صنعت ایران
ناصر مزيني mozayani@iust.ac.ir

مدلهای بازگشتي

Recurrent networks

Introduction



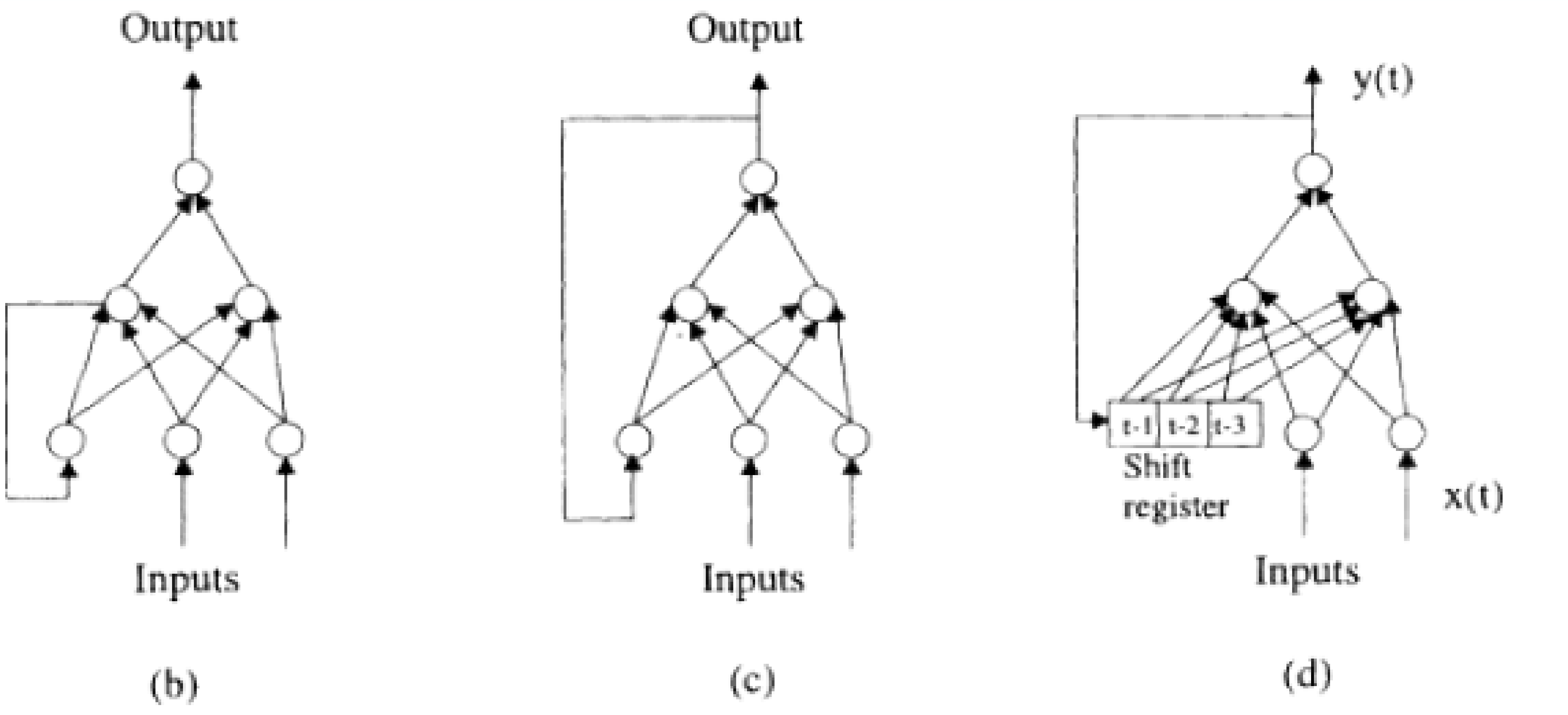
- Feed forward networks:
  - Information only flows one way
  - One input pattern produces one output
  - No sense of time (or memory of previous state)
- Recurrency
  - Nodes connect back to other nodes or themselves
  - Information flow is multidirectional
  - Sense of time and memory of previous state(s)
- Biological nervous systems show high levels of recurrency (but feed-forward structures exists too)

Introduction

- Inspired from biology
- Common features:
  - Nonlinear processing units
  - Many feedback connections
- Most common models:
  - Partially connected
    - Elman Network
    - Jordan Network
    - LSTM
  - Fully connected
    - Hopfield Network
    - Boltzman machine
    - Mean-Field-Theory machine
    - BSB

Introduction

– Partially connected

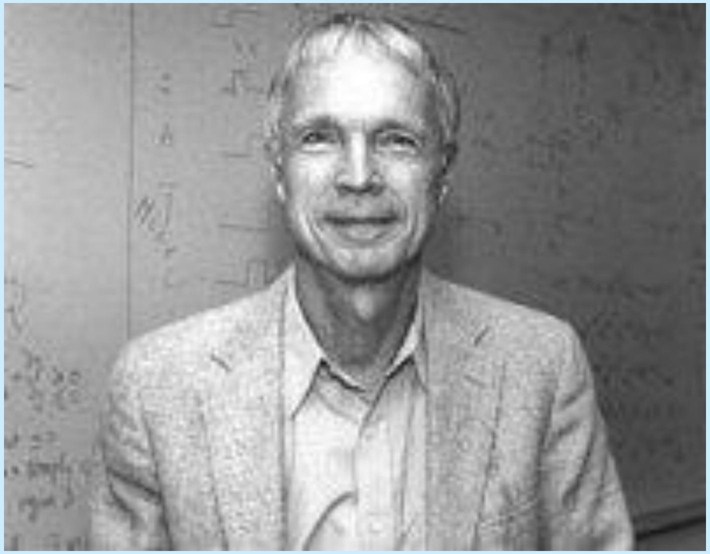


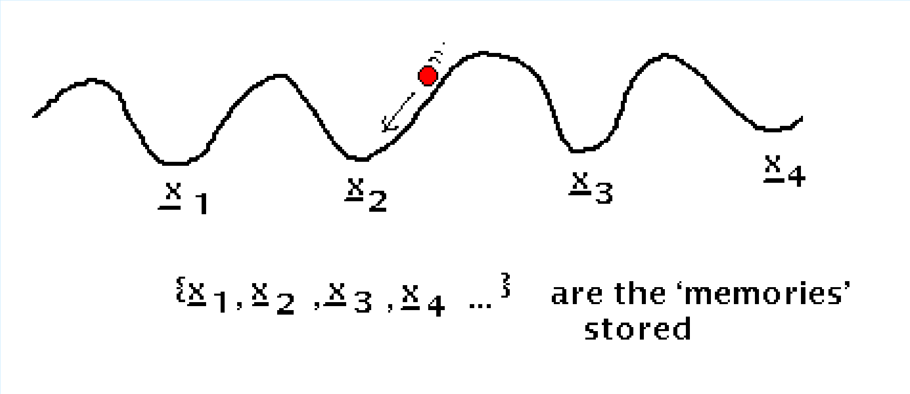
Elman

Jordan

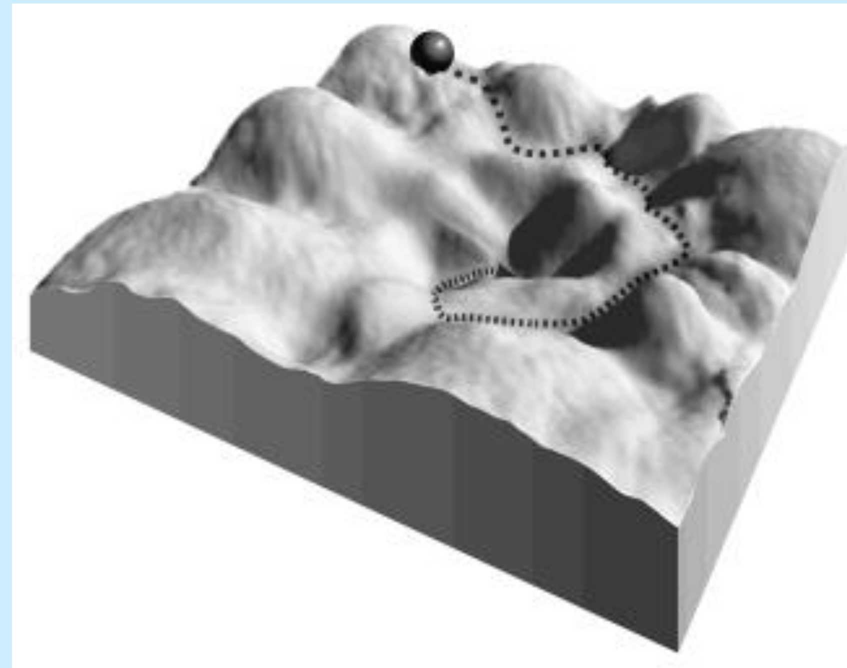
Hopfield model

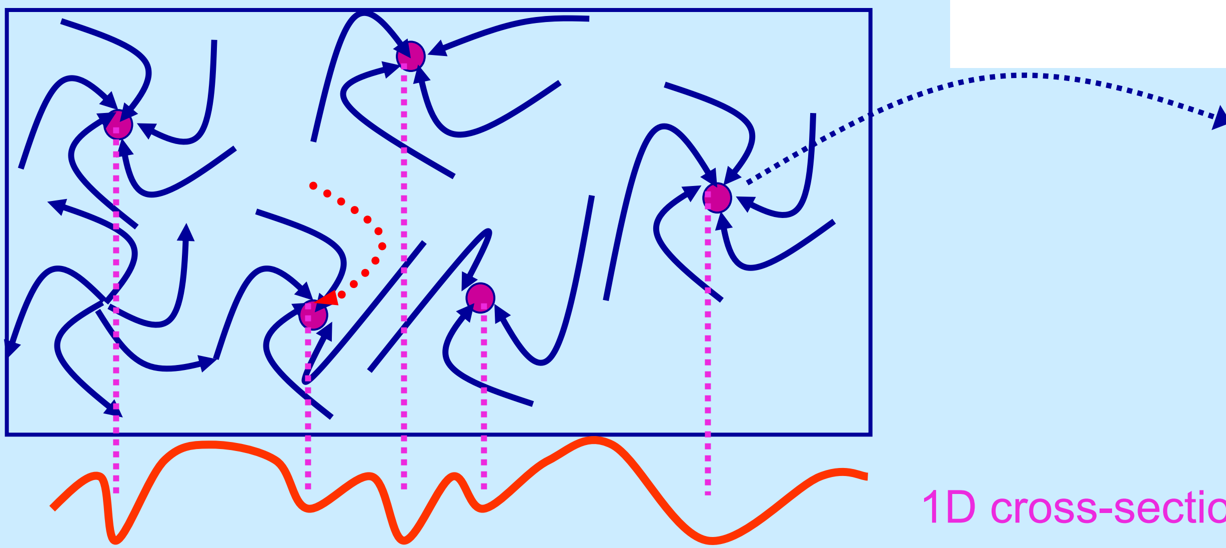
- Inspired from Physics
- unsupervised
- Dynamic interactive behaviour
- Natural learning protocols





Hopfield model

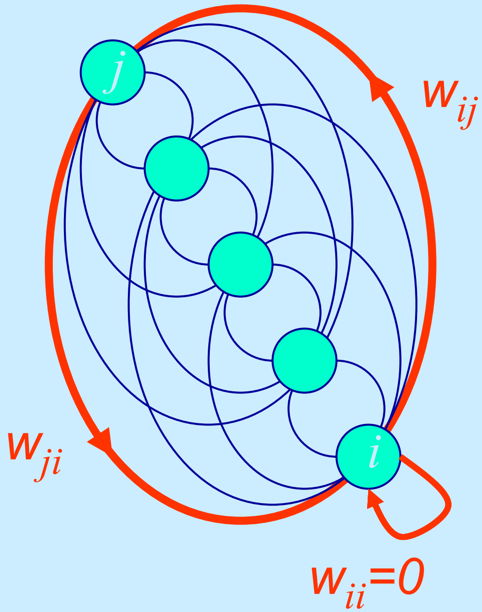




Attractors

1D cross-section of energy landscape

Hopfield Model



The restrictions imposed on the recurrent net are now:

- All connections are symmetric $\boldsymbol{w_{ij} = w_{ji}}$
- No self-connectedness

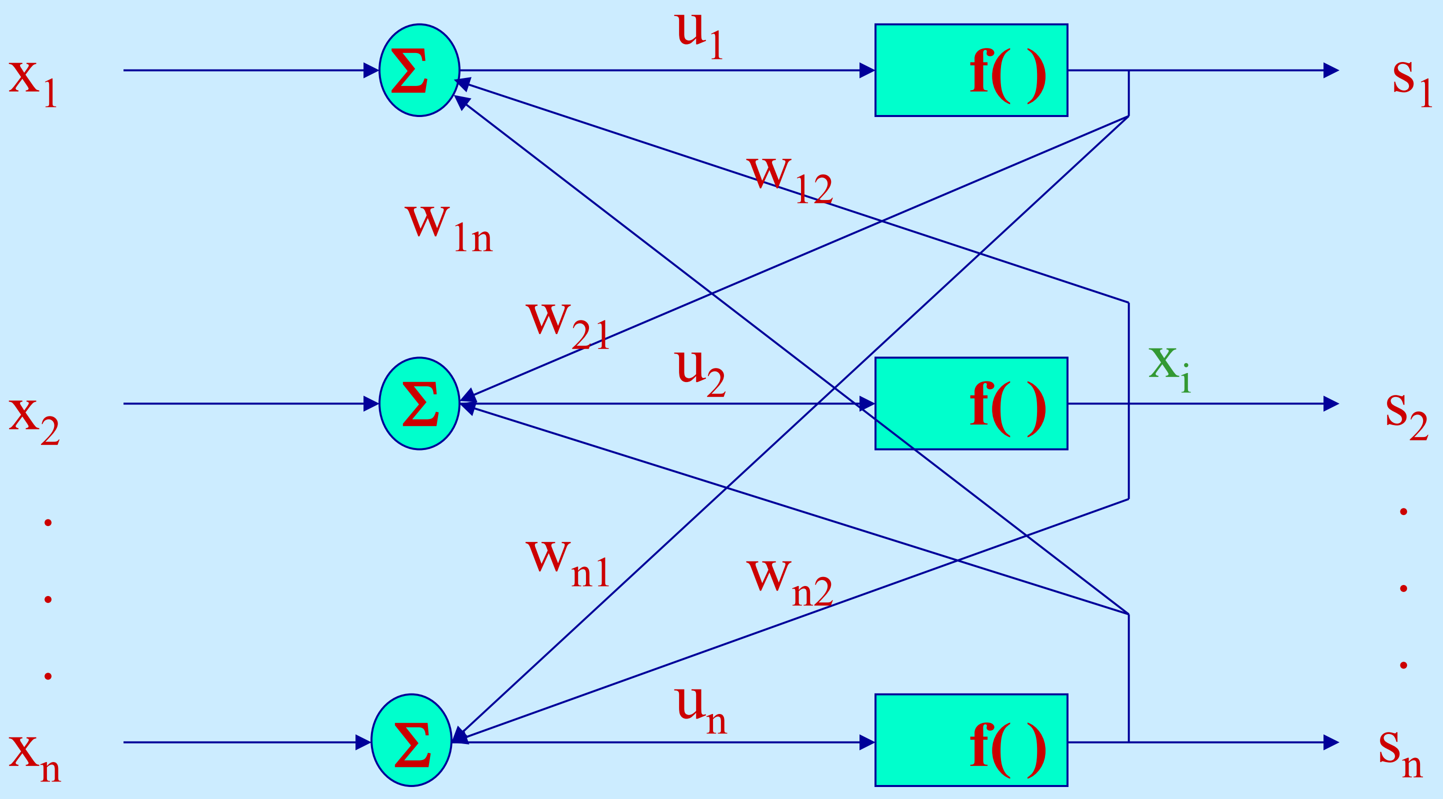
"Gerard Toulouse has called Hopfield's use of symmetric connections a 'clever step backwards from biological realism'. The cleverness arises from the existence of an energy function."

We will now see why having such an energy function provides us with:

- A formalism of the process of memory storage and recall
- A tool to visualise the activity (both learning and recall)
- A straightforward way to train the net
- Once trained, a guaranteed solution (recall of the correct memory).

Hopfield model

Network Architecture:



How Does It Work?

-Every node is modelled by a conventional binary McPitts neuron.

-Each neuron serves both as an input and output unit. (There are no hidden units.)

-Each state is given by the pattern of activity of the neurons (e.g. 101 for a network with three neurons). Therefore, the number of neurons in the net sets the maximum length for a bit-string to be remembered.

-Different patterns can be simultaneously stored in the network weights.

How Does It Work?

-The number of independent patterns or words that can be remembered is less than or equal to the number of nodes in the net.

-Memory recall corresponds to a trajectory taking the system from some initial state (input) to the local energy minimum (closest association) .

-Each step along the (recall) trajectory results in the same or lower energy.

-Since energy is bounded from below, a solution is guaranteed for every input.

Activation function

f( ) **can be**

- **Binary valued hard limiter {0, 1}**

$$f_t(x) = \begin{cases} 1 & x > t \\ 0 & x \le t \end{cases}$$

- **Bipolar valued hard limiter {-1, 1}**

$$f_t(x) = \begin{cases} 1 & x > t \\ -1 & x \le t \end{cases}$$

- **Binary sigmoid [0, 1]**

$$f_{t,\beta}(x) = \frac{1}{2}[1 + \tanh \frac{x-t}{\beta}]$$

- **Bipolar sigmoid [-1, 1]**

$$f_{t,\beta}(x) = \tanh \frac{x-t}{\beta}$$

Hopfield Network

- Network Model and Operation
  - *The formal McCulloch-Pits neuron*
  - *One of the two definite states "**on**" or "**off**"*
  - *Synaptic connection w_{ji} between neuron i and j, determines the contribution of output x_i of neuron i to activation potential of neuron j*
  - *Weights determined according to a Hebbian rule:*

$$w_{ij} = \sum_{k=1}^{P} x_i^k x_j^k$$

Network relations

Network dynamics is determined by some differential equations:

$$C_j \frac{dv_j(t)}{dt} = -\frac{v_j(t)}{R_j} + \sum_{i=1}^{N} w_{ji} x_i(t) + I_j, \quad j = 1, 2, \ldots, N$$

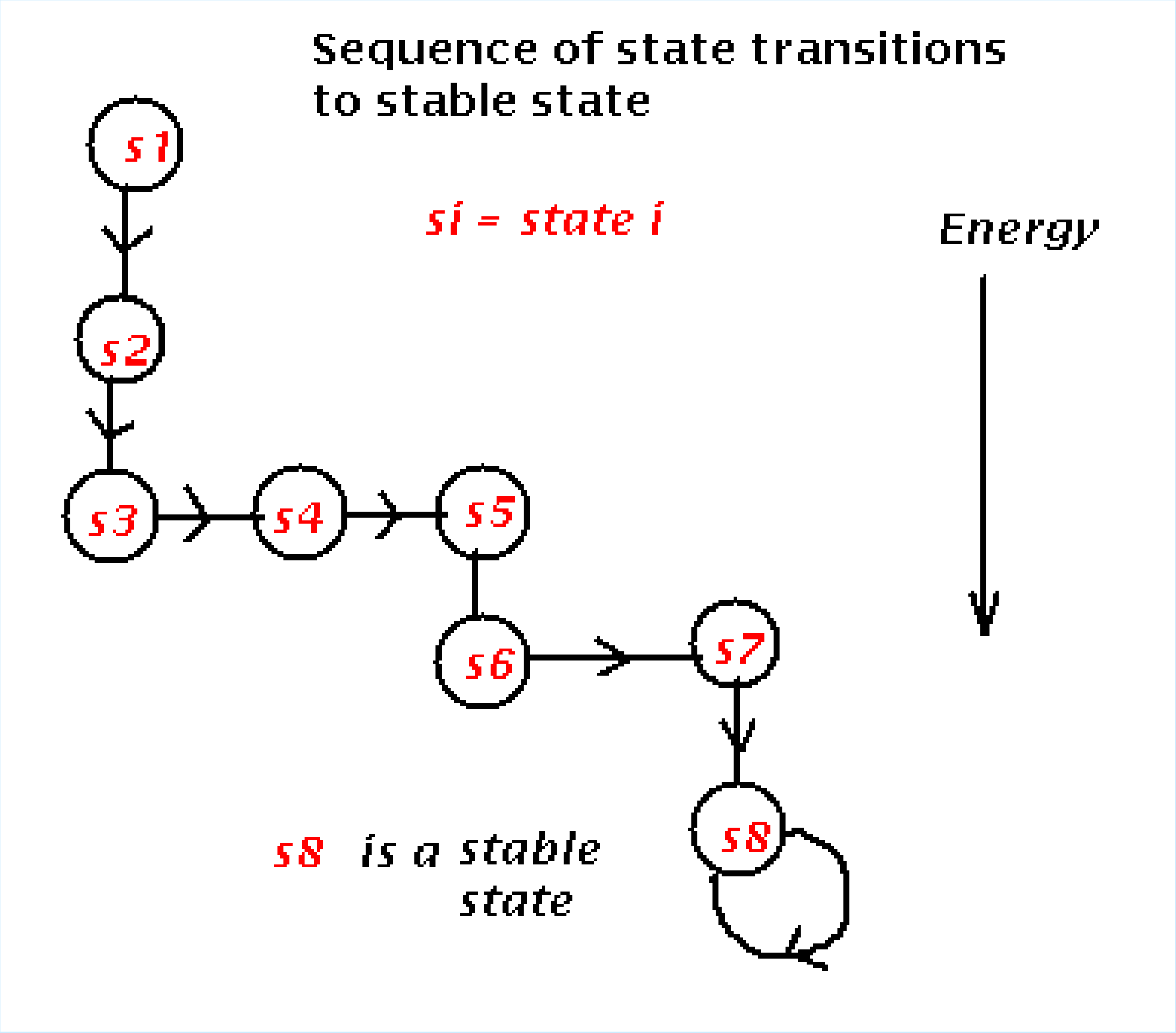
An energy function is defined (Liapunov):

$$E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ji} x_i x_j + \sum_{j=1}^{N} \frac{1}{R_j} \int_0^{x_j} \varphi_j^{-1}(x)\, dx - \sum_{j=1}^{N} I_j x_j$$

Associative memory

- Retrieval of a stored pattern
  - *Update of a randomly selected neuron*
  - *This asynchronous approach continues till there will not be any available neuron for state changes*
- Associative memory model
  - *Network should memorize the most resembling one of a set of P patterns while a pattern is presented to its input*
  - Resembles = Hamming distance
  - *Content-addressable structure or autoassociative*
  - *Robustness against to small errors in the input patterns*
  - *Recognition and reconstruction of images, and retrieval of data with incomplete information*

Associative memory



Simple example

Weight matrix

(neuron j) \ (neuron i)	1	2	3	4	5
1	0	1	-1	2	-3
2	1	0	3	-1	0
3	-1	3	0	1	-2
4	2	-1	1	0	1
5	-3	0	-2	1	0

threshold = 0

	1	2	3	4	5
Input (t=0)	0	1	0	0	0
t=1	1	1	1	0	1
t=2	0	1	1	1	0
t=3	1	1	1	1	0
t=4	1	1	1	1	0
⋮					
t	1	1	1	1	0
$\sum_i x_i w_{ij}$	1	0	3	-1	0
	-3	4	0	3	-5
	2	2	4	0	-1
	2	3	3	2	-4

The stable pattern represents a fixed point in the dynamics. While stable solutions are guaranteed, in fact, not all stable solutions are fixed point solutions.

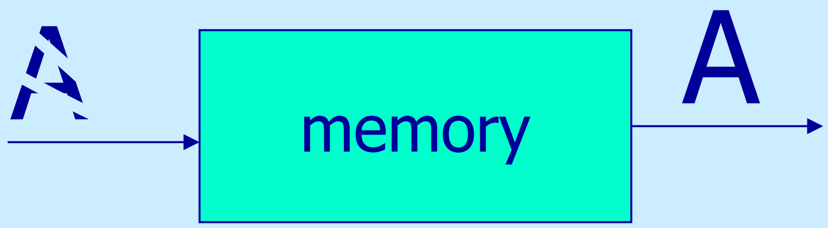
Exercise: repeat this example with an initial input of [0 1 0 1 0].

Note: this example used a "synchronous updating" method. Asynchronous (sequential or random) updating methods can also be implemented.

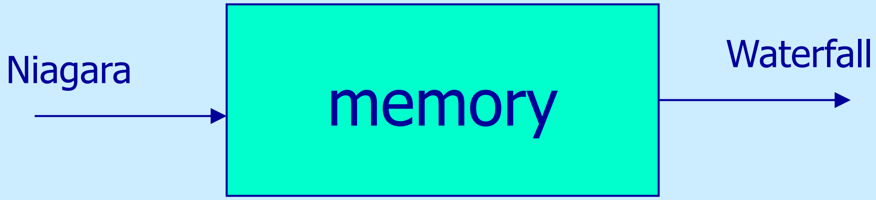
Associative memory

- Concept:
  - Object or pattern *A* (input) reminds the network of object or pattern *B* (output)
- Hetero-associative Vs. auto-associative memories
  - If *A* and *B* are different, the system is called hetero-associative net
  - If *A* and *B* are the same, the system is called auto-associative net

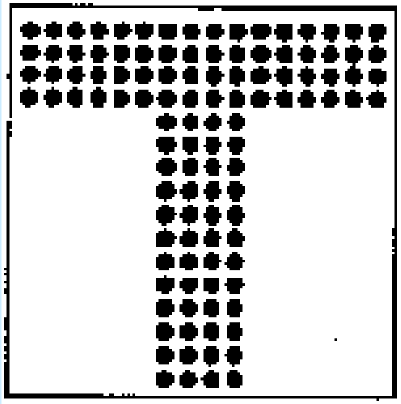
Auto-association:



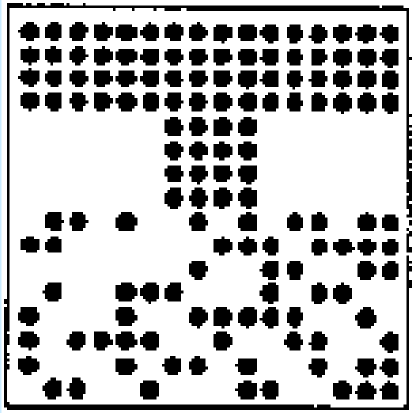
Hetero-association:



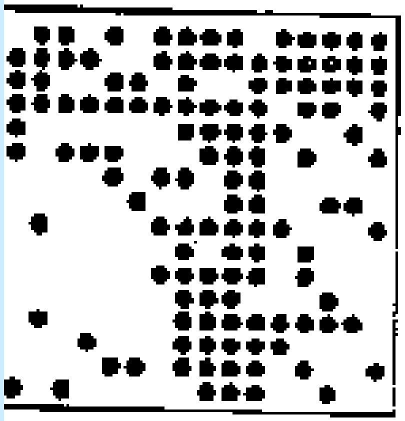
Associative memory



Original 'T'

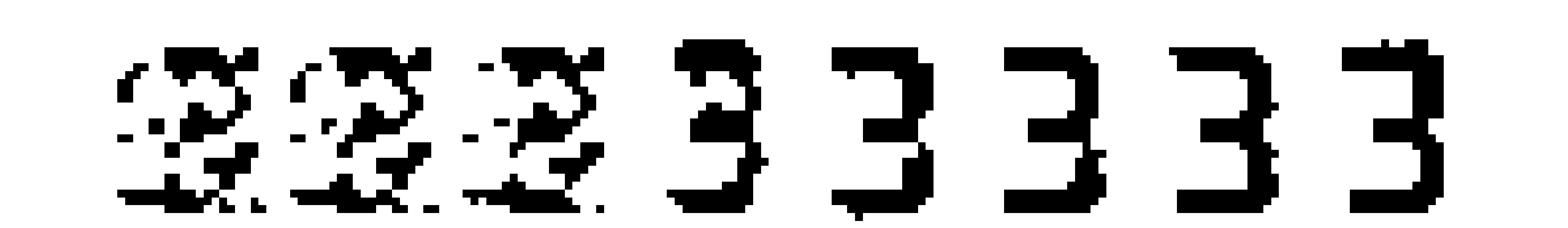


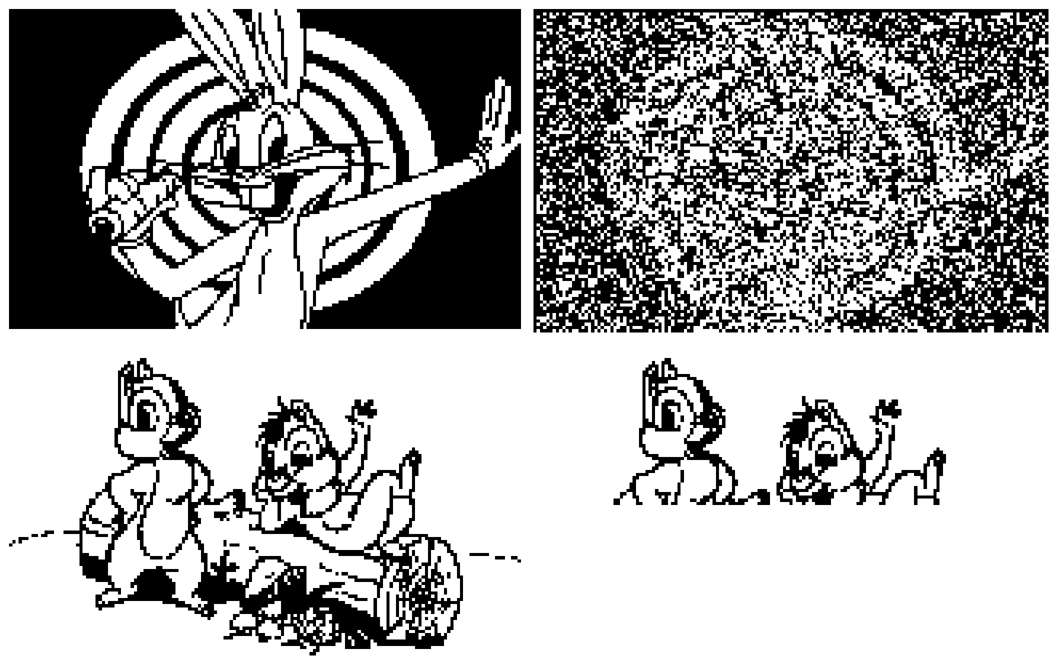
half of image corrupted by noise



20% corrupted by noise (whole image)

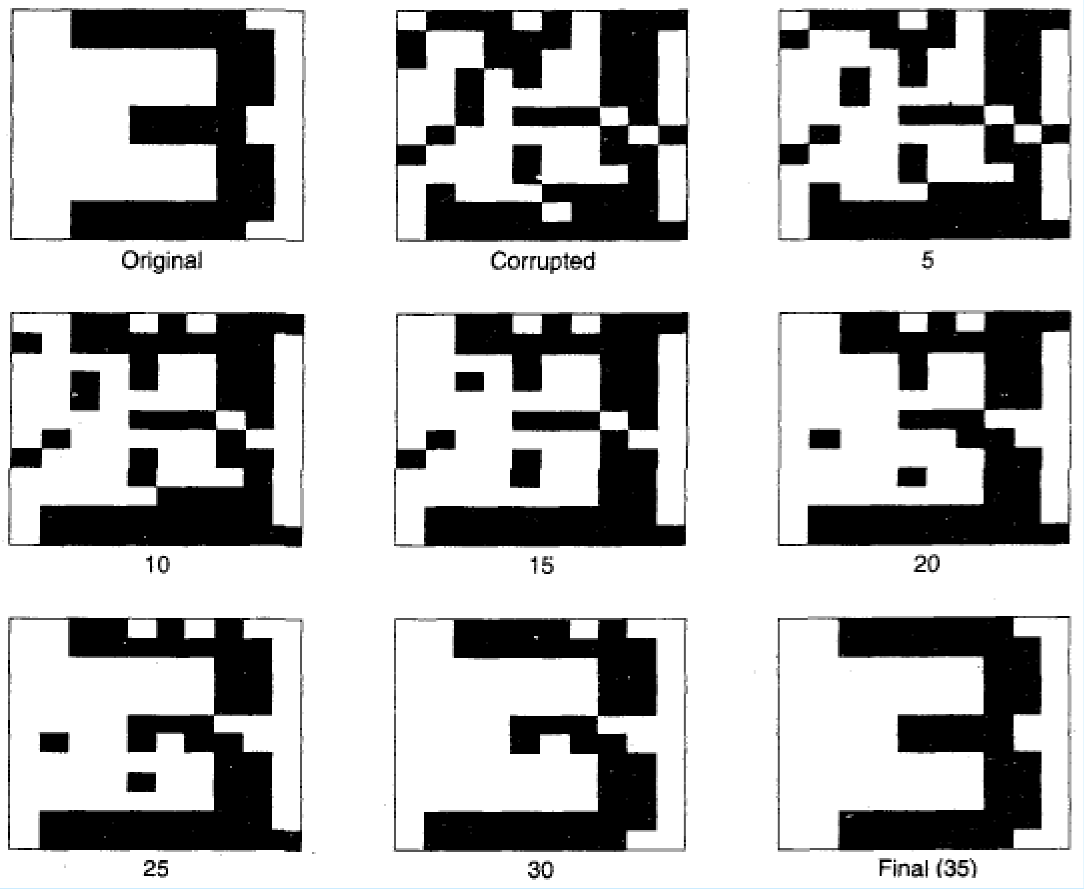
Hopfield Networks



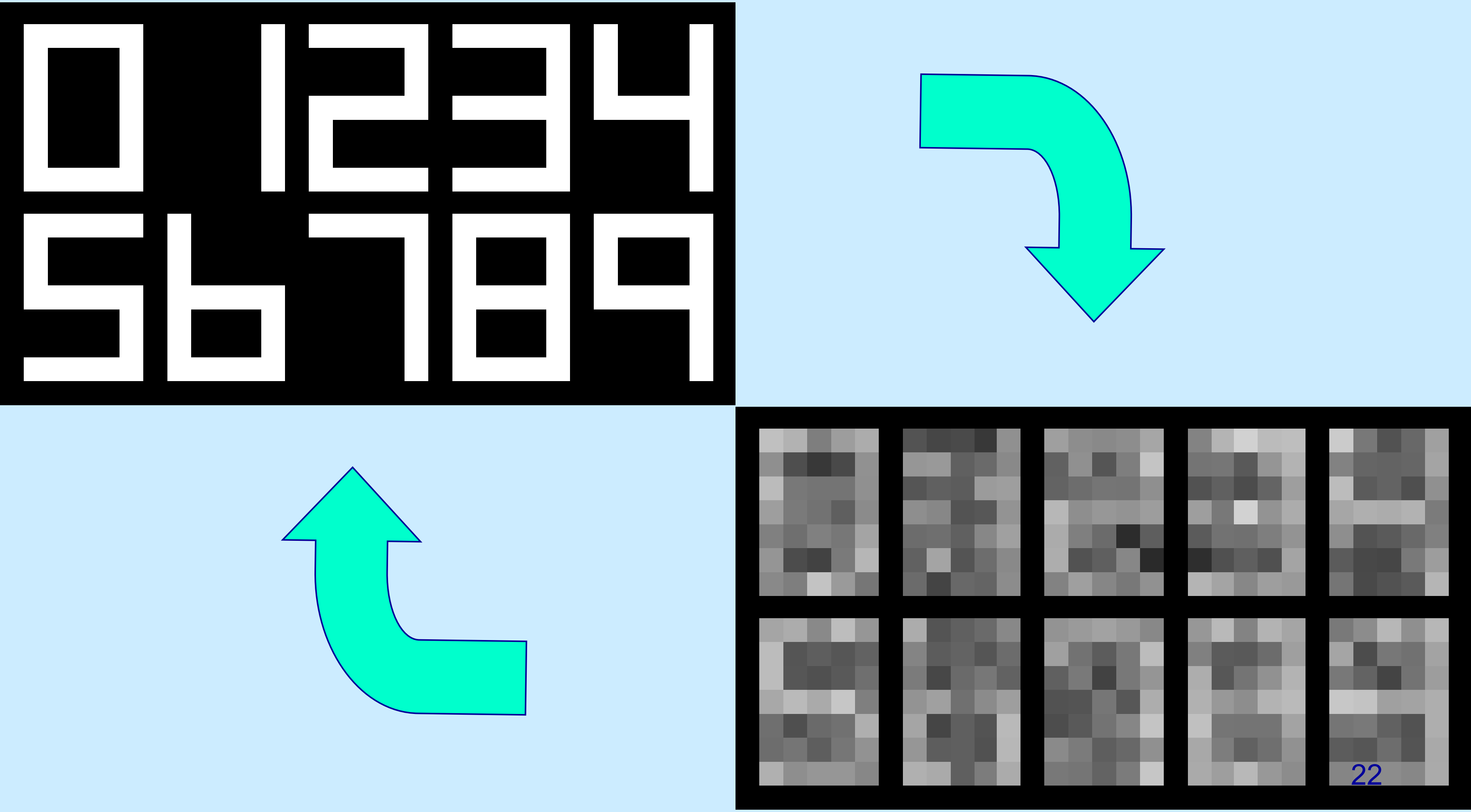


Original Degraded

Pattern retrieval



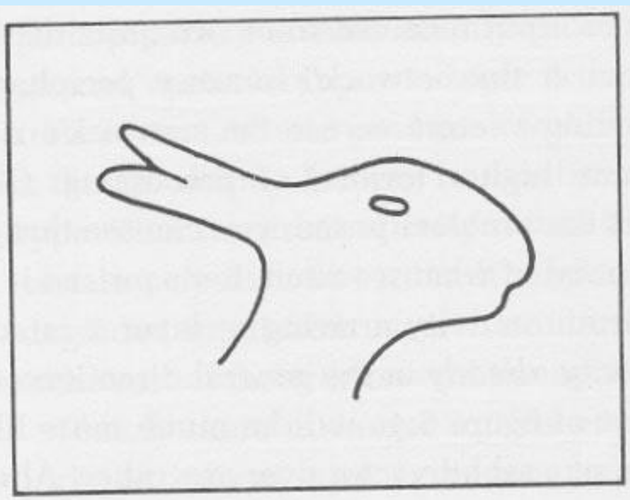
Hopfield Networks



Hopfield Nets in the Brain??

- The cerebral cortex is full of recurrent connections, and there is solid evidence for Hebbian synapse modification there. Hence, the cerebrum is believed to function as an associative memory.

- Flip-flop figures indicate distributed hopfield-type coding, since we cannot hold both perceptions simultaneously (binding problem)



(b)

Some properties for Hopfield

- Capacity C= *the maximum number of patterns that can be stored without unacceptable error*

  C=N/4lnN (Weisbuch and Fogelman-Soulie 1985)
- The network contains some spurious attractors (local minima of the energy)
  - Reversed sign
  - A linear combination of true patterns – mixture states
- All starting patterns with more that half the bits different from the original one will end up in reverse state
- Used for optimization problems

Phase diagram