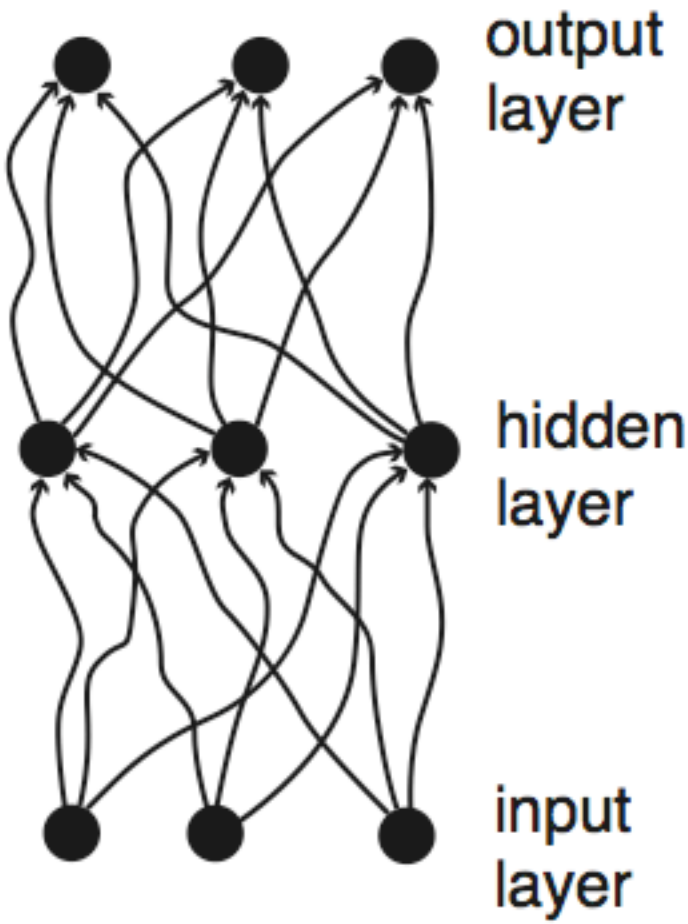
Multilayer perceptrons

Sebastian Seung

Layered networks

- Two layers of LT neurons
  - (three layers if input neurons are included)
- Two layers of synapses.
- No loops



Any Boolean function can be computed by a perceptron with two layers of synapses.

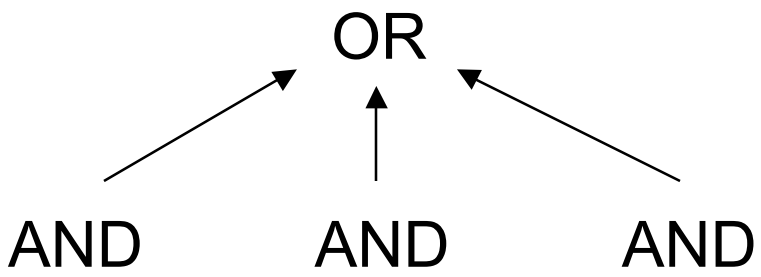
Truth table

- N variables
- 2^N rows
- 2^{2^N} possible functions

x_1	x_2	x_3	$f(x_1,x_2,x_3)$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Disjunctive normal form (DNF)

- Any boolean function can be written in disjunctive normal form.
- Disjunction of conjunctions



DNF construction

x_1	x_2	x_3	$f(x_1,x_2,x_3)$	
0	0	0	0	
0	0	1	0	
0	1	0	0	
0	1	1	0	
1	0	0	0	
1	0	1	1	$\longrightarrow x_1 \wedge \overline{x}_2 \wedge x_3$
1	1	0	1	$\longrightarrow x_1 \wedge x_2 \wedge \overline{x}_3$
1	1	1	1	$\longrightarrow x_1 \wedge x_2 \wedge x_3$

$$f = (x_1 \wedge \overline{x}_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge \overline{x}_3) \vee (x_1 \wedge x_2 \wedge x_3)$$
$$= x_1 \wedge (x_2 \vee x_3)$$

Any DNF can be written as a perceptron

- LT neuron
  - AND of N variables or their negations
  - OR of N inputs
- Two layers of LT neurons required

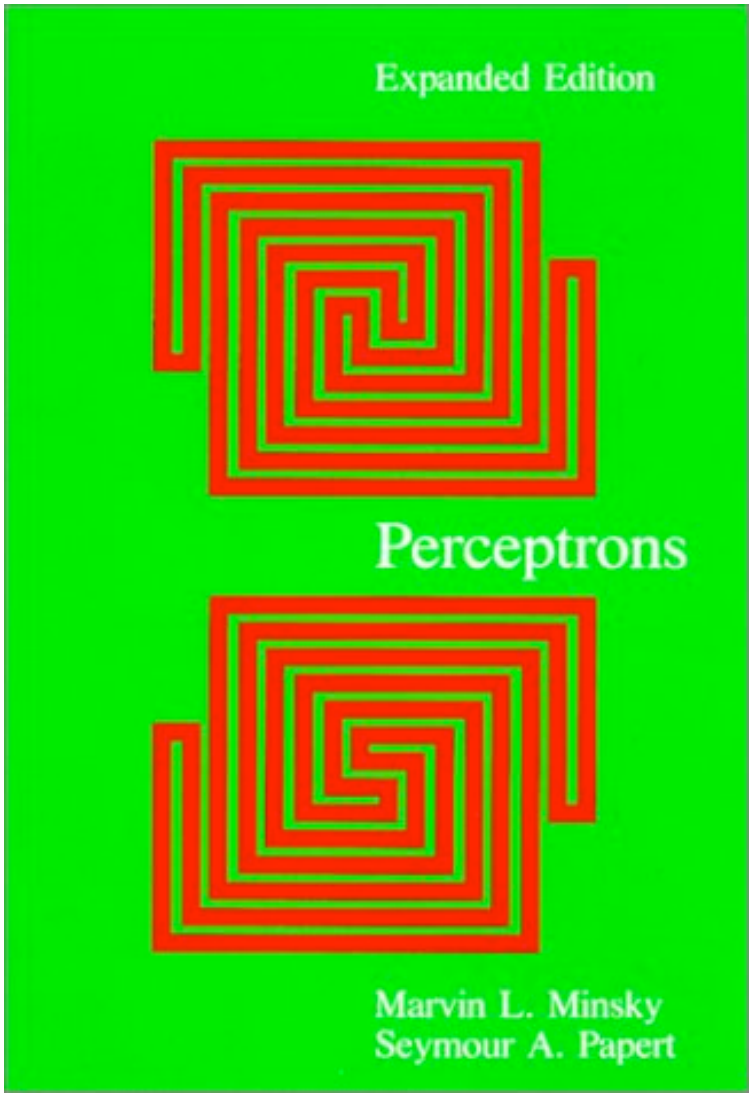
What's the catch?

- The number of conjunctions required may be exponentially large.
- I.e., there is no guarantee that the perceptron representation is efficient.

Efficiency

- Number of synapses
- Serial computer
  - Time
- Parallel computer
  - Space
  - Energy

Connectedness



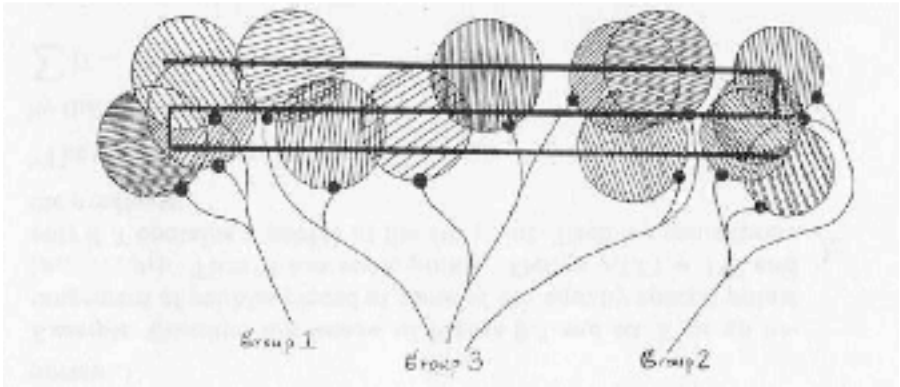
Minsky-Papert definition

- Requires only that the final step in the computation is an LT neuron.

$$H\left(\sum_{a} w_a \varphi_a(\mathbf{x})\right)$$

Diameter-limited perceptrons

- Assume that the input vector is organized as a 2d image.
- Each φ depends on a set of pixels with a limited diameter.

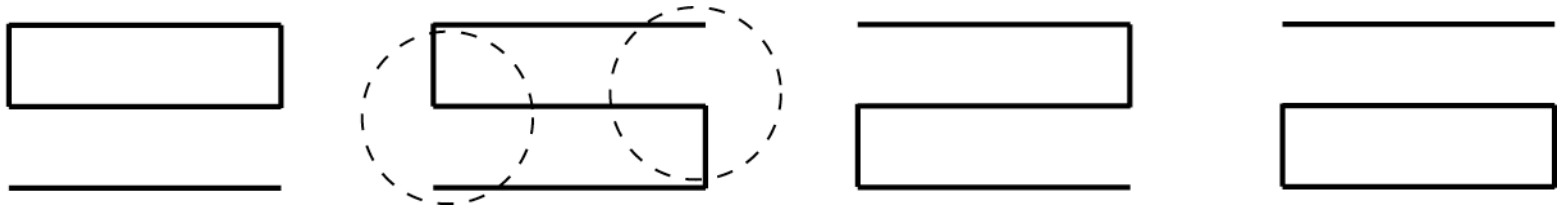


Theorem

No diameter-limited perceptron can compute connectedness.

Proof

- Reduce to XOR



Any smooth function can be approximated by a perceptron with two layers of synapses and a sigmoidal activation function.

Further reading

- Minsky and Papert, Perceptrons, expanded edition, MIT Press (1988).