Connectionist Computing
COMP 30230/41390

Gianluca Pollastri

office: E0.95, Science East.

email: gianluca.pollastri@ucd.ie

Credits

- Geoffrey Hinton, University of Toronto.
  - borrowed some of his slides for “Neural Networks” and “Computation in Neural Networks” courses.

- Ronan Reilly, NUI Maynooth.
  - slides from his CS4018.
- Paolo Frasconi, University of Florence.
  - slides from tutorial on Machine Learning for structured domains.

Lecture notes on Brightspace

- Strictly confidential...
- Slim PDF version will be uploaded later, typically the same day as the lecture.
- If there is demand, I can upload onto Brightspace last year's narrated slides.. (should be very similar to this year's material)

Books

- No book covers large fractions of this course.

- Parts of chapters 4, 6, (7), 13 of Tom Mitchell's "Machine Learning"
- Parts of chapter V of Mackay's "Information Theory, Inference, and Learning Algorithms", available online at:

http://www.inference.phy.cam.ac.uk/mackay/itprnn/book.html

- Chapter 20 of Russell and Norvig's "Artificial Intelligence: A Modern Approach", also available at:

http://aima.cs.berkeley.edu/newchap20.pdf

- More materials later..

Marking

- **3 landmark papers to read, and submit a 10-line summary on Brightspace about: each worth <u>6-7%</u>**
- **a connectionist model to build and play with on some sets, write a report: <u>30%</u>**
- **Final Exam in the RDS (<u>50%</u>)**

Connectionism

- **A computational approach to modelling *the brain* which relies on the interconnection of *many simple units* to produce complex behavior**
- **Not the usual paradigm where there is a powerful central processor that executes serially a static program:**
  - **simple elements,**
  - **parallel processing,**
  - **learning..**

Connectionism, AI and Deep Learning

- Connectionism starts from Cognitive Science and the study of the brain.
- Between the mid 80s and circa 2005 it produced the vast majority of what we now call Deep Learning.
- While between 2006 and now Deep Learning has come up with new algorithms, much of what differentiates it from Connectionism is about emphasis and packaging rather than genuinely new science.

Facts about the brain

- The brain consists of around 10^{11} neurons.
- Neurons are connected: each neuron receives between 10^3 and 10^4 connections. Hence there are 10^{14} to 10^{15} connections in the brain (100-1000 Tbytes to store 1 number for each of them).

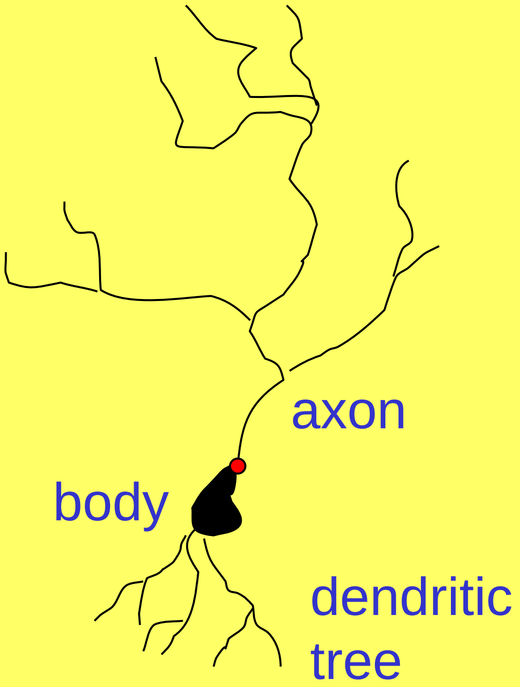
- The "currency" of the brain is the action potential or voltage spike.
- There appears to be considerable localisation of function in the brain.

A typical cortical neuron

- **Gross physical structure:**
  - **One axon that branches**
  - **A dendritic tree that collects input from other neurons**

- **Axons typically contact dendritic trees at synapses**
  - **A spike of activity in the axon causes charge to be injected into the post-synaptic neuron**

- **Spike generation:**
  - **Outgoing spikes whenever enough charge has flowed in at synapses to depolarise the cell membrane**



Some artificial neurons

activation

synaptic weight

$$z = b + \sum_i x_i w_i$$

bias

input

$$\begin{array}{ll} f(z) = z & linear \\ f(z) = step(z) & McCulloch - Pitts \\ f(z) = \sigma(z) & sigmoid \end{array}$$

$$step(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases}$$

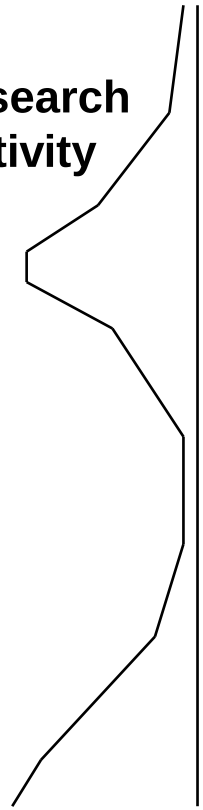
$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Learning

- We have a few simple models of neurons
- Their behaviour depends on the value of their synaptic weights (free parameters).
- How do we learn/adjust parameters?

A bit of history first

research activity



McCulloch & Pitt's formal neuron (1943)

Hebb's learning rule (1949)

Rosenblatt's Perceptron (1957)

Associators: Anderson (1972), Kohonen (1972)

Hopfield's network (1982)

Hinton, Sejnowski, & Ackley's Boltzmann learning algorithm (1984)

Rumelhart, Hinton & Williams error backpropagation learning algorithm (1986)

"old" connectionism

"new" connectionism

Hebbian learning

"When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes takes place in one or both cells such that A's efficiency as one of the cells firing B, is increased ." **(Hebb, 1949)**

Rosenblatt’s perceptron (’57)

- **A binary neuron, inputs possibly real-valued:**

$$z_j = \sum_i w_{ji} x_i$$

$$y_j = \begin{cases} 1 & \text{if } z_j \geq 0 \\ 0 & \text{if } z_j < 0 \end{cases}$$

- **This is really just a McCulloch-Pitts neuron.. but Rosenblatt actually implemented it on a computer..**

Rosenblatt’s perceptron

- .. and introduced the idea of training in practice by applying something close to the Hebbian learning rule:

learning rate

desired output – actual output
(supervision!)

$$\Delta w_{ji} = \eta(d_j - y_j)x_i$$

input

learning (weight change)

Rosenblatt's perceptron

- This rule assumes that we have examples
- Notice that, since this is a binary neuron, (d-y) can only be: -1, 0, 1

learning rate

desired output – actual output
(supervision!)

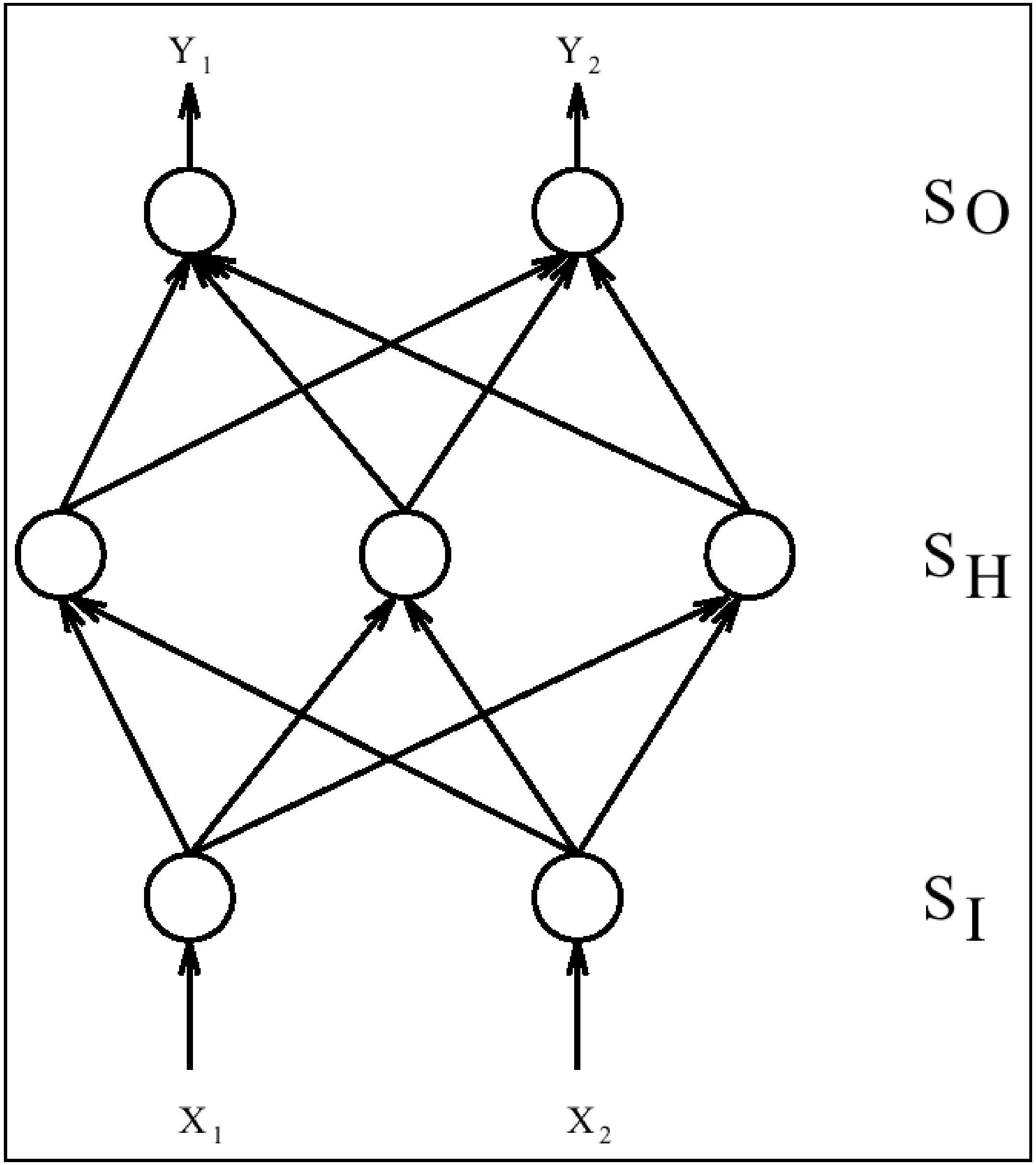
$$\Delta w_{ji} = \eta(d_j - y_j)x_i$$

input

learning (weight change)

Rosenblatt's contributions

- McCulloch-Pitts neuron ('43) and Hebbian learning ('49) implemented in practice.
- Computer simulations to study the behaviour of perceptrons
- Mathematical analysis of their properties
- Rosenblatt was probably the first who talked of "connectionism".
- He studied also perceptrons with multiple layers (see next slide..)
- He called "error backpropagation" a procedure to extend the Hebbian idea to multiple layers.



Connectionist Computing
COMP 30230

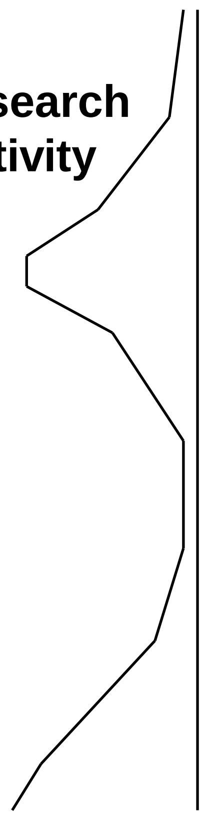
Rosenblatt's contributions: the dark side

- Expectations were raised that the perceptron would be the endgame of AI.
- When it became clear that wouldn't happen, research in the field went into the wilderness for over a decade.

- There is a history of boom and bust in the field. Wait 'til we discover that Deep Learning can't do general purpose AI.

Minsky-Papert ‘69

research activity



McCulloch & Pitt's formal neuron (1943)

Hebb's learning rule (1949)

Rosenblatt's Perceptron (1957)

Associators: Anderson (1972), Kohonen (1972)

Hopfield's network (1982)

Hinton, Sejnowski, & Ackley's Boltzmann learning algorithm (1984)

Rumelhart, Hinton & Williams error backpropagation learning algorithm (1986)

"old" connectionism

"new" connectionism

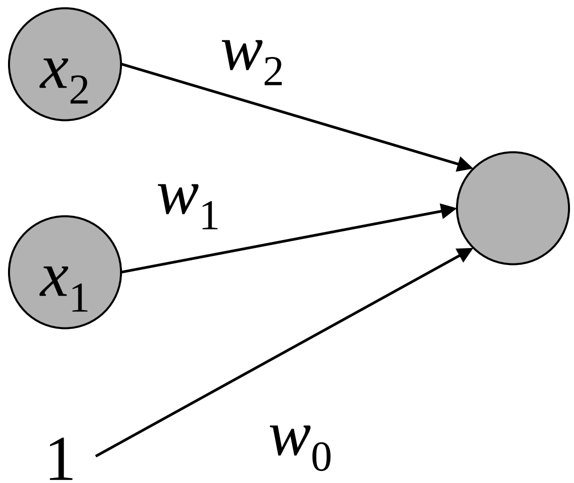
Minsky and Papert

- Marvin Minsky and Seymour Papert published a book on perceptrons in 1969.
- The book dominated the discussion on perceptrons in AI until the mid ‘80s.
- While the book did a whole lot more than trying to discredit perceptrons, what stuck were proofs that specific types of perceptron cannot implement some simple functions.

Minsky and Papert’s analysis

- **Studied simple, one-layered perceptrons**
- **Proved mathematically that some tasks just cannot be performed by such perceptrons:**
  - **whether there is an odd or an even number of inputs firing it**
  - **XOR**
- **While it was clear to the authors that wider/deeper networks of perceptrons could solve these problems, they maintained these networks were not affordable.**

Minsky & Papert: XOR



(This is how we draw neurons)

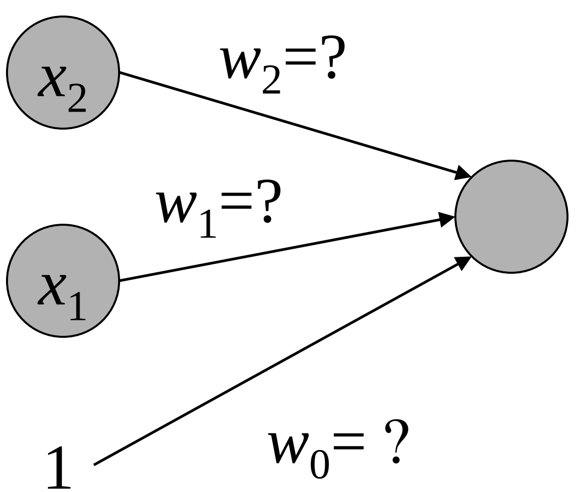
- Output is binary-thresholded linear combination of inputs:

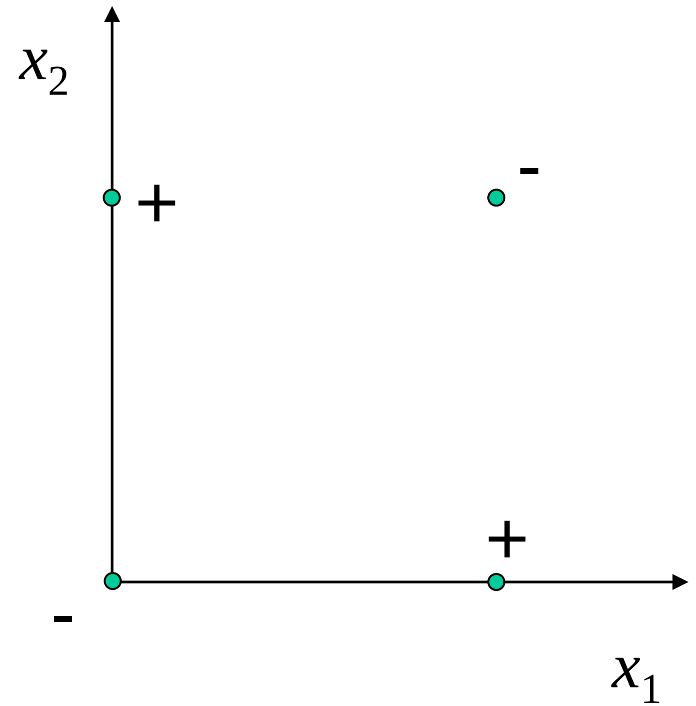
$$y = step(w_0 + w_1 x_1 + w_2 x_2)$$

- The perceptron draws a line in the output space.

Minsky & Papert: XOR (2)

x_1	x_2	$f(x)$
0	0	0
0	1	1
1	0	1
1	1	0





No linear separation surface exists!!

The aftermath

- After the 1969 critique by Minsky and Papert many AI researchers (and funding agencies) perceived connectionism as fruitless.
- (this is a little bizarre since their proofs applied only to a very basic kind of perceptron)

The aftermath (2)

- Only during the 80s was there a renaissance of connectionism:
  - "other" (rule-based) AI was perceived as stagnant
  - computers were finally getting a lot faster/cheaper
  - mathematicians and statisticians took connectionism over from cognitive scientists:
    - new developments (to be continued)
  - (at this point even Minsky partially retracted..)

- Minsky and Papert's book even became a springboard for new connectionism: "We have overcome all these problems!"