Harvey Mudd College

CS 152
Neural Networks
Fall 2000

Bob Keller, Professor
keller@cs.hmc.edu
621-8483
Office Hours (1249 Olin):

- Note: 1249 is in the SW corner of Olin
- Monday, Tuesday, Wednesday 4-5
- By drop-in (as available)
- By appointment:
  - email keller@cs.hmc.edu
  - phone 621-8483
- Crisis center: 621-2373

Course Outline

Please Refer to Web Page for details:

http://www.cs.hmc.edu/courses/current/cs152
Neural Networks: an Eclectic Discipline

**Biology**
Neurophysiology

**Psychology**
Learning & Cognition

**Engineering**
Pattern Classification, Control, Robotics

**Computer Science**
Artificial Intelligence

**Physics**
Dynamical Systems

**Mathematics**
Optimization

**Economics**
Forecasting
Time Series

**Statistics**
Regression Analysis
Classification
Biological Intelligence

- Intelligence, the ability to make decisions based upon input from the environment.

- Intelligence is realized by a network of neurons, for example the brain and the attendant sensory and motor neurons.
“Neurons R US”

- Not only our intelligence, but all aspects of our behavior, are the result of neural activity:
  - emotions
  - memory
  - reflexes
  - habits
  - likes and dislikes
  - addictions
Approaches to Artificial Intelligence

- Reverse Engineering of Biology
  - Understand real neurons well enough to model
  - *Simulate* neural behavior

- Simulated Evolution
  - Provide basic evolutionary mechanism for neurons
  - *Evolve* intelligent behavior

- Artificial Neural Networks
  - Develop a parameterized model for a class of problems
  - *Learn* the parameters
Fundamental Problems for a Given Neural Model

- How to represent information?
- How to characterize the computational capability of the model?
- How to achieve learning in the model?
Some Applications of Artificial Neural Networks (1 of 5)

- **Optical character recognition**
  - U.S. mail zip-code recognizer
  - Kanjii: 4000 chars in 15 fonts, 99% accurate, 100k chars/sec (Sharp & Mitsubishi)

- **Communications**
  - Adaptive noise cancellation
    - Headphones
    - Conference telephones
Applications (2 of 5)

- **Process control**
  - Electric arc furnace control: 30MVA, 50kamp transformer, $2M savings
  - Steel-rolling mill controller
  - Copier uniformity control (Ricoh)
  - Anti-lock brakes, etc. (Ford)
  - Food process control (M&M)
  - Particle beam focusing (SLAC)
  - Fluorescent bulb mfg. (GE)
Applications (3 of 5)

- **Financial analysis**
  - Prediction of commodities market (18% vs. 12.3% by traditional methods)
  - Mortgage risk evaluator (AVCO, Irvine)
  - Real-estate evaluation (Foster Onsley Conley)
  - Portfolio management (LBS Captial)
  - Currency trading (Citibank)

- **Crime prevention**
  - Bomb sniffer (JFK airport)
  - Credit card fraud detection (Visa, etc.)
Applications (4 of 5)

- Object classification
  - Grading grains from video images
  - Forensics: glass classification
  - High-energy physics: particle identification
- Warfare
  - Missile guidance
- Optical telescope focusing
Applications (5 of 5)

- Biomedical
  - Clinical diagnoses
  - Patient mortality predictions
  - Protein structure analysis
  - Electrode placement
- Speech recognition
- Game playing
  - World backgammon champion
Some Physiological Aspects of Neurons
Reference for Biological Aspects

The Neuron
Cell and Molecular Biology
THIRD EDITION

Irwin B. Levitan
Leonard K. Kaczmarek

Oxford University Press.
Neuron Cell
(top half)
Neuron Cell (bottom half)

Axon can be a meter or more long (e.g. spine-to-toe).

Information flow

Synaptic terminals (To dendrites of other neurons.)
Photomicrograph of one neural cell (from cerebral cortex)
Structure of one neuron

Photomicrograph of network of neural cells (from the hippocampus region of the brain)
Composition of the Brain

- 10% neurons
- 90% glial ("glue") cells
Myelin sheath around axon (consists of glial cells)

Myelin Sheath (cont’d)

- Acts as insulator

- Current can flow out only at junctions (called nodes of Ranvier) to other axons

- Demyleinating diseases:
  - Myelin deficit in newborns
  - MS (multiple sclerosis)
  - ALS (amyotrophic lateral sclerosis, “Lou Gehrig’s disease”)

---

Demyleinating diseases:
- ALS (amyotrophic lateral sclerosis, “Lou Gehrig’s disease”)
- MS (multiple sclerosis)
- Myelin deficit in newborns

---

Myelin Sheath (cont’d)
Dendrite Information Flow

- Normally dendrites receive information from synapses of other neurons.

- In some cells, both input and output can occur through the same set of dendritic structures.
In addition to signal, axon carries:

- Construction material (proteins)
- Nutrients (in the form of mitochondria)
- Enzymes
Experiment determining chemical nature of transmission, Loewi, 1921

slow-down due to direct stimulation

slow-down due to chemical change in solution

An ionic (electro-chemical) reaction carries the signal across the gap between a synapse of one neuron and a dendrite of the next.

The strength of this connection is an abstraction of the efficiency of the transfer.

In artificial neural networks, this strength is represented by a numeric weight.
Chemical Synapse

Electronmicrograph of one synapse/dendrite connection

Schematic sometimes used (symbolic of synaptic clefts)

individual neurons

information flow
Ionic Neurotransmitter Reaction

- Action potential at synapses
- Ca ions enter cell
- Vesicles move to membrane, fuse, spill contents into cleft
- Transmitter crosses cleft, causes voltage change on post-synaptic side

Terminology

- **neurotransmitters**: molecules that traverse from synapse to dendrite through ion diffusion.

- **spiking**: abrupt change of output voltage

- **depolarization**: change in net input voltage toward a threshold value, at which it will “spike”

- **action potential**: the voltage change produced when the neuron spikes
C. If a stimulus arrives during this phase of the potential it can not stimulate the neuron.
D. Only a very strong stimulus can stimulate the neuron during this phase.
F. Absolute refractory period.
G. Relative refractory period.
All-or-None Behavior

- The action potential is essentially binary-valued.
- The strength of the stimulus does not matter, except insofar as whether it is over or above a threshold.
Triggering phenomenon

Stimulus (summed inputs)

Response

Refractory Period

- A neuron cannot spike again while it is already spiking, or for some time afterward.
- The “waiting time” until the next spike can occur is called the refractory period.
Intensity

- Because of the all-or-none behavior, the neuron indicates intensity of stimulation by the **frequency** of spikes, not amplitude.

- Because of the refractory period, there is a maximum frequency, at which the frequency **saturates**.
Spiking Frequency of a Neuron as a Function of (Artificial) Stimulus Magnitude

Larger stimulus higher frequency of output spiking

Intensity Abstraction

- In *artificial* neural networks, we associate an output value with a neuron, which might be continuous, but have an upper limit.

- This can be viewed as a convenient abstraction of what is really the value of spiking frequency.
• In addition to frequency, information can be encoded based on whether the neuron is spiking regularly ("beating") or in bursts ("bursting").
Various Firing Patterns

Silent

Beating

Bursting

Different Patterns of Responses to a Given Stimulation

Sizes, Scale

- Human estimated to have $10^{10} - 10^{11}$ neurons.
- One neuron may connect to $10^2 - 10^3$ others.
- Therefore $10^{12} - 10^{14}$ connections are present.
Speeds

- Switching speed ~ 1 kHz
  (1 million times slower than a computer)

- Conduction speed ~ 100 m/s
  (vs. near speed of light in a computer)

- Switching energy ~ $10^{-16}$ joules/op
  (vs. $10^{-5}$ joules/op for today’s computers)
Human Nervous System

- Accounts for 1-2% of body’s weight
- Consumes ~ 25% of body’s energy
Transition to *Artificial Neural Networks*
Neural network schematic

Synaptic “weights”
(strength of connection)

Many-to-many connections
Neural network (further schematized)
Generality

- Do we lose generality assuming a regular connection pattern?

- Do we lose generality assuming no cycles?
Abstract Functional Characteristics of Neurons

- Weighted sum multiple synaptic inputs
  - positive weight: “excitatory”
  - negative weight: “inhibitory”

- Threshold triggering phenomenon:
  - weighted sum of inputs must exceed threshold in order to cause an event.
How might a neural network learn?
The Organization of Behavior

A NEUROPSYCHOLOGICAL THEORY

D. O. HEBB

McGill University

1949

New York · JOHN WILEY & SONS, Inc.
London · CHAPMAN & HALL, Limited
A NEUROPHYSIOLOGICAL POSTULATE

Let us assume then that the persistence or repetition of a reverberatory activity (or “trace”) tends to induce lasting cellular changes that add to its stability. The assumption can be precisely stated as follows: When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.
Hebb’s Postulate

When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it,

some growth process or metabolic change takes place in one or both cells

such that A’s efficiency, as one of the cells firing B, is increased.
“When a postsynaptic neuron becomes depolarized [fires], it generates a biochemical reaction or a trophic factor that stabilizes [strengthens] the excitatory synapses that are firing at that time.”
Colloquial Hebb

Neurons that fire together wire together.
“An important aspect of [Hebb’s] hypothesis is that a given presynaptic input to a cell need not, by itself, be of sufficient strength to induce a large depolarization in its target.

If that input is fired at the same time as a number of other inputs, and their combined action depolarizes the cell, all of these inputs will tend to be stabilized.”
“If, in contrast, a given input fires asynchronously with most of the other inputs onto that cell, this input will tend to be eliminated.”

[This could be called “anti-Hebbian” learning.]
Axon inactive or active only at times other inputs are not active.

Simultaneously active axons.

Depolarization of postsynaptic cell.

No stabilization of inactive synapses.

Stabilization of all active synapses.

Elimination of unstabilized inputs.

Hebb’s rule. Excitatory synapses that successfully stimulate a post-synaptic neuron, or are active when the postsynaptic neuron is depolarized, are selectively stabilized.
Some NN Historical Highlights

- 1943  McCulloch and Pitts, Linear Threshold Logic Gate models
- 1949  Hebb, proposed Learning principle
- 1957  Rosenblatt's Perceptron
- 1960  Widrow & Hoff's Adaline
- 1969  Minsky & Papert (MIT), Limitations of perceptrons
Historical Highlights (cont’d)

- **1970-1980** The “neural-net winter"


- **1982** Hopfield (Princeton, then Caltech)
  Hopfield networks

- **1986** Rumelhart and McClelland, popularized backpropagation in multi-layer perceptrons, published "Parallel Distributed Processing"
Characteristics of Simple ANN Models

- “weight” = strength of connection
- threshold = value of weighted input below which no response is produced
- signals may be:
  - real-valued, or
  - binary-valued:
    - “unipolar” \( \{0, 1\} \)
    - “bipolar” \( \{-1, 1\} \)
McCulloch-Pitts Model, 1943

- Synchronous operation
- Binary (uni-polar) signals
- Linear threshold gates
McCulloch-Pitts Neural Model

output = \begin{cases} 
1 & \text{if } w_1 x_1 + w_2 x_2 - w_3 x_3 > q \\
0 & \text{otherwise}
\end{cases}

if we allow weights to be signed
How Powerful is a Network of McCulloch-Pitts Neurons?
Can *any* switching function be represented?
Kleene’s paper, 1956

- “Representation of Events in Nerve Nets”
- Used McCulloch-Pitts model with possible feedback connections
- Assumed synchronous model (not realistic?)
- “Events” are essentially what we now call regular expressions
- Provides an exact characterization of what McCulloch-Pitts network can do