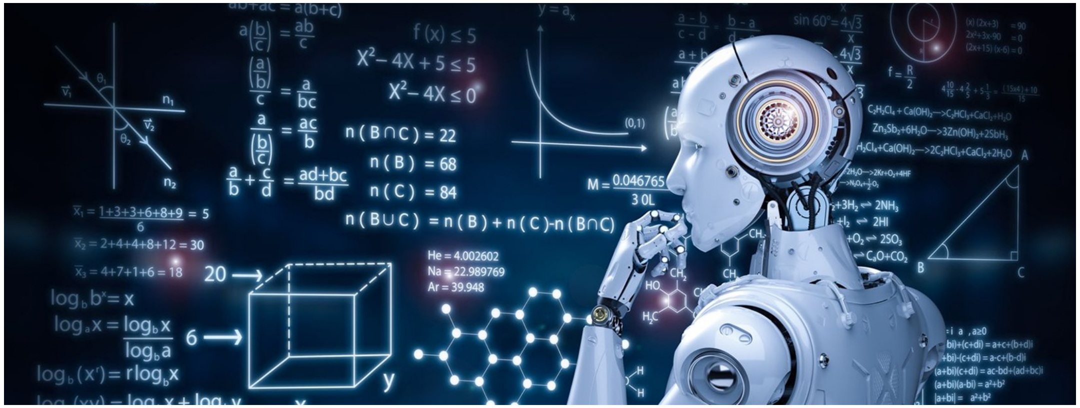
Tutorial 9

CS5486: Intelligent Systems

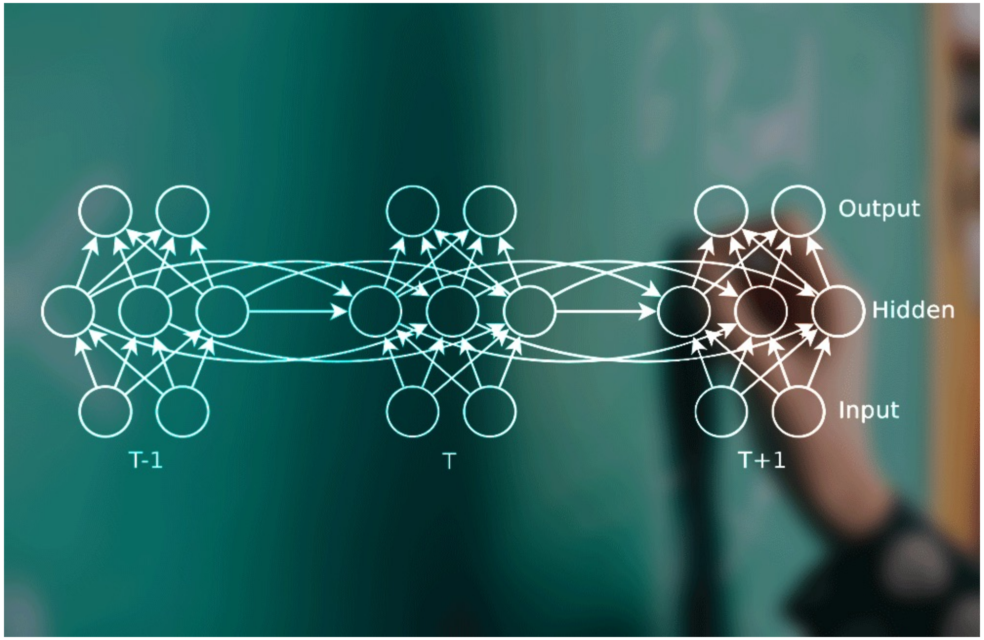


Instructor: Yixuan LI
Yixuan.Li@my.cityu.edu.hk
Computer Science
City University of Hong Kong

Outline

1. **Recurrent Neural Network**
2. **Discrete Hopfield Network**
3. **Continuous Hopfield Network**

1. Recurrent Neural Network



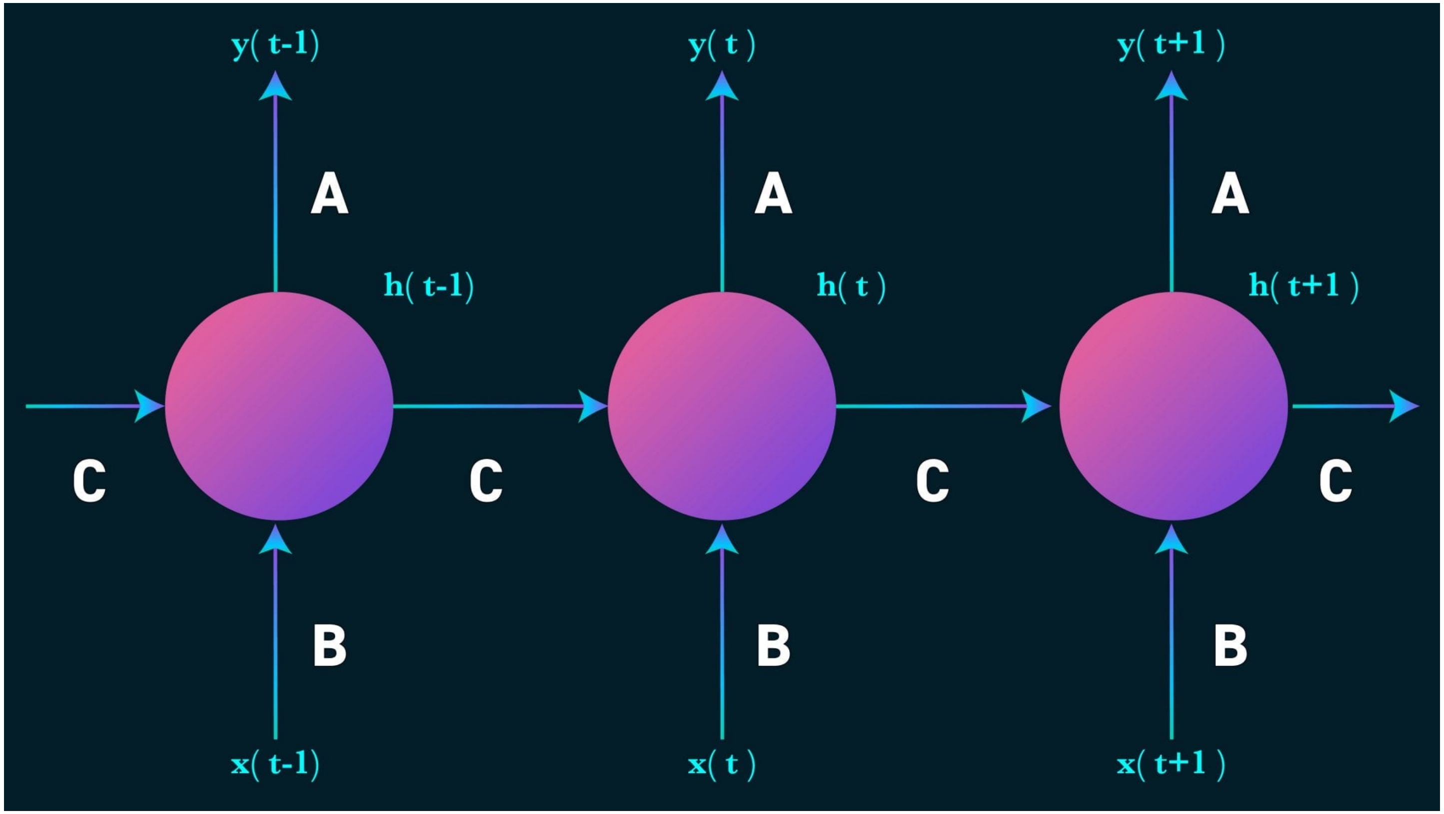
A recurrent neural network (RNN) is a type of artificial neural network which uses sequential data or time series data.

These deep learning algorithms are commonly used for ordinal or temporal problems, such as language translation, natural language processing (NLP), speech recognition, and image captioning; they are incorporated into popular applications such as Siri, voice search, and Google Translate.

1. Recurrent Neural Network

What is a Recurrent Neural Network?

The logic behind an RNN is to consider the sequence of the input. For us to predict the next word in the sentence we need to remember what word appeared in the previous time step. These neural networks are called Recurrent because **this step is carried out for every input**. As these neural network consider the previous word during predicting, it acts like a memory storage unit which stores it for a short period of time.

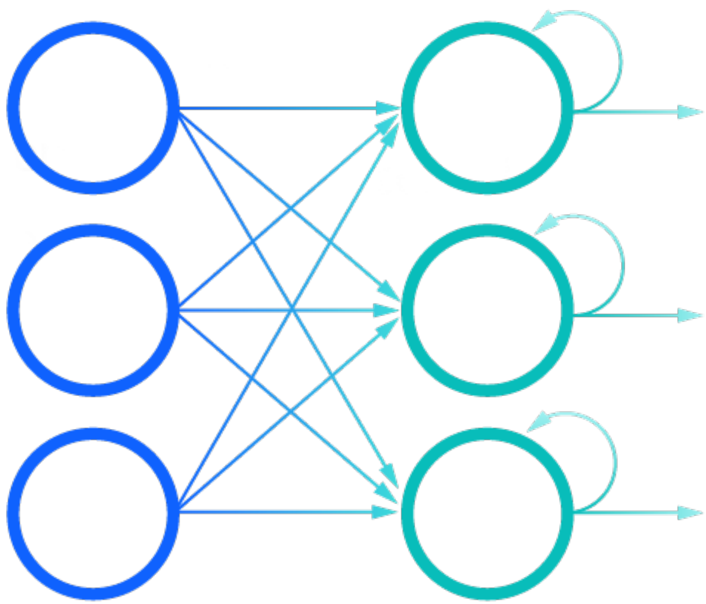


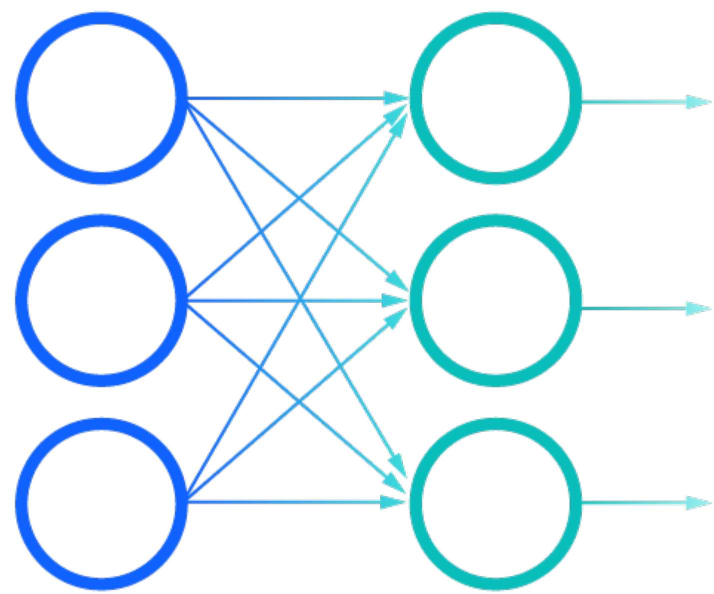
1. Recurrent Neural Network

Recurrent Neural Network vs. Feedforward Neural Network

take information from prior inputs to influence the current input and output

depend on the prior elements within the sequence.





2. Discrete Hopfield Network

The Hopfield neural network model is a kind of recurrent neural network. It was invented by Dr. John J. Hopfield in 1982. It consists of a single layer which contains one or more fully connected recurrent neurons. The Hopfield network is commonly used for auto-association and optimization tasks.

2. Discrete Hopfield Network

According to the discrete or continuous output of the network, the Hopfield network is divided into two types:

- discrete Hopfield neural network (DHNN)
- continuous Hopfield neural network (CHNN)

2. Discrete Hopfield Network

Discrete Hopfield Network: It is a fully interconnected neural network where each unit is connected to every other unit. It behaves in a discrete manner, i.e. it gives finite distinct output, generally of two types:

- **Binary (0/1)**
- **Bipolar (-1/1)**

The discrete Hopfield network (DHN) is a classic recurrent neural network operating with binary or bipolar states and activation function in discrete time as follows:

$$u(t) = Wx(t) - \theta, \tag{6}$$
$$x(t+1) = \sigma(u(t)), \tag{7}$$

where $u \in \mathbb{R}^n$ is the net-input vector, $x \in \mathbb{R}^n$ is the state vector, $W \in \mathbb{R}^{n\times n}$ is the connection weight matrix, $\theta \in \mathbb{R}^n$ is the threshold vector, and $\sigma(\cdot)$ is a vector-valued discontinuous activation function defined element-wisely as follows:

$$x_i(t+1) = \sigma(u_i) = \begin{cases} 0, & \text{if } u_i(t) \le 0 \\ 1, & \text{if } u_i(t) > 0 \end{cases} \qquad x_i(t+1) = \begin{cases} -1, & \text{if } u_i(t) \le 0 \\ +1, & \text{if } u_i(t) > 0 \end{cases}$$

2. Discrete Hopfield Network

The DHN[1] is globally convergent to a local minimum of the following combinatorial optimization problem:

$$\min \quad -\frac{1}{2}x^T Wx + \theta^T x \quad \text{s.t.} \quad x \in \{0,1\}^n. \tag{7}$$

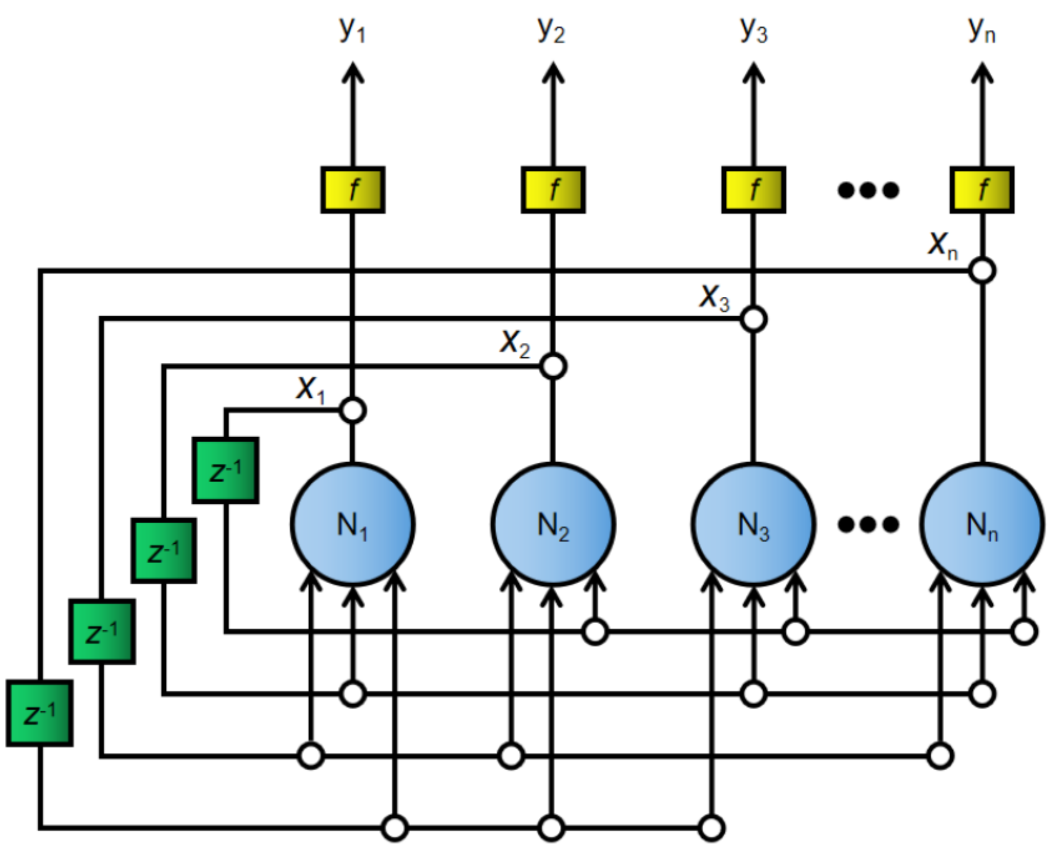
[1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," Proceedings of the National Academy of Sciences, vol. 79, no. 8, pp. 2554–2558, 1982.

In the examination, you are probably required to compute a simple Hopfield network. So you need to understand the **weights, stability and updating rules**.

2. Discrete Hopfield Network

$$u(t) = Wx(t) - \theta,$$

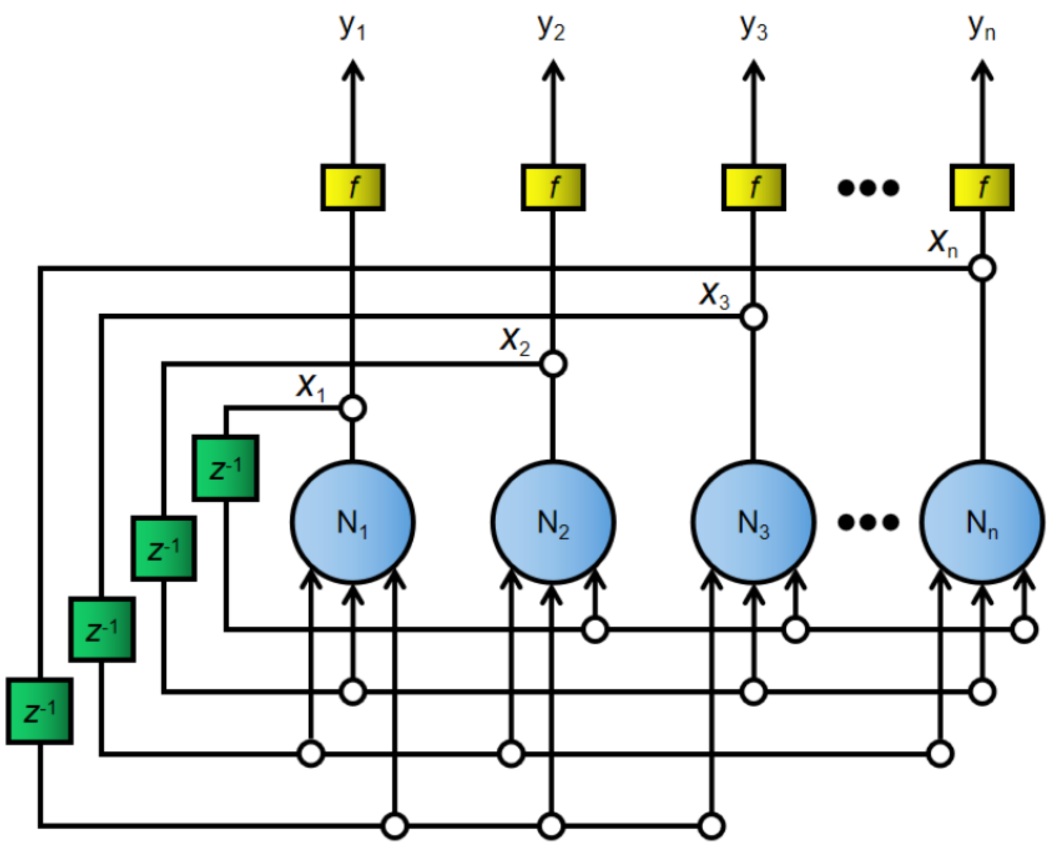
$$x(t+1) = \sigma(u(t)),$$



2. Discrete Hopfield Network

- This model consists of neurons with one inverting and one non-inverting output.

- The output of each neuron should be the input of other neurons but not the input of self.

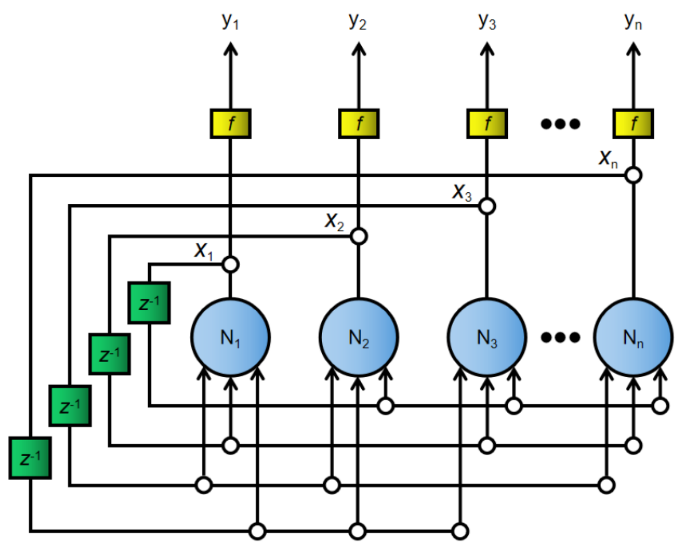


2. Discrete Hopfield Network

- Weight/connection strength is represented by $\mathbf{w_{ij}}$.

- Connections can be excitatory as well as inhibitory. It would be excitatory, if the output of the neuron is same as the input, otherwise inhibitory. Weights should be symmetrical, i.e. $\mathbf{w_{ij} = w_{ji}}$

- Weights should be symmetrical, i.e. $\mathbf{w_{ij} = w_{ji}}$



2. Discrete Hopfield Network: Examples

Let's take an example: Design a network of 3 neurons that remembers a pattern (1,1,-1).

$$\vec{\xi} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

$$M_{ij} = \frac{1}{N}\vec{\xi}\,\vec{\xi}^T = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

Is the pattern (1,1,-1) a stable state?

$$v(t+1) = \mathrm{sgn}\left[\frac{1}{3}\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}\right] = \mathrm{sgn}\left[\frac{1}{3}\begin{pmatrix} 3 \\ 3 \\ -3 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

- Does our network have an 'attractor' at the pattern (1,1,-1)?
  - Let's start the network at a different state and see what happens…

$$\vec{v}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$v(t+1) = \mathrm{sgn}\left[\frac{1}{3}\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right] = \mathrm{sgn}\left[\frac{1}{3}\begin{pmatrix} 3 \\ 3 \\ -3 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

In comparison with Discrete Hopfield network, continuous network has time as a continuous variable. It is also used in auto association and optimization problems such as travelling salesman problem.

Model − The model or architecture can be build up by adding electrical components such as amplifiers which can map the input voltage to the output voltage over a sigmoid activation function.

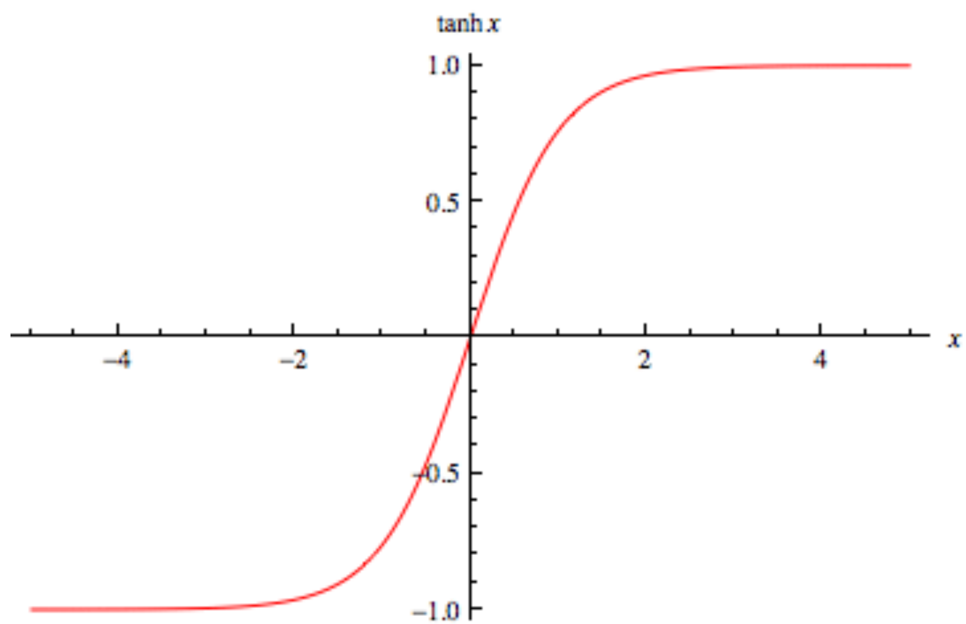
3. Continuous Hopfield Network

Consequently, $\mathbf{u}, \mathbf{v}, \mathbf{i}^{\mathrm{b}}$ will be the vectors of neuron states, outputs and biases of the network. The dynamics of the CHN is described by the differential equation system

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = -\frac{\mathbf{u}}{\tau} + \mathbf{T}\mathbf{v} + \mathbf{i}^{\mathrm{b}} \tag{1}$$

and the output function $v_i = g(u_i)$ is a hyperbolic tangent

$$g(u_i) = \frac{1}{2}\left(1 + \tanh\left(\frac{u_i}{u_0}\right)\right) \quad u_0 > 0.$$



3. Continuous Hopfield Network

The existence of equilibrium points for the CHN is guaranteed if a Lyapunovor Energy function exists. Hopfield showed that,if matrix T is symmetric, then the following Lyapunov function exists:

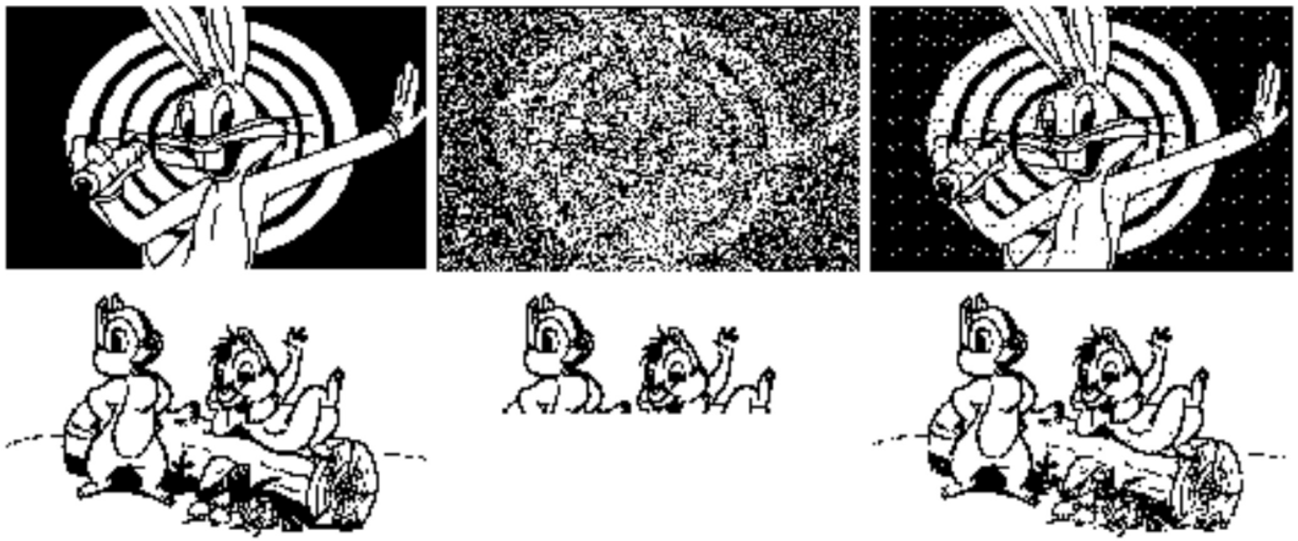
$$E = -\frac{1}{2}\mathbf{v}^{\mathrm{t}}\mathbf{T}\mathbf{v} - (\mathbf{i}^{\mathrm{b}})^{\mathrm{t}}\mathbf{v} + \frac{1}{\tau}\sum_{i=1}^{n}\int_{0}^{v_i} g^{-1}(x)\,\mathrm{d}x.$$

The CHN will solve those classification and optimization problems which can be expressed as the constrained minimization of

$$E = -\tfrac{1}{2}\mathbf{v}^{\mathrm{t}}\mathbf{T}\mathbf{v} - (\mathbf{i}^{\mathrm{b}})^{\mathrm{t}}\mathbf{v}, \tag{4}$$

In the examination, you are probably required to identify all equilibrium or equilibria of a given discrete or continuous Hopfield network.

Examples:



Original Degraded Reconstruction

Thank You !