Lecture 3
Basic Definitions of ANN

Outline

- What is an Artificial Neural Network (ANN)?
- Basic Definitions:
  - Neuron
    - Net functions
    - Activation functions
  - ANN Configurations

What is an Artificial Neural Network?

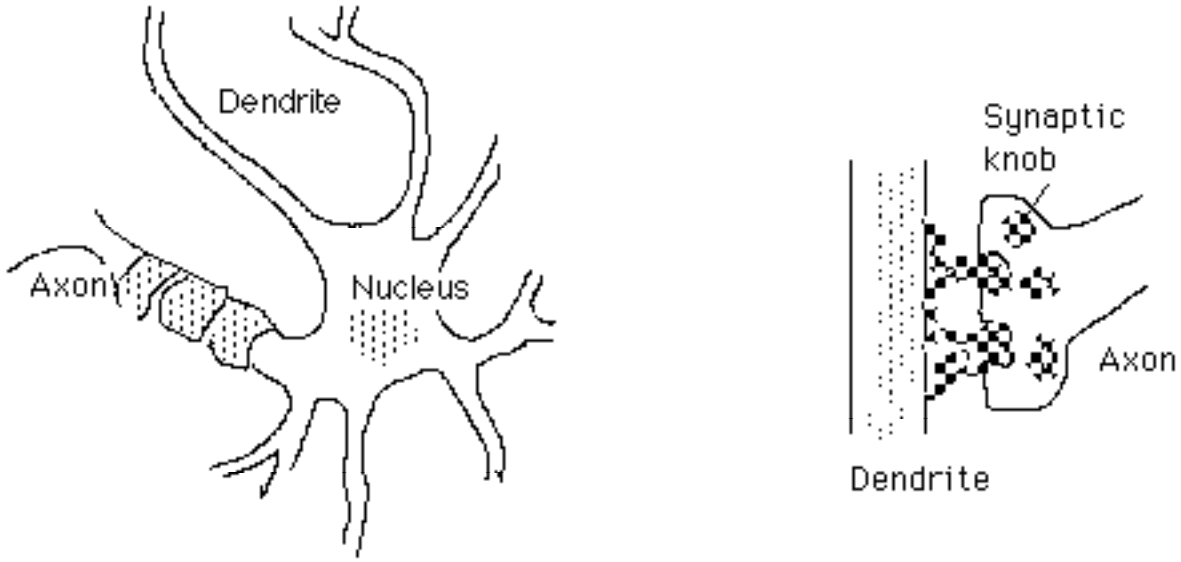
An artificial neural network (ANN) is a massively parallel distributed computing system (algorithm, device, or other) that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two aspects:

1). Knowledge is acquired by the network through a learning process.

2). Inter–neuron connection strengths known as synaptic weights are used to store the knowledge.

– Aleksander & Morton (1990), Haykin (1994)

Schematic of a Biological Neuron



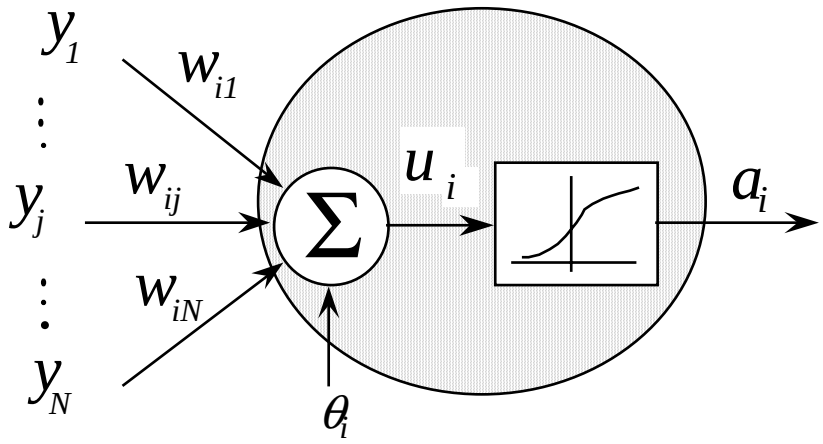
In a neuron cell, many dendrites accept information from other neurons, and a single axon output signals to other neurons. This is an extremely simplified model. However, it is still a popular and useful one.

WHY ANN?

- Has a potential to solve difficult problems current methods can not solve well (realistic reasons):
  - Pattern classification: hand-written characters, facial expression, engine diagnosis, etc.
  - Non-linear time series modeling, forecasting: Stock price, utility forecasting, ecg/eeg/emg, speech, etc.
  - Adaptive control, machine learning: robot arm, autonomous vehicle
- Requires massive parallel implementation with optical devices, analog ICs.
- Performance degrades gracefully when portions of the network are faulty.

NEURON MODEL

- McCulloch-Pitts (Simplistic) Neuron Model

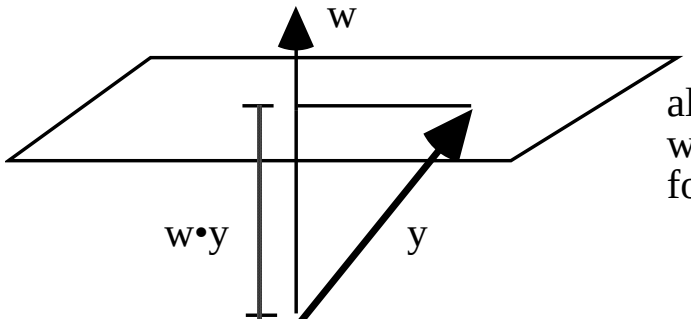


- The network function of a neuron is a weighted sum of its input signals plus a bias term.

Neuron Model

- The net function is a linear or nonlinear mapping from the input data space to an intermediate feature space (in terms of pattern classification).
- The most common form is a *hyper-plane*

$$u_i = \sum_{j=1}^{N} w_{ij} y_j + \theta_i = \mathbf{w}_i^T y + \theta_i = \begin{bmatrix} \theta_i & \mathbf{w}_i^T \end{bmatrix} \begin{bmatrix} 1 \\ y \end{bmatrix}$$



all the y such that
w•y = constant
form this plane

Side note: A Hyper Plane

Let $y = [y_1, y_2, \ldots, y_N]^T$ be a point (vector) in the N dimensional space, and $w = [w, w, \ldots, w_N]^T$ be another vector. Then, the *inner product* between these two vectors,

$$w^T y = c$$

defines a (N–1) dimensional ***hyperplane*** in the N-dimensional space. In 2-dimensional space, a hyperplane is a straight line that has the equation

$$w_1 y_1 + w_2 y_2 = c$$

In a 3D space, a hyperplane is just a plane.

A hyperplane partitions the space into two halves.

Other Net Functions

- Higher order net function: Net function is a linear combination of higher order polynomial terms. for example, a 2nd order net function has the form:

$$u_i = \sum_{j=1,k=1}^{N} w_{ijk} y_j y_k + \theta_i$$

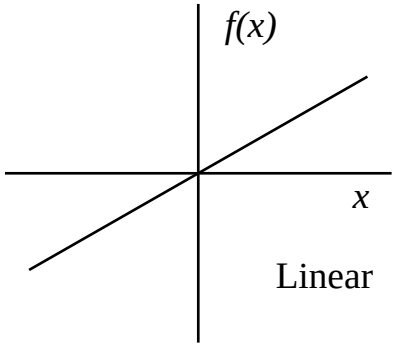
- Delta (S-P) net function – instead of summation, the product of all weighted synaptic inputs are computed:

$$u_i = \prod_{j=0}^{N} w_{ij} y_j$$

NEURON Activation Function

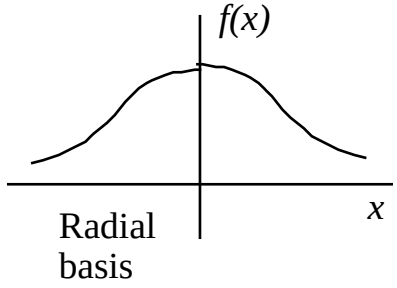
- Linear – $f(x) = a\,x + b$,

$$\frac{df(x)}{dx} = \alpha : \text{a constant}$$



- Radial basis – $f(x) = exp(-\alpha||x - x_o||^2)$;

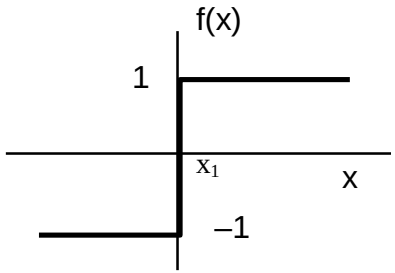
$$\frac{df(x)}{dx} = -2\alpha(x - x_o) \bullet f(x)$$



More Activation Function

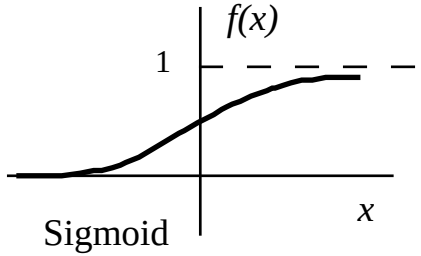
- Threshold –

$$f(x) = \begin{cases} 1 & x > x_1; \\ -1 & x < x_1. \end{cases}$$



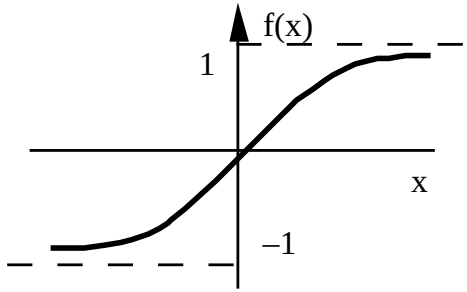
- Sigmoid – $f(x) = \dfrac{1}{1+\exp(-x/T)}$; T: temperature

$$\frac{df(x)}{dx} = \frac{1}{T} f(x)[1 - f(x)]$$



More Activation Function

- Hyperbolic tangent – $f(x) = \tanh(\frac{x}{T})$



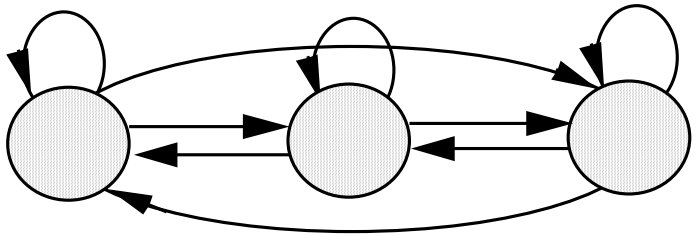
$$\frac{df(x)}{dx} = T[1 - f^2(x)]$$

- Inverse tangent – $f(x) = \frac{2}{\pi}\tan^{-1}\left(\frac{x}{T}\right)$

$$\frac{df(x)}{dx} = \frac{2}{\pi T}\frac{1}{1+(x/T)^2}$$

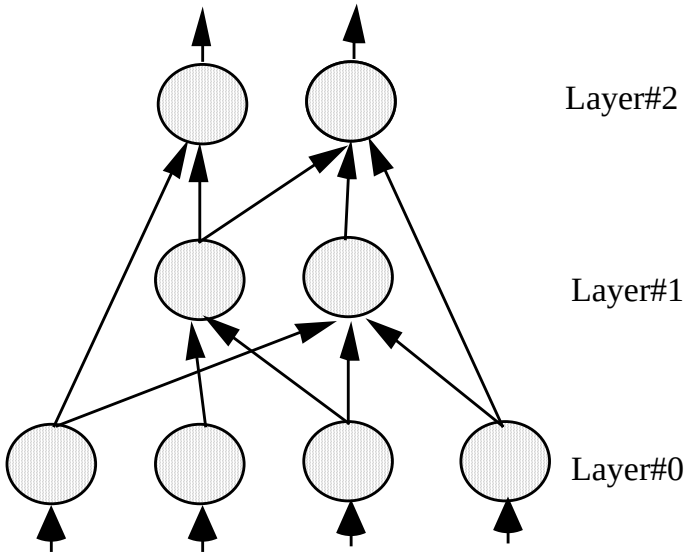
ANN CONFIGURATION

- Uni-directional communication links represented by directed arcs. The ANN structure thus can be described by a directed graph.
- Fully connected – a cyclic graph with feed-back. There are N^2 connections for N neurons.

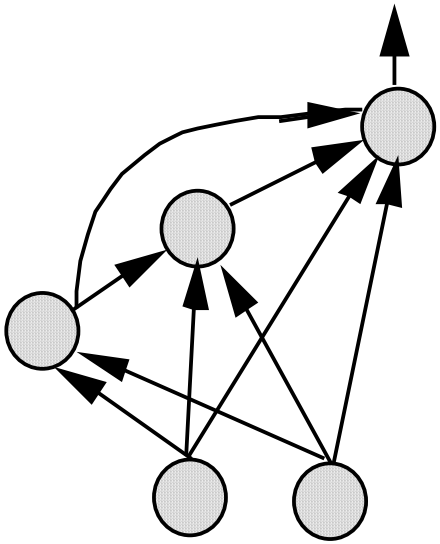


ANN CONFIGURATION

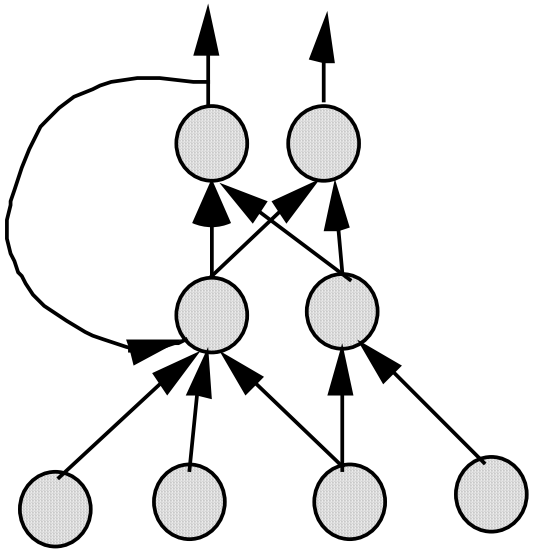
- Feed-forward, layered connection – acyclic directed graph, no loop or cycle.



ANN Configuration



A net with unidirection lateral connection



A net with feedback

Feed-back Dynamic System

- Without Delay, feedback cause causality problem: an unknown variable depends on an unknown variable!

  $a_2 = g(a_1) = g(g(a_2)) = \ldots$
- To break the cycle, at least one delay element must be inserted into the feedback loop.
- This effectively created a nonlinear dynamic system (sequential machine).