Harvey Mudd College

CS 152
Neural Networks
Fall 2000

Bob Keller, Professor
keller@cs.hmc.edu
621-8483

Office Hours (1242 Olin):

- Note: 1242 is inside of 1240, near the SE entry of Olin
- Monday and Wednesday 3-4; Tuesday 4-5
- By drop-in (as available)
- By appointment:
  - email keller@cs.hmc.edu
  - phone 621-8483
- Crisis center: 621-2373

Text


which I will abbreviate “NAS”.

Course Outline

Please Refer to Web Page for details:

http://www.cs.hmc.edu/courses/2000/fall/cs152

Neural Networks: an Eclectic Discipline

Intelligence

- It is generally accepted that intelligence, the ability to make decisions based upon input from the environment, is realized by a network of neurons, for example the brain and the attendant sensory and motor neurons.
Approaches to Artificial Intelligence

- Reverse Engineering
  - Understand real neurons
  - Build simulator, simulate neural behavior
- Simulated Evolution
  - Provide basic evolutionary mechanism for neurons
  - Evolve intelligent behavior
- Traditional 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 the model can learn

Some Applications of Artificial Neural Networks (1 of 5)

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

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)
  - Mortgage risk evaluator (AVCO, Irvine)
  - Real-estate evaluation (Foster Onsley Conley)
  - Portfolio management (LBS Captial)
  - Currency trading (Citibank)
  - 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

Neuron Cell (top half)

Neuron Cell (bottom half)

Structure of one neuron

Photomicrograph of one neural cell (from cerebral cortex)

Some Physiological Aspects of Neurons

Neuron Cell (top half)

Neuron Cell (bottom half)

Structure of one neuron

Photomicrograph of one neural cell (from cerebral cortex)
Photomicrograph of network of neural cells (from the hippocampus)

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.

Electronmicrograph of one synapse/dendrite connection

Schematic sometimes used (symbolic of synaptic clefts)

In addition to signal, axon carries:

- Construction material (proteins)
- Nutrients (in the form of mitochondria)
- Enzymes

Material Flow in Neuron Established by Experiment ~1930

Paul Weiss: Vesicles flow from soma to terminal synapse. When axon is ligated, vesicles observed to accumulate on side opposite synapse. (Note: not information flow.)
Inter-Neuron Signaling

- An ionic chemical reaction (electrical in invertebrates) carries the signal across the gap between a synapse of one neuron and a dendrite of the next.
- The strength of this connection is determined by the efficiency of the transfer.

Neurotransmitters

- The molecules that traverse from synapse to dendrite.
- A process of ion diffusion is involved.

Chemical Synapse

Ionic Neurotransmitter Reaction

Composition of the Brain

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

- Acts as insulator
- Current can flow out only at nodes (called nodes of Ranvier) where it can “jump” to other axons
- Demyelinating diseases:
  - Myelin deficit in newborns
  - MS (multiple sclerosis)
  - ALS (amyotrophic lateral sclerosis, “Lou Gehrig’s disease”)

Neural network schematic

- Synaptic “weights” (strength of connection)
- Many-to-many connections

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.

Generality

- Do we lose generality assuming a regular connection pattern?
- Do we lose generality assuming no cycles?
### Triggering phenomenon

**Stimulus (summed inputs)**

**Response**


### Signal Encoding

- Magnitude?
- Frequency?
- Phase?
- Combinations of the above?
- Other, e.g. patterns?

### Signal Encoding: Signal Encoding

- **Signal Encoding**

### Spiking Frequency of a Neuron as a Function of Stimulus Magnitude


- Larger stimulus higher frequency of output spiking

### Various Firing Patterns


### 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.

### Different Types of Responses to a Given Stimulation

### Speeds

- Switching speed ~ 1 kHz
- Conduction speed ~ 100 m/s
- 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

### How might a neural network learn?

- **Hebb's Postulate, 1949**

  *The Organization of Behavior*
  
  A NEUROPSYCHOLOGICAL THEORY

  D. O. Hebb

  Bern: CHAPMAN & HALL, Limited

  New York - JOHN WILEY & SONS, Inc.

  London - CHAPMAN & HALL, Limited

  1949

  A NEUROPSYCHOLOGICAL POSTULATE

  Let us assume then that the persistence or repetition of a postsynaptic activity (or "trace") leads to some long-lasting physical change that adds to its excitability. The assumption can be made

  simply 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.*

  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 Restated
(Levitan and Kaczmarek, p 351)

“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.”

Levitan and Kaczmarek, p 351

“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.”

Levitan and Kaczmarek, p 351 (cont’d)

- “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.]

Levitan and Kaczmarek, p 351 (cont’d)

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)

- 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

\[
\text{output} = \begin{cases} 
1 \text{ if } w_1 x_1 + w_2 x_2 + w_3 x_3 > \theta \\
0 \text{ otherwise} 
\end{cases}
\]

How Powerful is a Network of McCulloch-Pitts Neurons?

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

Can any switching function be represented?