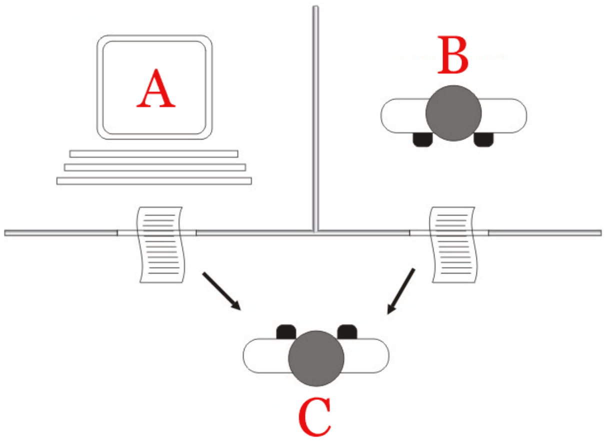
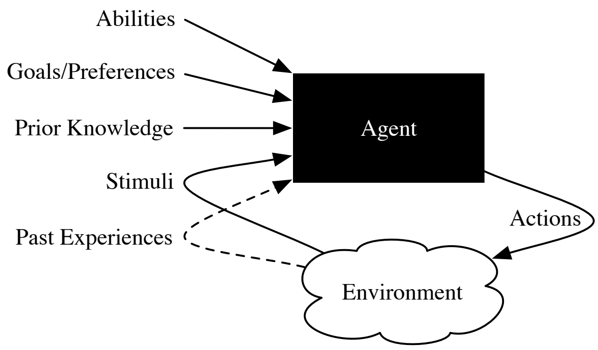
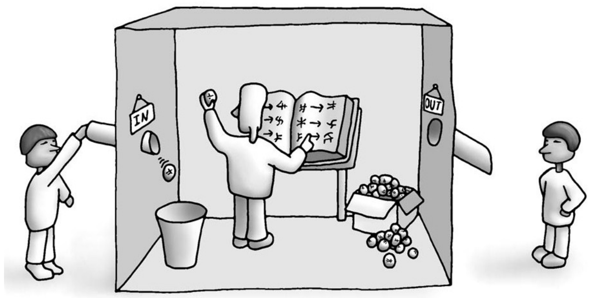
Peering inside the black box



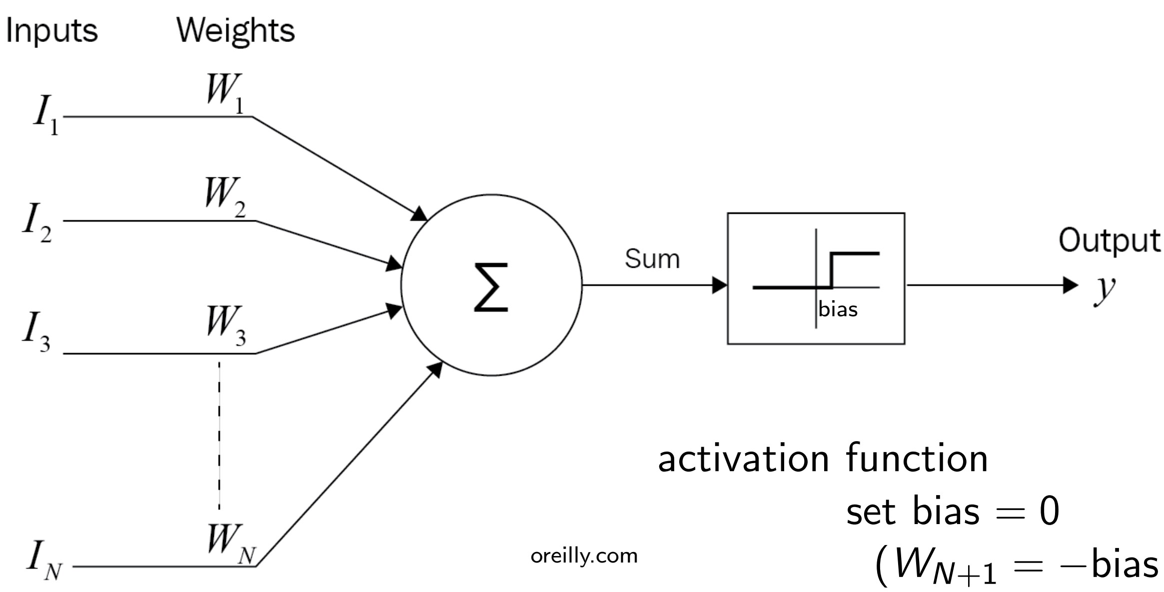




Agents from neurons (perceptrons)

$$\text{Sum} = W_1 I_1 + W_2 I_2 + \cdots + W_N I_N$$

$$y = \begin{cases} 1 & \text{if Sum} \geq \text{bias} \\ 0 & \text{if Sum} < \text{bias} \end{cases}$$



activation function

set bias $= 0$

$(W_{N+1} = -\text{bias}$

$I_{N+1} = 1)$

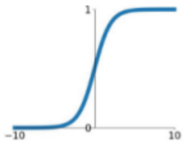
Beyond the binary threshold unit

$y = \textbf{step}(\sum_i W_i I_i)$ where

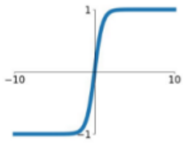
$$\textbf{step}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

non-linear

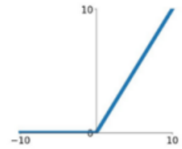
Sigmoid
$\sigma(x) = \frac{1}{1+e^{-x}}$



tanh
$\tanh(x)$



ReLU
$\max(0, x)$



Values other than 1,0

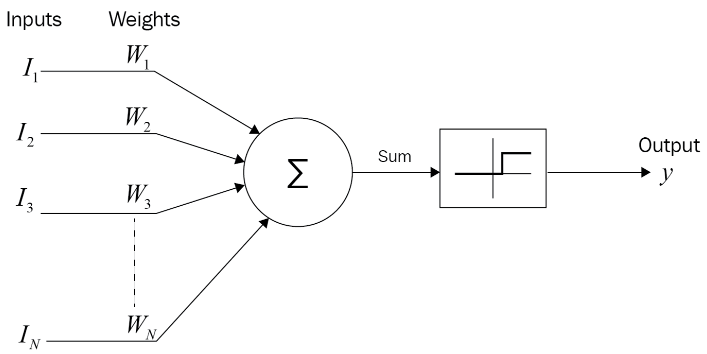
digital $\rightsquigarrow$ analog devices (gradients ... calculus)

Output y as input I_i to any number of neurons

often layered: input, hidden, output

structured: away from biology ...

Finite state aggregations



Kleene 1956

$$I_1, I_2 \cdots I_N \;=\; q, a$$

$q \overset{a}{\rightsquigarrow} y \iff$
output y on
input a at q

input a given by k *input cells* $N_1 \dots N_k$
state $q = (v_1 \dots v_m)$ records the
values v_i of m *inner cells* $M_1 \dots M_m$

McCulloch & Pitts 1943: neural activity is *all-or-none*

Kleene 1956: input cells are binary, forming a string a of k bits
inner cell M_i can take one of s_i many values

$$q \overset{a}{\to} (y_1 \dots y_m) \iff q \overset{a}{\rightsquigarrow}_i y_i \text{ for } 1 \leq i \leq m$$

Rabin & Scott 1959: work with any finite sets A, Q and $\to$

Logical abstractions away from physics and biology

1976 Turing Award to Rabin & Scott for their joint paper

> The internal workings of an automaton will not be analyzed too deeply. We are not concerned with how the machine is built but with what it can do. The definition of the internal structure must be general enough to cover all conceivable machines, but it need not involve itself with problems of circuitry. The simple method of obtaining generality without unnecessary detail is to use the concept of internal states. No matter how many wires or tubes or relays the machine contains, its operation is determined by stable states of the machine at discrete time intervals. An actual existing machine may have billions of such internal states, but the number is not important from the theoretical standpoint — only the fact that it is finite.

page 115

Back to Turing 1950: §7. Learning Machines

In considering the functions of the mind or the brain we find certain operations which we can explain in purely mechanical terms. This we say does not correspond to the real mind: it is a sort of skin which we must strip off if we are to find the real mind.
... Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain ... We have thus divided our problem into two parts. The child-programme and the education process.
... The processes of inference used by the machine need not be such as would satisfy the most exacting logicians.
... Processes that are learnt do not produce a hundred per cent. certainty of result; if they did they could not be unlearnt.
It is probably wise to include a random element in a learning machine ... A random element is rather useful when we are searching for a solution of some problem.

Knowledge Representation & Automata

Logicomix: an epic search for truth

a graphic novel about Bertrand Russell's foray into the foundations of mathematics

Topics for CSU44060

- Logic & Agents as Turing machines
- Description Logics
- Finite-state methods (transducers & MSO)
- Action, change & time
- Trakhtenbrot's & Fagin's theorems
- Entropy (cross, relative, conditional)