B TECH –III yrs

Sem -5th 2020-21

NEURAL NETWORK
AND
FUZZY LOGICE

UNIT-1

Biological Inspirations

- Humans perform complex tasks like vision, motor control, or language understanding very well.
- One way to build intelligent machines is to try to imitate the (organizational principles of) human brain.

Human Brain

The brain is a highly complex, non-linear, and parallel computer, composed of some 1011 neurons that are densely connected (~104connection per neuron). *We have just begun to understand how the*

brain works…

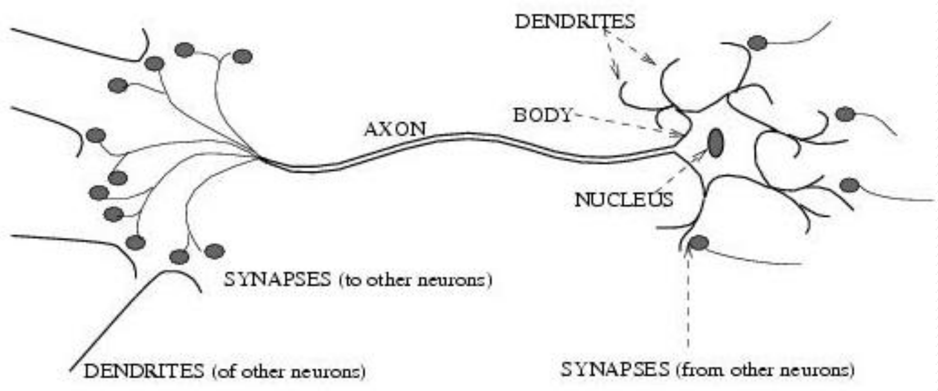
- A neuron is much slower (10-3sec) compared to a silicon logic gate(10-9sec), however the massive interconnection between neurons make up for the comparably slow rate.

– Complex perceptual decisions are arrived at quickly (within a few hundred milliseconds)

Human Brain

- 100-Steps rule: Since individual neurons operate in a few milliseconds, calculations do not involve more than about 100 serial steps and the information sent from one neuron to another is very small (a few bits)
- Plasticity: Some of the neural structure of the brain is present at birth, while other parts are developed through learning, especially in early stages of life, to adapt to the environment (new inputs).

Biological Neural Activity



- Neural activities passes from one neuron to another neuron in terms of electrical from body down to axon by means of electro chemical process
- Axon acts as connection wire

- Each neuron has a body, an axon, and many dendrites
  - Neuron's nucleus is where the neural computation takes place
  - Can be in 2 states: firing and rest
  - Neuron fires if incoming stimulus exceeds the threshold

- Synapse: Thin gap between axon of one neuron and dendrites of another one
  - Signal exchange

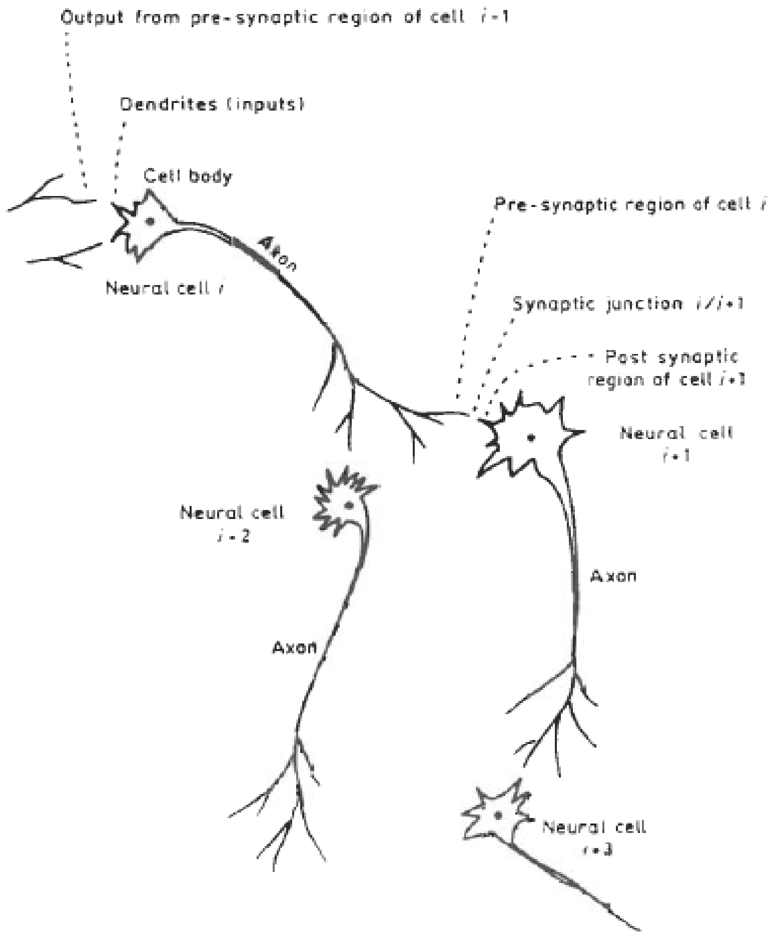
Biological Neuron

- A variety of different neurons exist (motor neuron, on center off-surround visual cells…), with different branching structures.
- The connections of the network and the strengths of the individual synapses establish the function of the network.

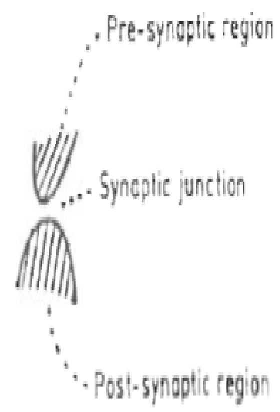
Biological Neuron

- Dendrites: nerve fibres carrying electrical signals to the cell
- Cell body: computes a non-linear function of its inputs
- Axon: single long fiber that carries the electrical signal from the cell body to other neurons
- Synapse: the point of contact between the axon of one cell and the dendrite of another, regulating a chemical Connection whose strength affects the input to the cell.

Biological Neural Activity



Interconnection of Biological Neural Nets



Synaptic Junction

Artificial Neural Networks

Computational models inspired by the human brain:

– Massively parallel, distributed system, made up of simple processing units (neurons)

– Synaptic connection strengths among neurons are used to store the acquired knowledge.

– Knowledge is acquired by the network from its environment through a learning process

Properties of ANNs

Learning from examples

– labeled or unlabeled

Adaptivity

– changing the connection strengths to learn things

Non-linearity

– the non-linear activation functions are essential

Fault tolerance

– if one of the neurons or connections is damaged, the whole network still works quite well

Thus, they might be better alternatives than classical solutions for

problems characterised by:

– high dimensionality, noisy, imprecise or imperfect data; and

– a lack of a clearly stated mathematical solution or algorithm

Neuron Model
and
Network Architectures

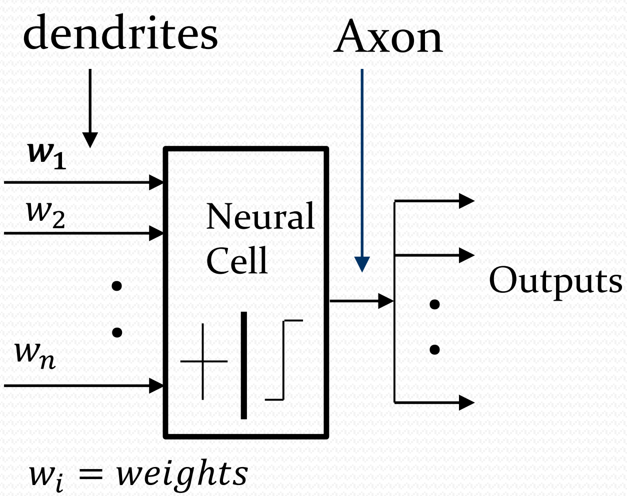
Artificial vs Biological NN

- **Artificial NN**
  - Nodes
    - Input
    - Output
    - Activation/Node function
  - Connections
    - Connection strength

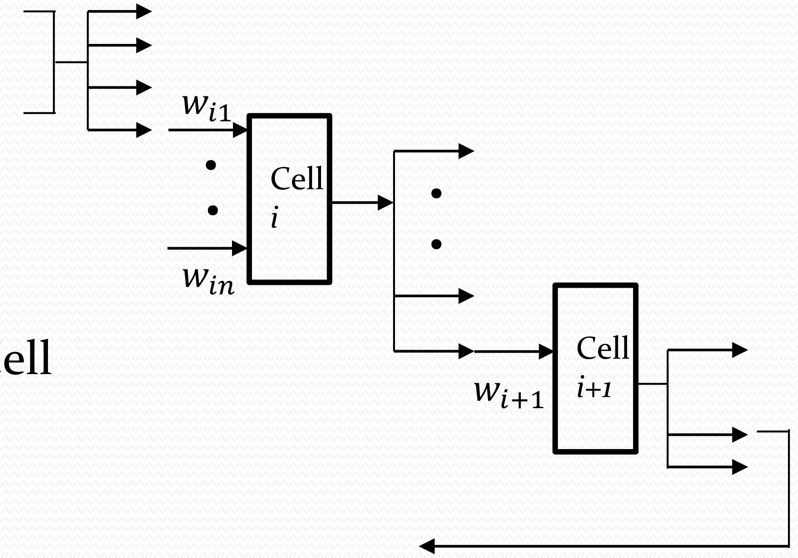
- **Biological NN**
  - Cell body
    - Signal from other neurons
    - Firing frequency
    - Firing mechanism
  - Synapses
    - Synaptic strength.

- Highly parallel, simple local computation at the neuron level, achieve global results as emerging property of the interaction at the network level
- Pattern directed (meaning of individual nodes only in the context of a pattern)
- Fault-tolerant/graceful degrading
- Learning/adaptation plays an important role

Analog Representations



Schematic Analog of a Biological Neural Cell



Schematic Analog of a Biological Neural Network

Artificial vs Biological NN

Characteristics	ANN	Biological NN
Speed	Fast	Slow
Processing	Exhibit some degrees of sequential instructions	Massive Parallel
Size and Complexity	Do not involve much computational neurons. Difficult to perform complex tasks	Large number of computing neurons ~ 10^{11} neurons 10^{15} connections which give the brain to perform complex tasks
Storage	Fixed size. New information may overwrite old memory storage	Information is stored in the strength of the interconnections without destroying old information
Fault Tolerance	Inherently not	Exhibit fault tolerance. Info is distributed in the connections
Control Mechanism	There is a control limit to monitor all activities of computing	No specific control mechanism external to the computing task

History of ANN

- **McCulloch and Pitts (1943)**
  - First mathematical model for biological neurons using simple binary functions. This model has been modified and widely applied in subsequent work
- **Landahl, McCulloch, and Pitts (1943)**
  - All Boolean operations can be implemented by McCulloch and Pitt model (with different threshold and excitatory/inhibitory connections)
- **Hebb (1949)**
  - Hebbian rule of learning: increase the connection strength between neurons *i* and *j* whenever both *i* and *j* are activated
  - Or increase the connection strength between nodes *i* and *j* whenever both nodes are simulated ON or OFF

History of ANN

- **Early booming (50's to early 60's)**
  - Rosenblatt (1958)
    - Perceptron: network of threshold nodes for pattern classification
    - Perceptron learning rule
    - Perceptron convergence theorem: Every thing that can be represented by a perceptron can be **learned**
  - Widrow and Hoff (1960, 1962)
    - Learning rule based on gradient descent
  - Minsky's attempt to build a general purpose machine with Pitts/McCulloch units

History of ANN

- **The setback (mid 60's to late 70's)**
  - Serious problems with perceptron model (Minsky's book 1969)
    - Single layer perceptrons cannot represent/learn a simple function such as XOR
    - Multi-layer of non-linear units may have greater power but there is no learning rule for such nets
    - Scaling problem: connection weights may grow infinitely
    - The first two problems overcame in 80's but the scaling problem remains
  - Death of Rosenblatt (1964)

History of ANN

- **Renewed enthusiasm and flourish (80's to present)**
  - New techniques
    - Backpropagation learning for multi-layer feed forward nets (with non-linear, differentiable node functions)
    - Thermodynamic models (Hopfield net, Boltzmann machine, etc.)
    - Unsupervised learning
  - Impressive applications: character recognition, speech recognition, text-to-speech transformation, process control, associative memory etc.
  - Traditional approaches face difficult challenges

- **Caution:**
  - Do not underestimate difficulties and limitations
  - Pose more challenges than solutions

Principles and Model of ANN

Principles of ANN

- **First formulated by McCulloch and Pitts (1943) by 5 assumptions**
  1. Activity of a neuron (ANN) is all-or-nothing
  2. # excited synapses > 1 within a given interval for a neuron to be excited
  3. The only significant delay within the neural system is the synaptic delay
  4. Inhibitor synapse prevents the excitation of a neuron
  5. The structure of the interconnection does not change over time

- By assumption 1, the neuron is a binary element

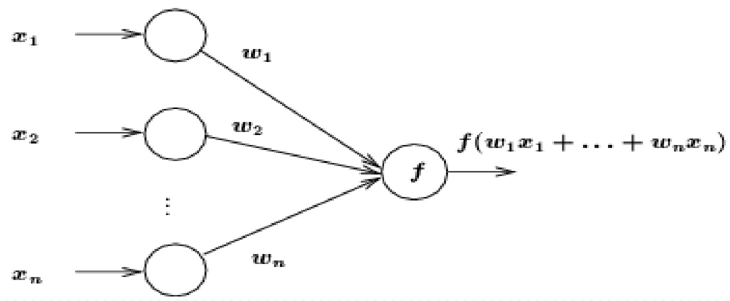
- **Hebbian rule** (1949)

- **Associative Memory (AM) Principle [Longuett, Higgins, 1968]**
  - Information vector (pattern, code) that is input into a group of neurons may (over repeated application of such input vectors) modify the weights at the input of that certain neuron in an array of neurons to which it is input, such that they more closely approximate the coded input

- **Winner Take All (WTA) Principle [Kohonen, 1984]**
  - If an array of N neurons receiving the same input vector, then only 1 neuron will fire. This neuron being the one whose weight best fit the given input vector
  - This principle saves firing from a multiple of neurons when only one can do the job

ANN Model

- Each node has one or more inputs from other nodes, and one output to other nodes
- Input/Output values can be
  - Binary {0, 1}
  - Bipolar {-1, 1}
  - Continuous (bounded or not)
- All inputs to one node come in at the same time and remain activated until the output is produced
- Weights associated with link
- $f(net)$ is the most popular function where $net = \sum_{i=1}^{n} w_i x_i$



Bias

An artificial neuron:

- computes the weighted sum of its input (called its **net input**)
- adds its bias

- passes this value through an activation function

- We say that the neuron "fires" (i.e. becomes active) if its output is above zero.

$$a_i = f(n_i) = [f(\Sigma w_{ij} x_j) + b_i]$$

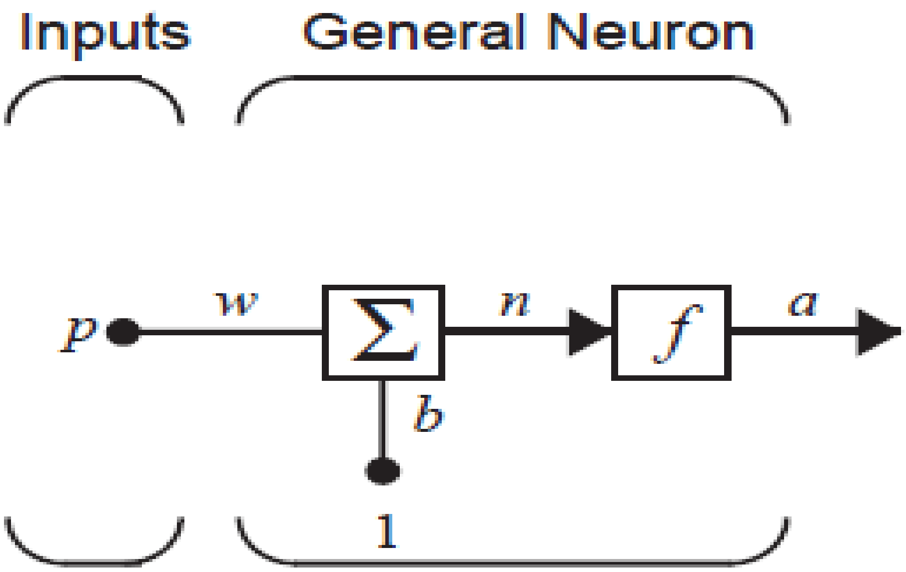
Bias

- Bias can be incorporated as another weight clamped to a fixed input of +1.0

- This extra free variable (bias) makes the neuron more powerful.

- $a_i = f(n_i) = f(\Sigma w_{ij} x_j) = f(\mathbf{w_i} . \mathbf{x_j})$

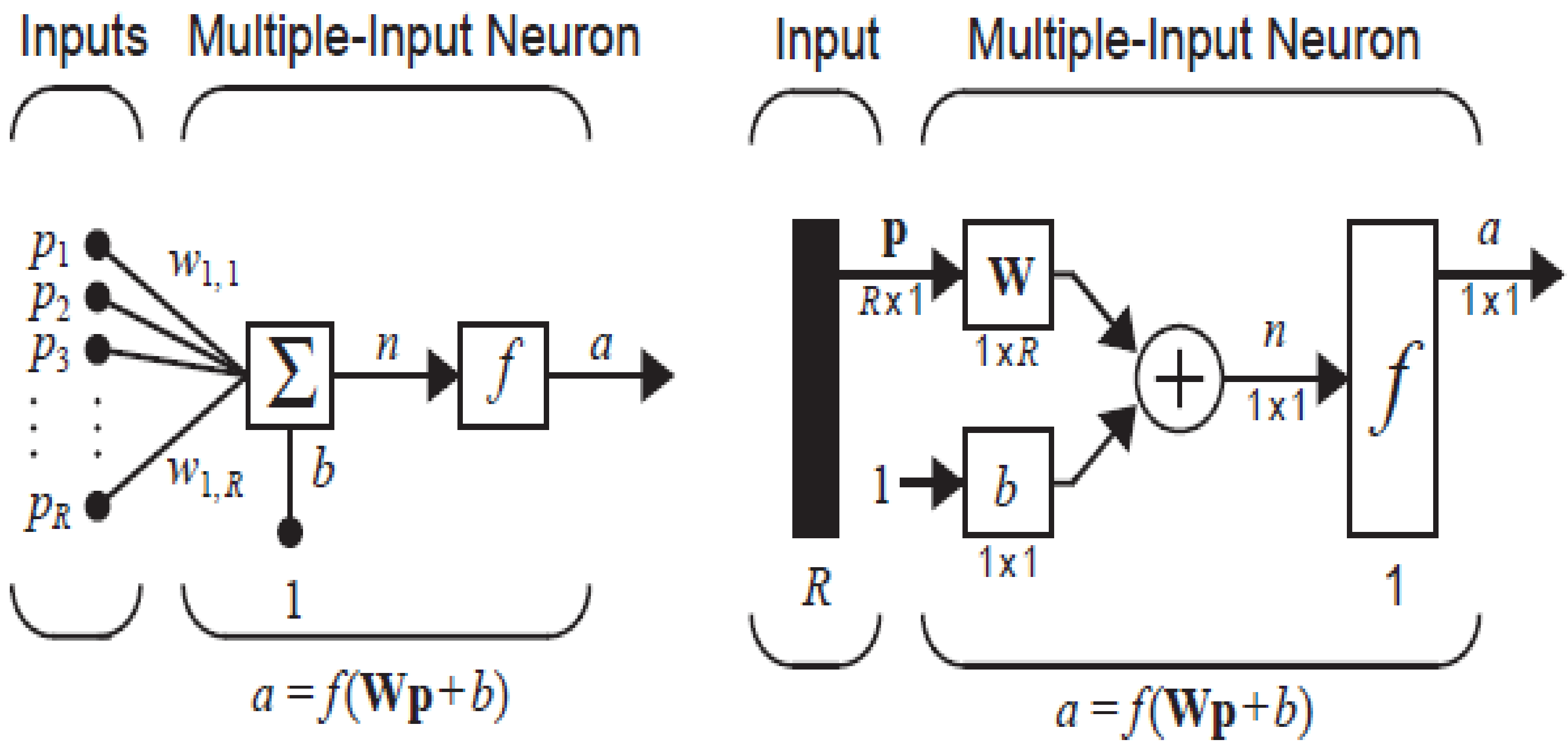
Single Input ANN Model



$$a = f(wp + b)$$

Single Input

Multiple Inputs ANN Model



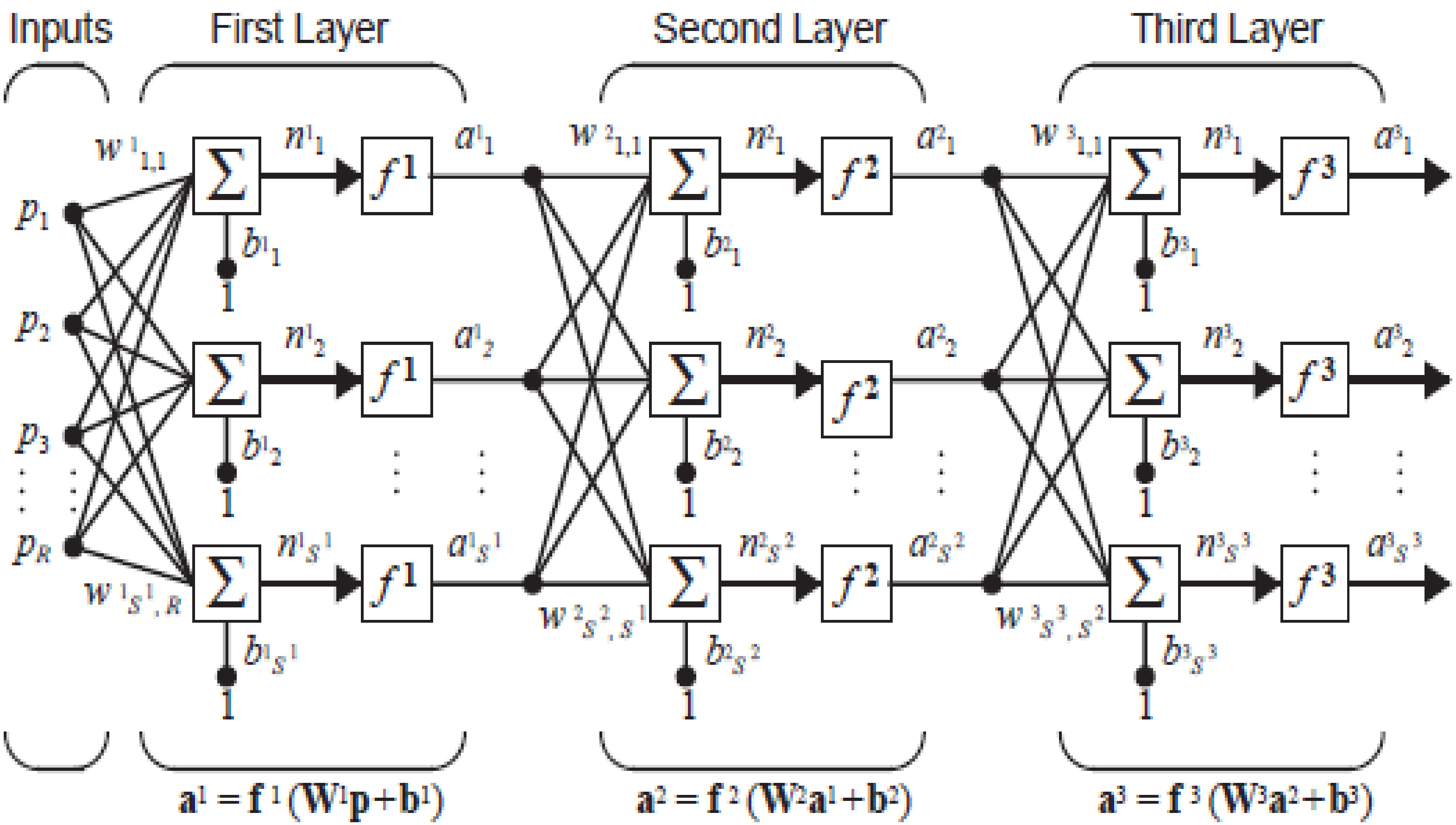
Multiple Inputs

Network Architectures

Multi-Layer Network Architectures

- **Four major multi-layer network architectures:**
  1. The Back-Propagation Network [Rumelhart et al, 1986]
     - A multi-layer Perceptron based ANN
     - Provide an elegant solution to hidden layers learning
  2. The Hopfield Network [Hopfield, 1982]
     - Feedback between neurons
     - Weight adjustment is based on Associate Memory
  3. The Counter Propagation Network [Hecht-Nielsen, 1987]
     - Self Organizing Mapping (SOM) is employed to facilitate unsupervised learning
     - Winner Take All (WTA) principle
  4. The LAMSTAR (Large Memory Storage and Retrieval) Network
     - A Hebbian network that use SOM layer and WTA principle

Multilayer Network Architecture

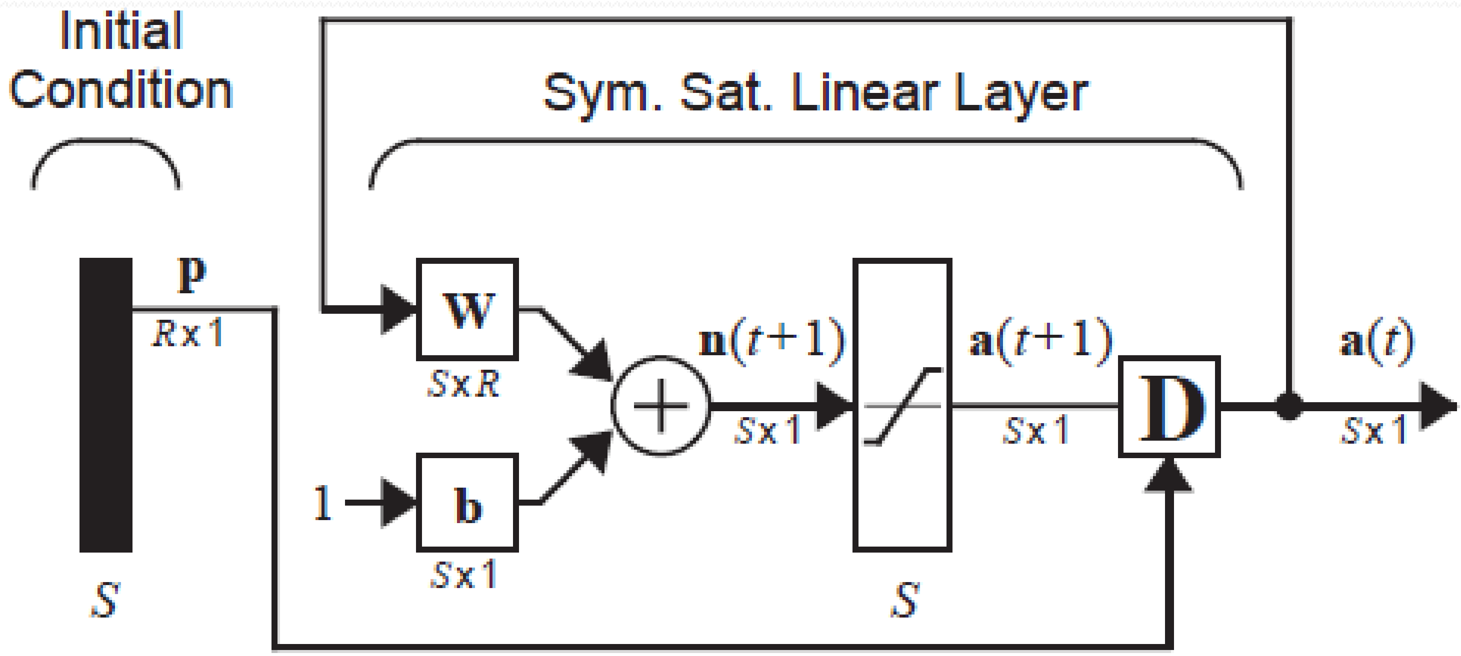


$$a^3 = f^3(W^3 f^2(W^2 f^1(W^1 p + b^1) + b^2) + b^3)$$

Classes

- **Acyclic Networks**
  - Connections do not form directed cycles
  - Multi-layered feedforward nets are acyclic
- **Recurrent Networks**
  - Nets with directed cycles (feedback)
  - Much harder to analyze than acyclic nets
  - Long short-term memory (LSTM)
- **Modular nets**
  - Consist of several modules, each of which is itself a neural net for a particular sub-problem
  - Sparse connections between modules

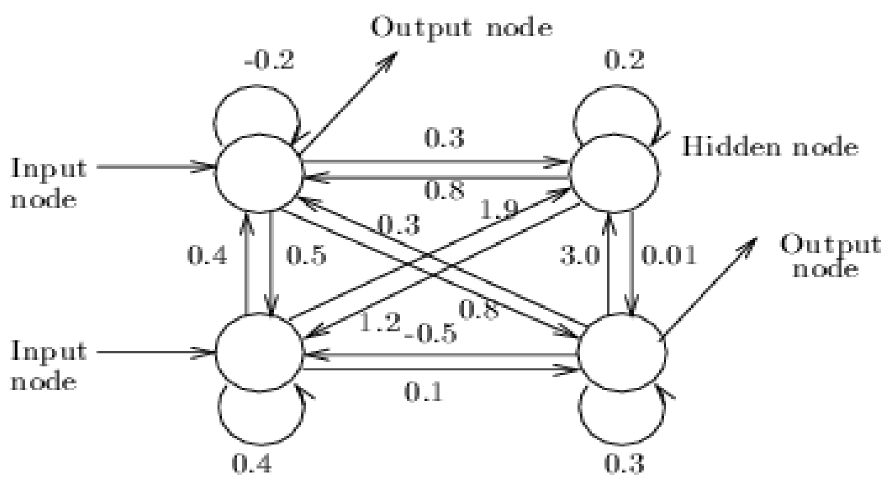
Recurrent Network Architecture



Fully Connected Network Architecture

Fully Connected Network

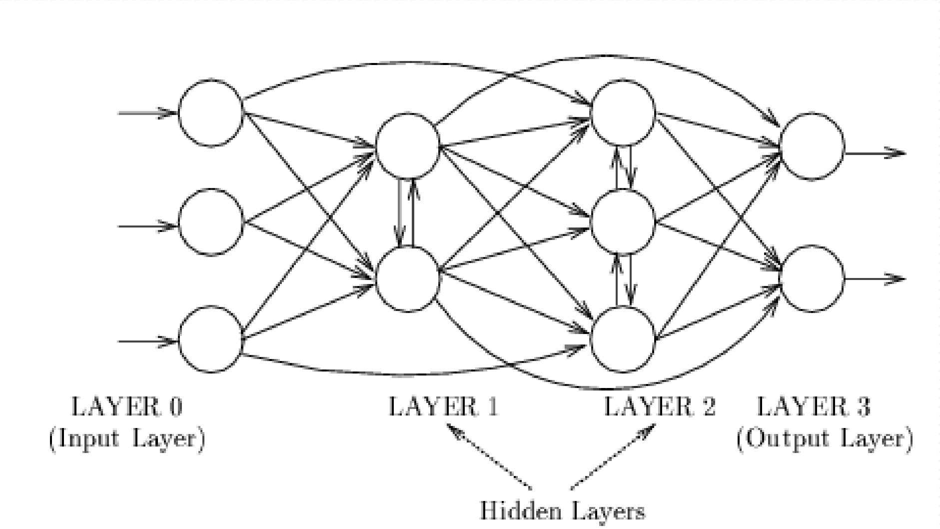
- Every node is connected to every other node
- Connection may be excitatory (positive), inhibitory (negative), or irrelevant (≈ 0)
- Most general
- Symmetric fully connected net: weights are symmetric $w_{ij} = w_{ji}$



- Input Nodes: Receive inputs from environment
- Output Nodes: Send signals to environment
- Hidden Nodes: No direct interaction to the environment

Layered Network Architecture

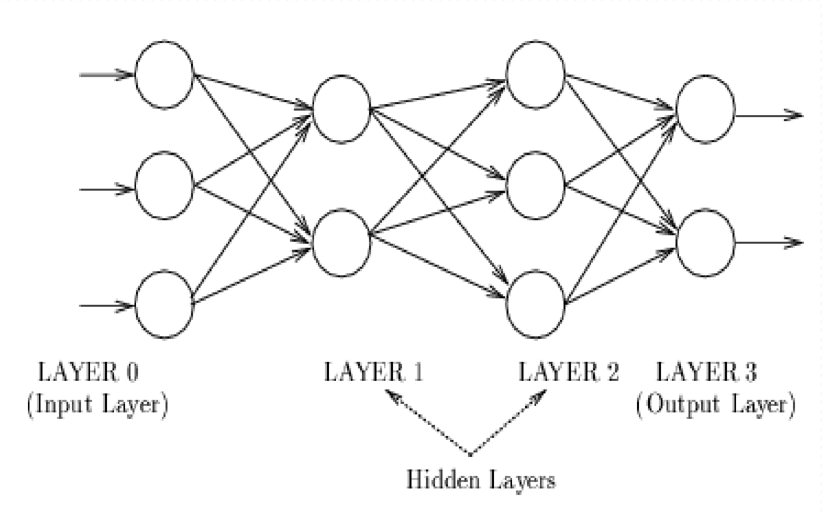
- **Layered Network**
  - Nodes that are partitioned into subsets, called layers
  - No connections that lead from nodes from layer j to layer k, if $j > k$ (no feedback)



- Input from the environment are applied to input layer (layer 0)
- Nodes in input layer are place holders with no computation occurring, i.e., their node functions are identity function

Feedforward Network Architecture

- **Feedforward Networks**
  - A connection is allowed from a node from layer i to nodes from layer $i + 1$ only
  - Most widely used



- Conceptually, nodes at higher levels successively abstract features from preceding layers