CSE 5526: Introduction to Neural Networks

Instructor: Michael Mandel

What is this course about?

- Artificial intelligence in the broad sense
  - in particular, learning
- Neurally-inspired algorithms that learn from data
- The human brain and its amazing abilities
  - e.g. vision, hearing, movement

Course Aims

- Understand the structure and uses of important neural network models
- Understand and implement their learning algorithms
- Use neural networks to solve learning problems
- Build foundation for taking other specialty courses on neural networks and related techniques

What is a neural network?

- Neural network: a collection of (biological) neurons that communicate with one another
- Artificial neuron: a model of the way that biological neurons communicate and compute
  - Range from very crude (McCulloch-Pitts)
  - To very sophisticated (Hodgkin-Huxley)
- Artificial neural network: a learning machine composed of artificial neurons
  - We will mainly use them as generic learners with crude neural models

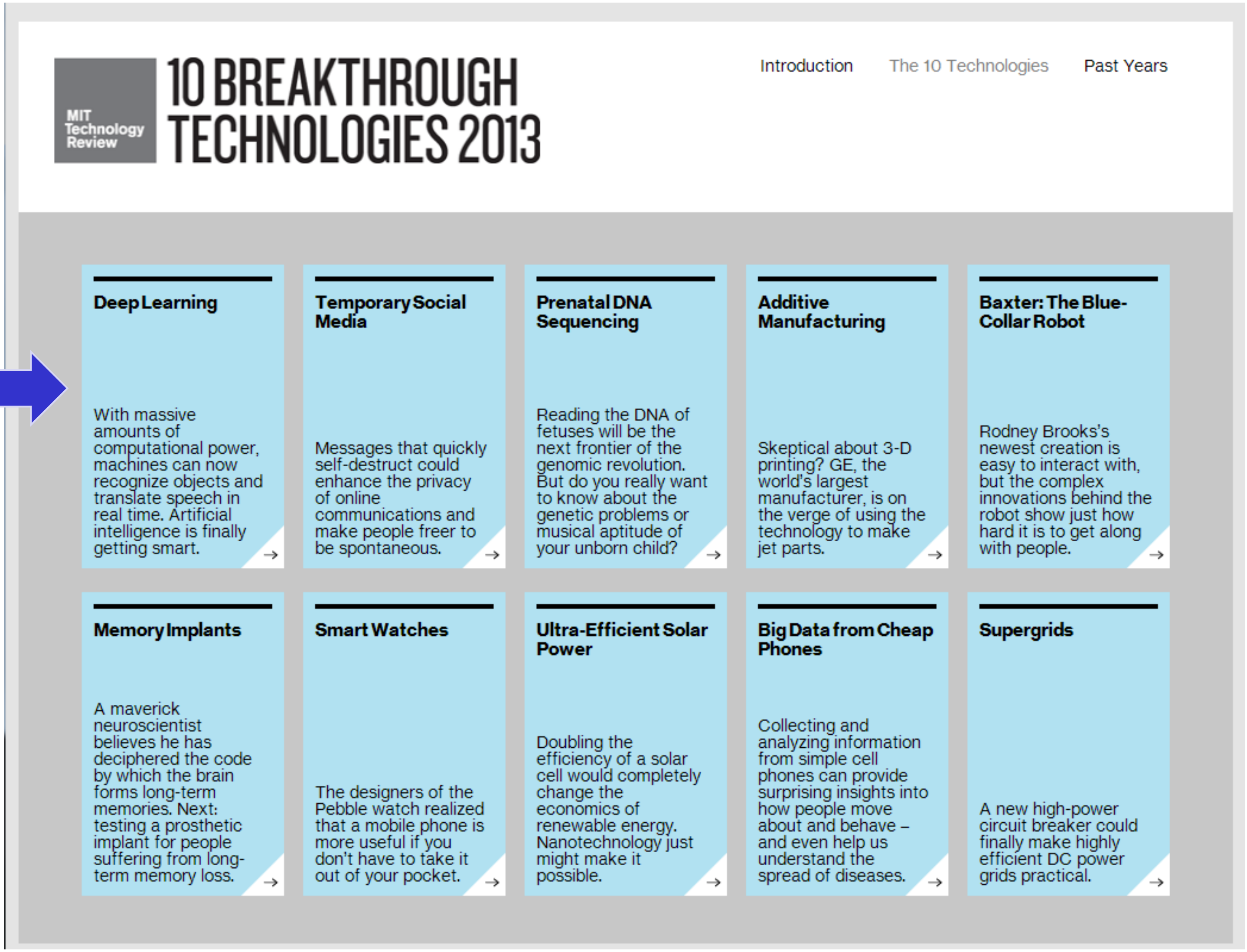
Why are artificial neural networks useful?

- They can approximate any function
- They generalize well to new (similar) data
- They learn from examples (less hand-coding)
- They can learn from lots of data (efficient training)
- Once trained, they are efficient to evaluate
  - And their parameters don’t take up much memory
- They are well suited to (massive) parallelization

How are ANNs used in practice?

- Visual object recognition
- Automatic speech recognition
- Handwritten digit recognition
- Language models for machine translation
- Time series modeling and prediction
- Document classification and retrieval
- Compression and denoising

Recent news…


MIT Technology Review
10 BREAKTHROUGH TECHNOLOGIES 2013
Introduction
The 10 Technologies
Past Years
Deep Learning
With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.
Temporary Social Media
Messages that quickly self-destruct could enhance the privacy of online communications and make people freer to be spontaneous.
Prenatal DNA Sequencing
Reading the DNA of fetuses will be the next frontier of the genomic revolution. But do you really want to know about the genetic problems or musical aptitude of your unborn child?
Additive Manufacturing
Skeptical about 3-D printing? GE, the world's largest manufacturer, is on the verge of using the technology to make jet parts.
Baxter: The Blue-Collar Robot
Rodney Brooks's newest creation is easy to interact with, but the complex innovations behind the robot show just how hard it is to get along with people.
Memory Implants
A maverick neuroscientist believes he has deciphered the code by which the brain forms long-term memories. Next: testing a prosthetic implant for people suffering from long-term memory loss.
Smart Watches
The designers of the Pebble watch realized that a mobile phone is more useful if you don't have to take it out of your pocket.
Ultra-Efficient Solar Power
Doubling the efficiency of a solar cell would completely change the economics of renewable energy. Nanotechnology just might make it possible.
Big Data from Cheap Phones
Collecting and analyzing information from simple cell phones can provide surprising insights into how people move about and behave – and even help us understand the spread of diseases.
Supergrids
A new high-power circuit breaker could finally make highly efficient DC power grids practical.

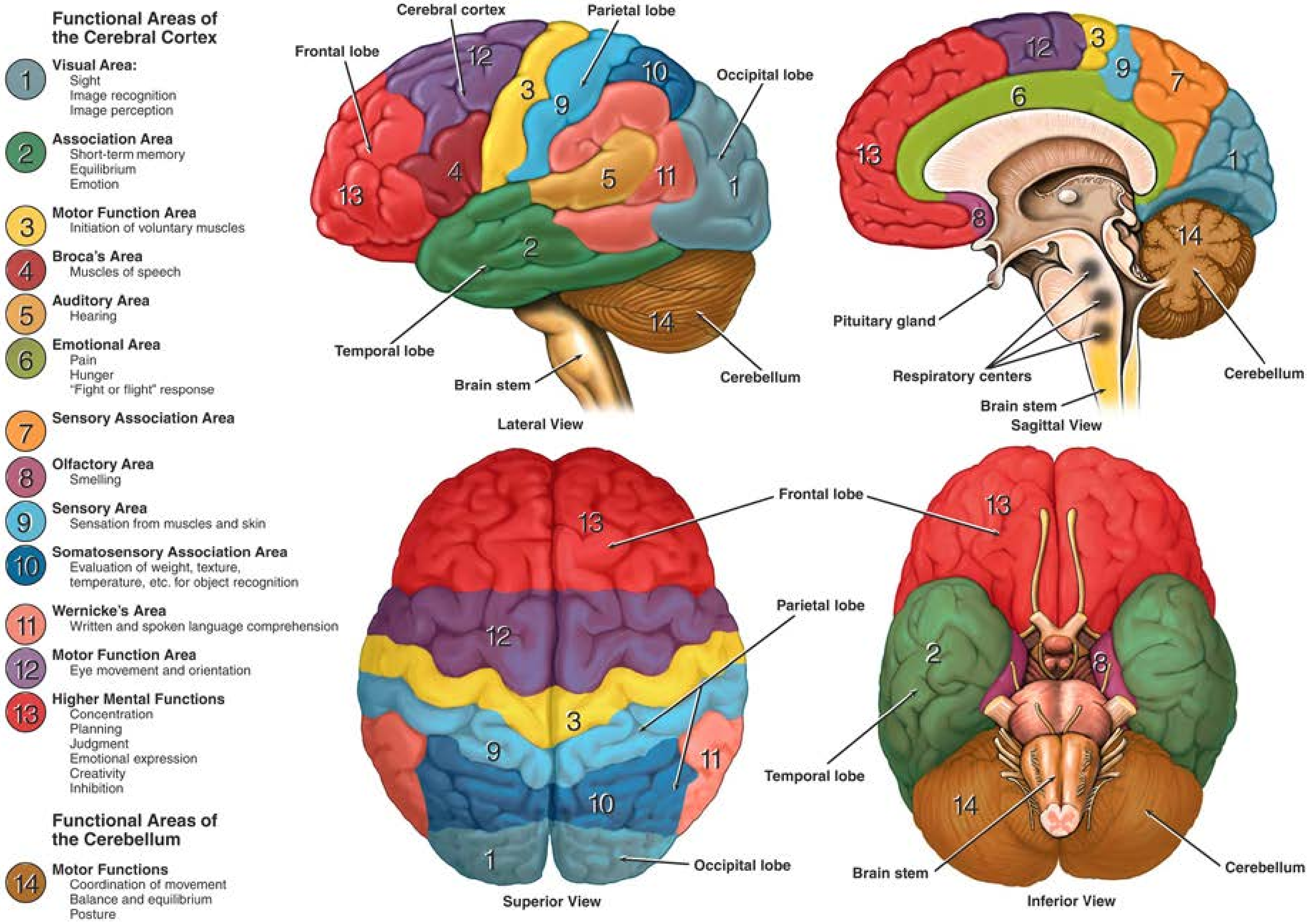
Syllabus

- Intro, McCulloch-Pitts networks, perceptrons
- Regression & least mean square algorithm
- Multilayer perceptrons & backpropagation
- Radial basis function networks
- Support vector machines
- Unsupervised learning and self-organization
- Hopfield networks
- Stochastic methods & Boltzmann machines
- Deep neural networks

Instructor: Dr Michael Mandel

- Research scientist in CSE since 2012
- Office hours: Mon&Thurs 11-12 in Dreese 258
- Studies “machine listening”
  - Noise robust speech recognition, noise suppression
  - Music classification and source separation
- Has used neural networks in many research projects
  - Automatically describing music
  - Recognizing and isolating speech in noise
- Postdoc in deep learning at University of Montreal

Human brain



Credit: Nucleus Medical Art, Inc./Getty Images

http://www.dana.org/News/Details.aspx?id=43515

Brain versus computer

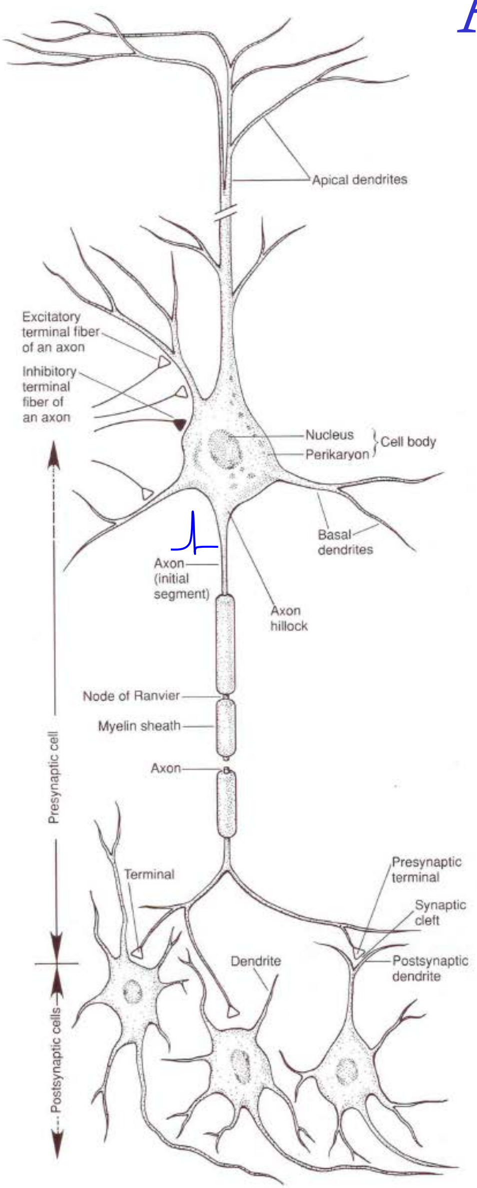
Brain

-
-
-
-
-

Computer

-
-
-
-
-

A single neuron

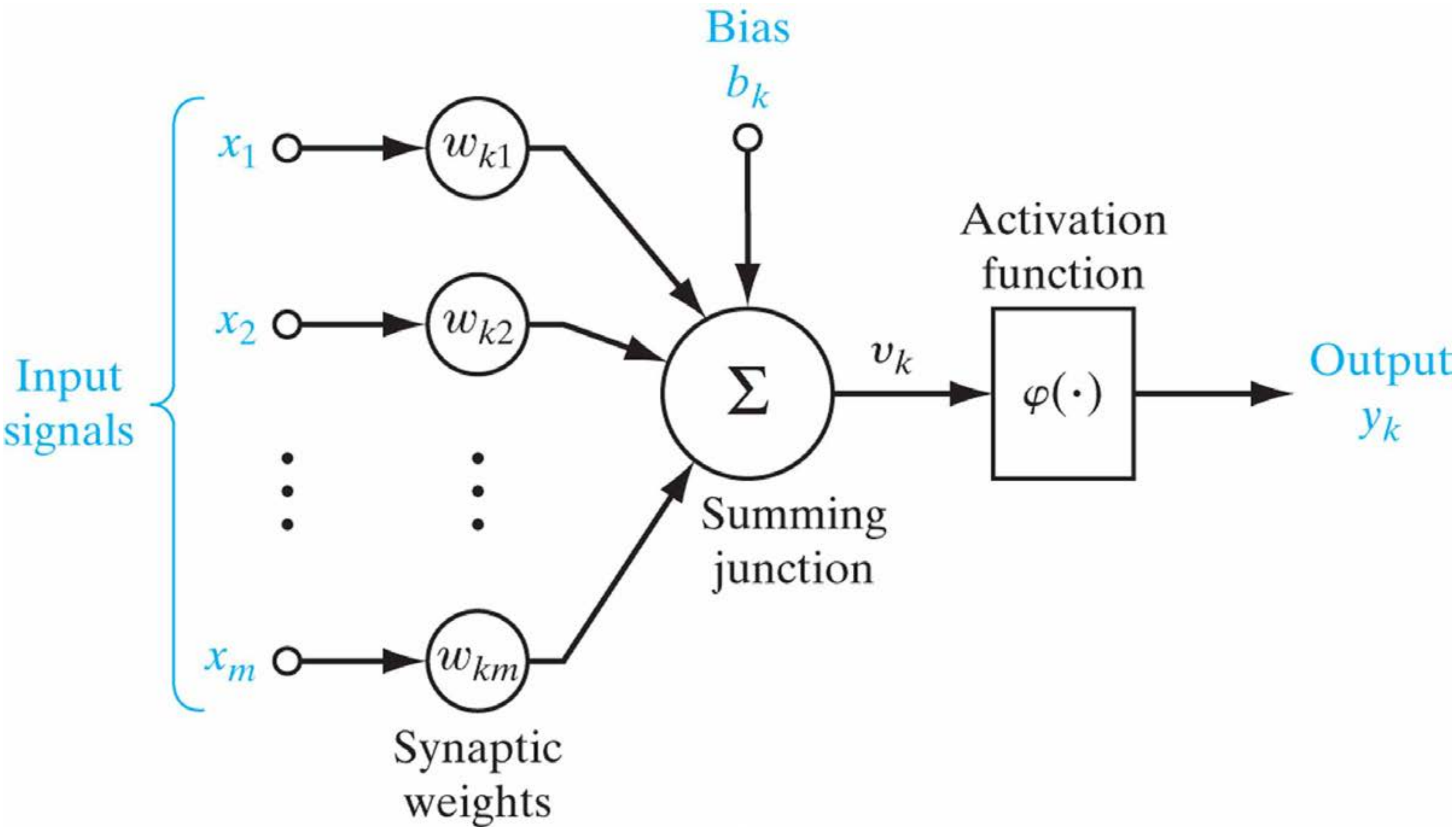


- Cell body (~50μm)
  - Initiates action potential
- Axon (0.2-20μm)
  - Transmits signal to up to 1000 other neurons
  - Insulated by myelin sheath
  - Up to 1m long
- Synapse: junction btw neurons
- Dendrites: receive signals

Real neurons, real synapses

- Properties
  - Action potential (impulse) generation
  - Impulse propagation
  - Synaptic transmission & plasticity
  - Spatial summation
- Terminology
  - Neurons – units – nodes
  - Synapses – connections – architecture
  - Synaptic weight – connection strength (either positive or negative)

Model of a single neuron



Neuronal model

$$u_k = \sum_{j=1}^{m} w_{kj} x_j$$

Adder, weighted sum, linear combiner

$$v_k = u_k + b_k$$

Activation potential; b_k: bias

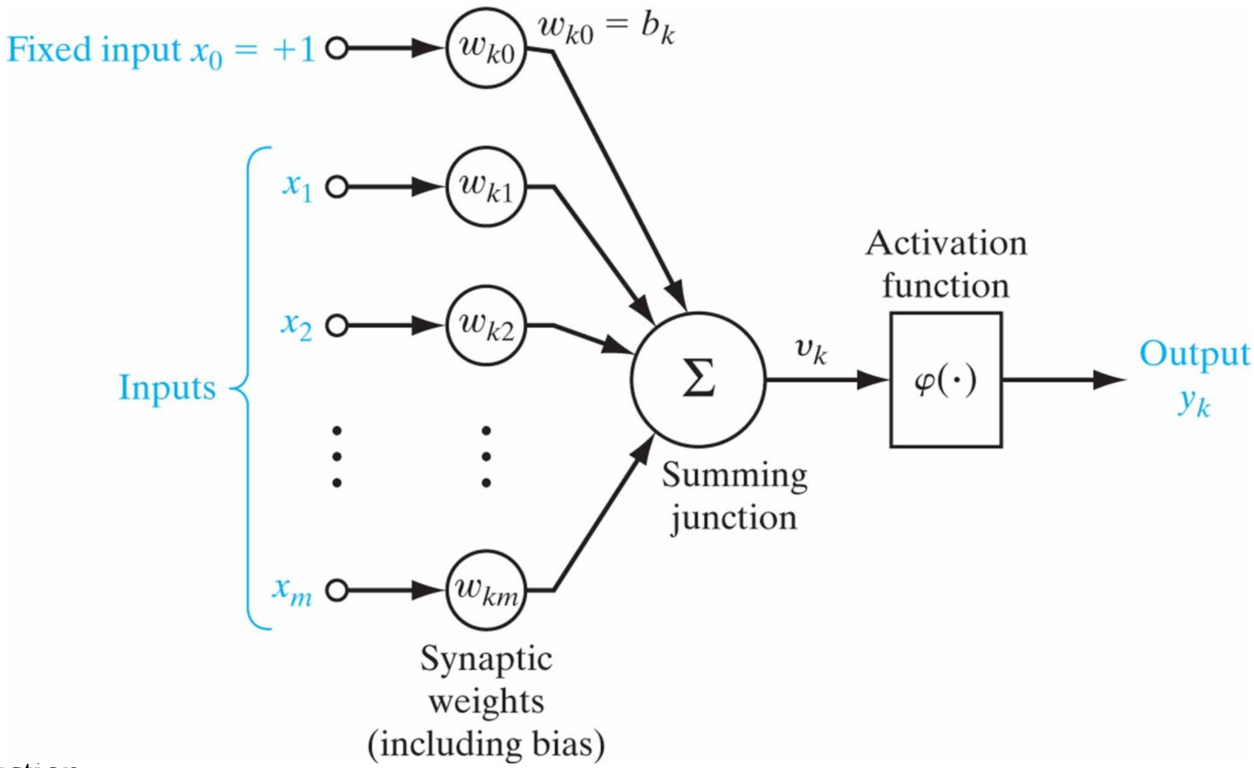
$$y_k = \varphi(v_k)$$

Output; φ: activation function

Another way of including bias

Set $x_0 = +1$ and $w_{k0} = b_k$

So we have $v_k = \sum_{j=0}^{m} w_{kj} x_j$



McCulloch-Pitts model

$x_i \in \{-1, 1\}$ Bipolar input

$$y = \varphi(\sum_{i=1}^{m} w_i x_i + b)$$

$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ -1 & \text{if } v < 0 \end{cases}$$

A form of signum (sign) function

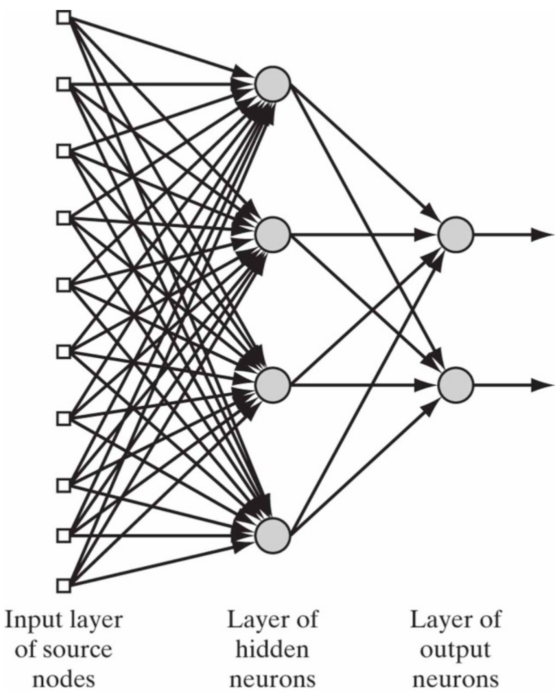
McCulloch-Pitts model (cont.)

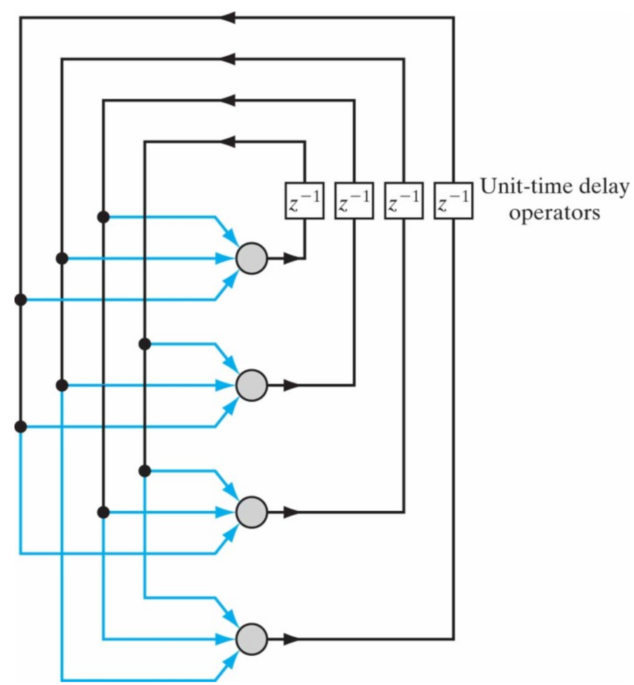
- Example logic gates (see blackboard)

- McCulloch-Pitts networks (introduced in 1943)
  - Can implement any finite state machine
  - Can compute any logic (Boolean) function
  - Can recognize any *regular* sequence

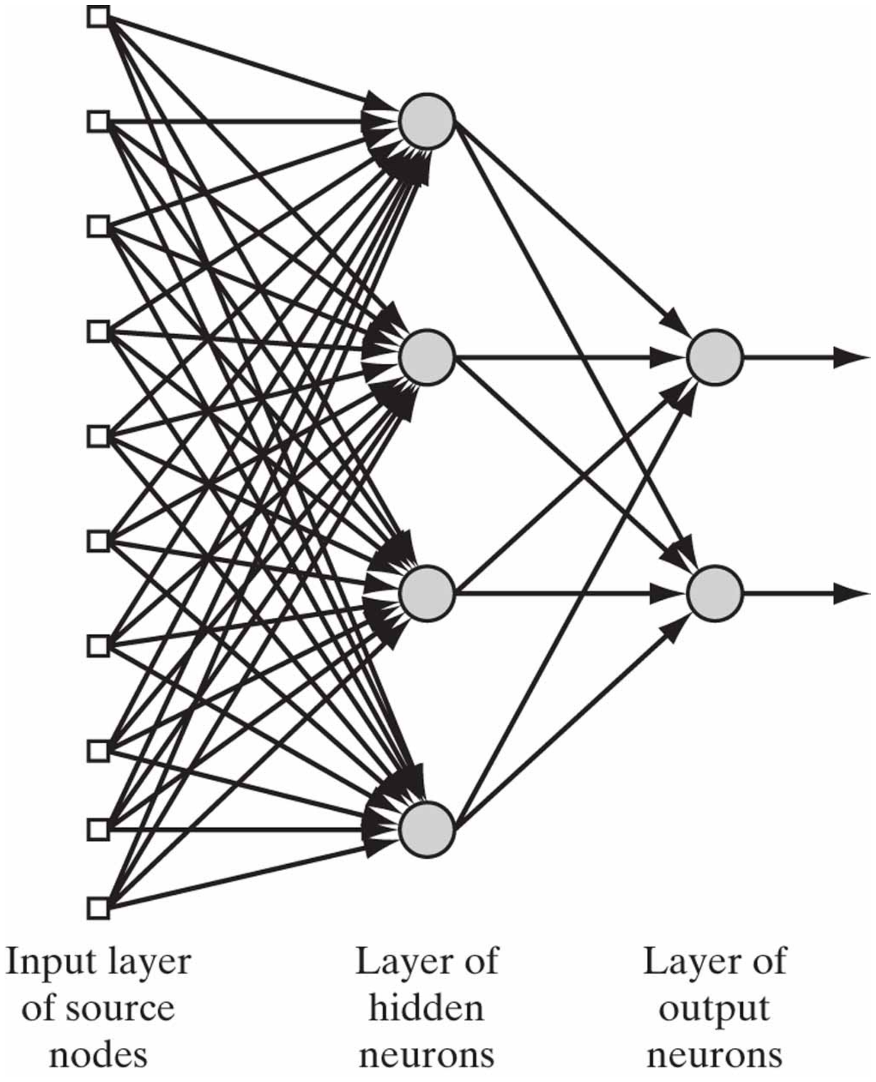
Network architecture

- View an NN as a directed graph (its architecture)
  - Feedforward nets: loop-free graph
  - Recurrent nets: loopy graph



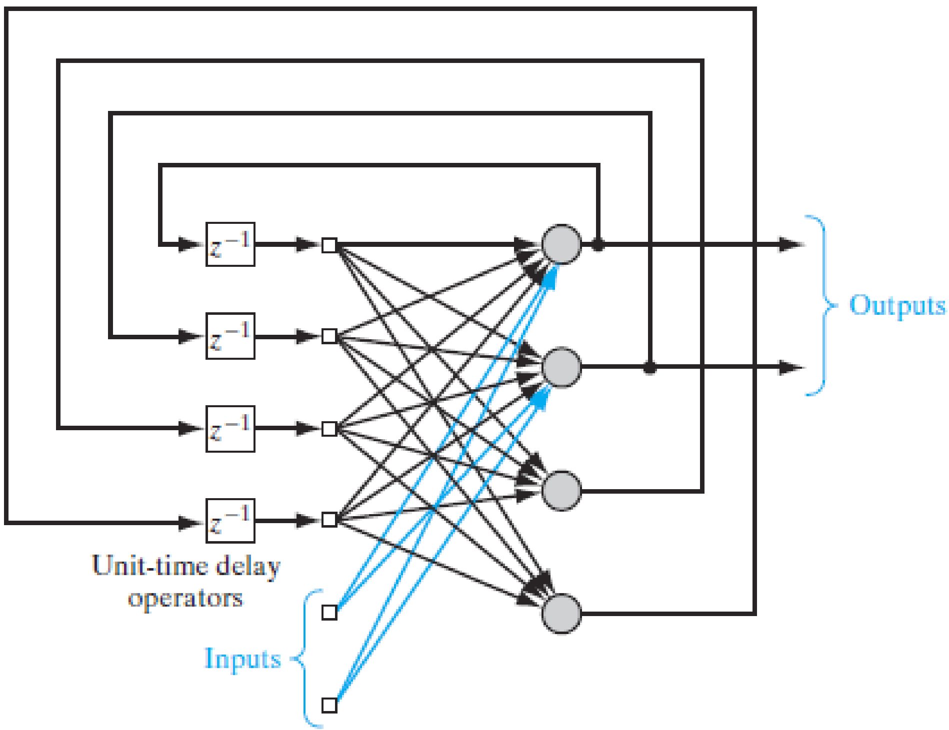


Feedforward net



- Describe architecture by number of nodes in each layer: e.g., 10-4-2
- And by number of layers (not counting input)
- So 10-4-2 counts as a two-layer net

An example recurrent net



- Example architecture
  - 2 Input units
  - 4 Hidden units
  - 2 Outputs
  - Feedback from previous hidden unit values

Sufficient for most purposes

Difficult to train

Network components

- Three components characterize a neural net
  - Architecture
  - Activation function
  - Learning rule (algorithm)