Introduction to Computer Science
Lecture 10: Artificial Intelligence

Department of Electrical Engineering
National Taiwan University

What Is AI?

John Searle

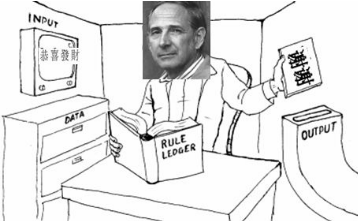
Strong AI : A physical symbol system can have a mind and mental states.

Weak AI : A physical symbol system can act intelligently.

Stuart Russell

	Human	**Rational**
Think	Thinking humanly	Thinking rationally
Act	Acting humanly	Acting rationally

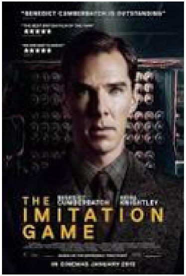
Chinese Room Argument by John Searle

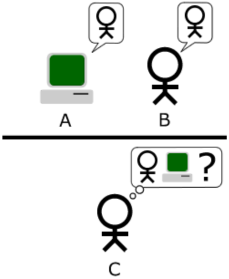


- Suppose it is possible to have a program $\mathcal{P}$ that is sufficient for understanding of Chinese.
- In principle a person in the Chinese room can carry out $\mathcal{P}$.
- But such a person would not understand Chinese.
- So $\mathcal{P}$ is not sufficient for producing understanding of Chinese.
- So there is no program sufficient for producing understanding of Chinese.

Turing Test (1950)

- Operational Test





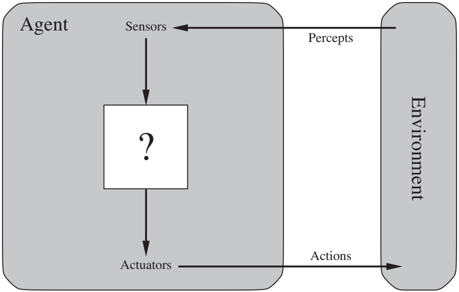
Turing Test Applications

- ELIZA (1965).
- Mitsuku (2016 Loebner prize winner).
- CAPTCHA

Acting Rationally: Rational Agent

- Rational behavior: maximizes the expected performance, given available information.
- Doesn't necessarily involve thinking – *e.g.*, blinking reflex
- Rational $\neq$ Omniscience
  - Percepts may not supply all relevant information.
- Rational $\neq$ Clairvoyant
  - Action outcomes may not be as expected.
- High-level rationality $\Rightarrow$ information gathering, exploration, learning, autonomy.

Agents and Environments



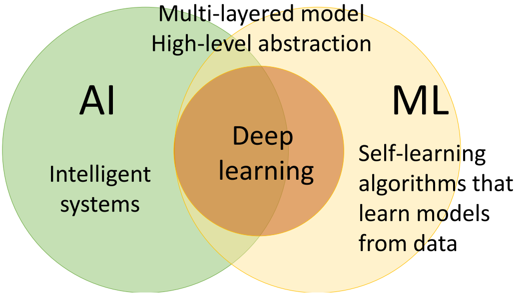
- An agent is an entity that perceives and acts.
- The agent function maps any percept sequences to an action.

$$f : \mathcal{P}^* \to \mathcal{A}$$

- An agent program is the implementation of an agent function.
- Computational limitations → limited rationality.

Different Levels of AI

- Level 0: Simple reflex, marketing “AI”.
- Level 1: Search, planning based on some knowledge base.
- Level 2: Learning, exploration.
- Level 3: Automatic feature generation, high-level abstraction.



3 Waves of AI

1950~1970

- Search
- Neural Networks

Debut

1980~1995

- Expert Systems
- Knowledge

Knowledge Driven

2010~now

- Machine Learning
- Automatic Feature Generation

Data Driven

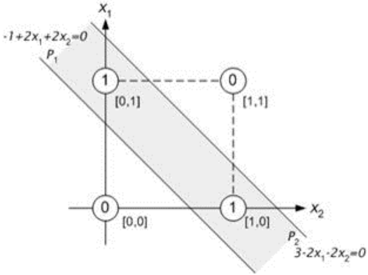
Achievement and Bottlenecks

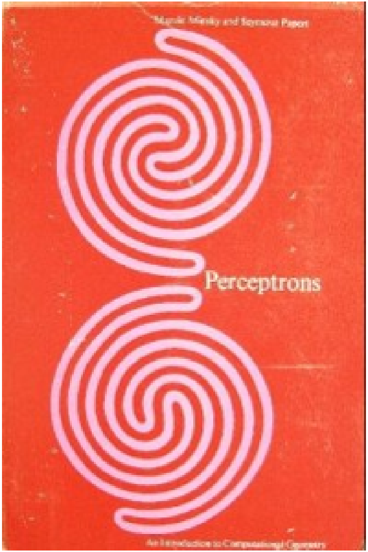
- Milestone achievement: Deep Blue beat Kasparov in 1997.

- “Seems” mainly for “toy” problems.
- Difficult to directly apply to real problems.
  - How to treat patients?
  - Which product should we develop?

Decline of ANN

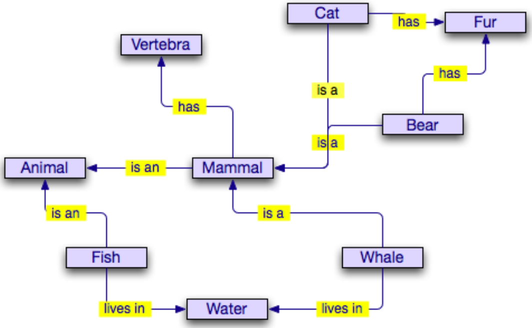
- Minsky and Papert, MIT (1969)
- XOR problem.
- Connection problem.





Expert Systems

- Standford's MYCIN (1960): Diagnosis of infections (69% vs. 80%).
- Cyc project (Lanet, 1984): Semantic network.



Achievement and Bottleneck

- Watson beat human champions in Jeopardy! in 2011.

- Ontology problem: e.g. what is “part of”?
- Common sense
  - He saw a woman in the garden with a telescope.
- Frame problem (McCarthy).
- Symbol grounding problem (Harnad): Embodiment.

Natural Language Processing

- Syntactic analysis
  - Mary gave John a birthday card. Subject: Mary
  - John got a birthday card. Subject: John
- Semantic analysis, contextual analysis
  - John drove me home.
  - John drove me crazy.

  - The pigpen was built by the barn.
  - The pigpen was built by the farmer.

  - Do you know what time it is?

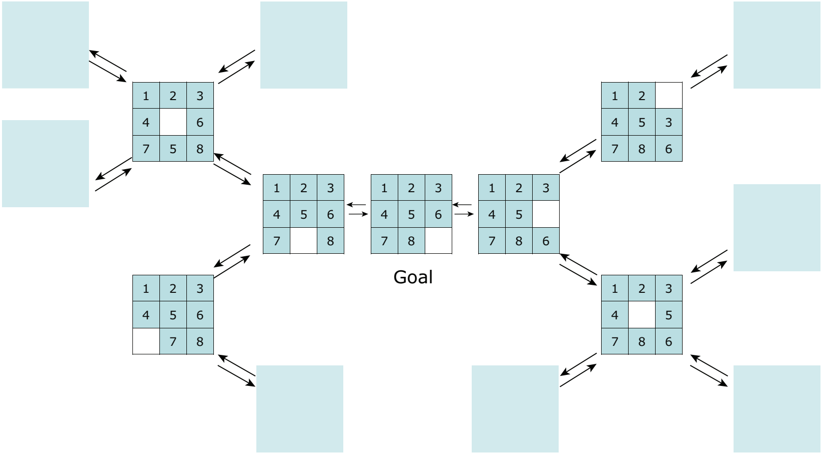
Natural Language Processing (contd.)

- Information retrieval / extraction
  - It's raining cats and dogs.
  - Ya, I just stepped on a poodle.

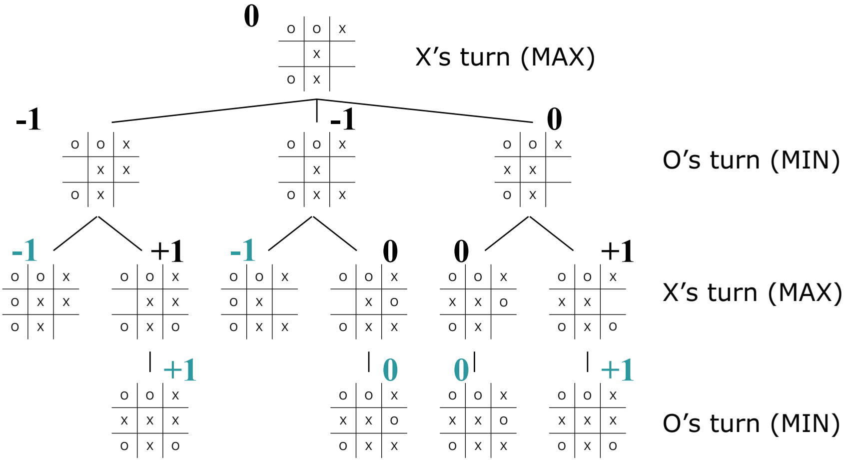
  - What do lawyers do when they die?
  - Lie still.

  - You can count on me.
  - Ya, right. That's comforting.

Search a Production System



Game Tree & Minimax Search



Heuristic

- For most games, a complete search is practically impossible.
  - Chess $\sim 10^{47}$; Chinese chess $\sim 10^{48}$; Go $\sim 10^{171}$
- A quantitative estimate of the distance to a goal is needed.
- Requirements for good heuristics
  - Much easier to compute than a complete solution
  - Reasonable estimate of proximity to a goal

Let's Define a Heuristic

XX_	100
X__	10
___	0
OX*	0
O__	-10
OO_	-100

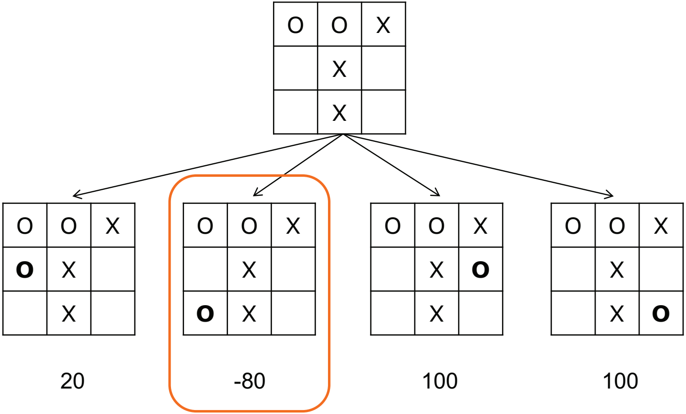
○	○	×
	×	
	×	

$\longrightarrow$

0+10+10-10+0+10+0+100 = 120

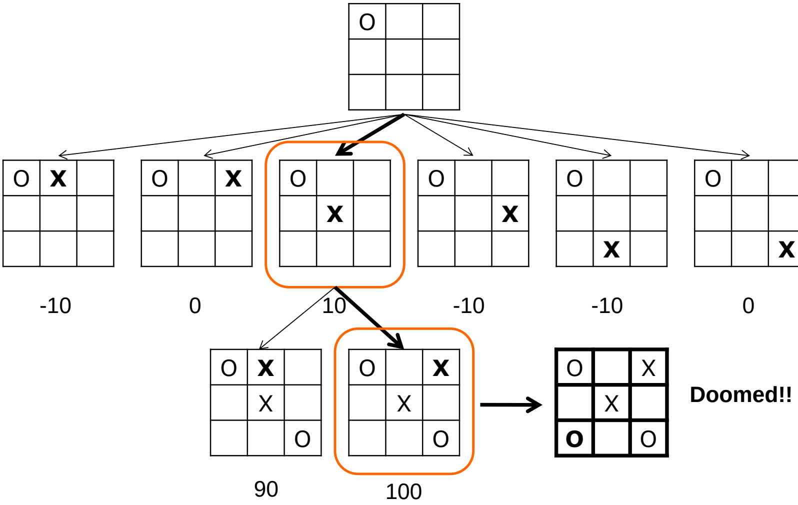
The board favors X

Does It Work?



This is the best choice for O based on our heuristic.

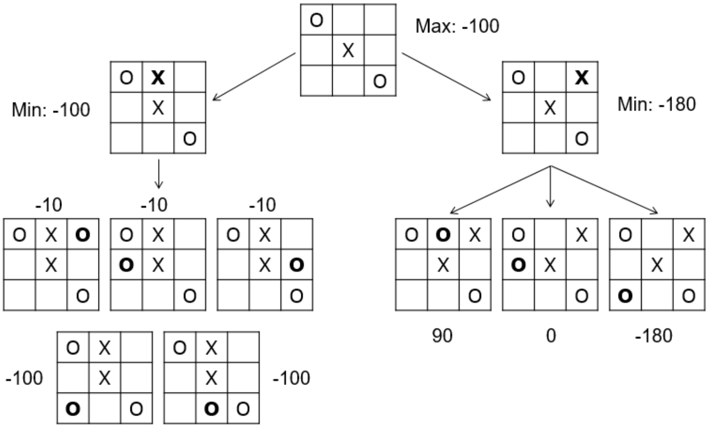
How About This?



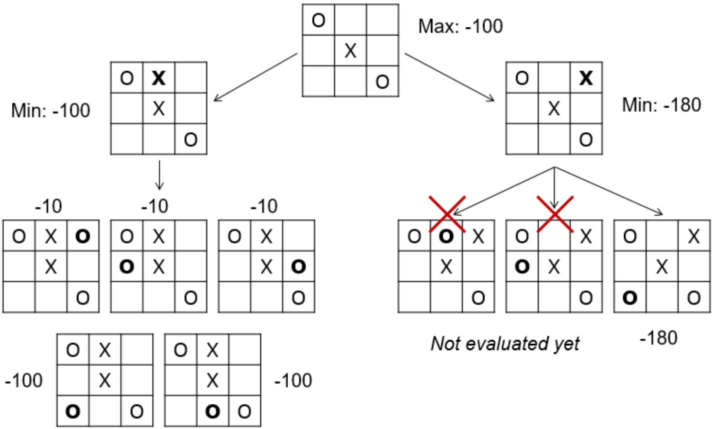
What's Wrong?

- Heuristics are not perfect
  - Otherwise, we'd call them solutions
- Heuristics are usually more accurate toward the end of the game.
- Need some search procedure for more accurate estimation.

Heuristic + Minimax Search



Alpha-Beta Pruning

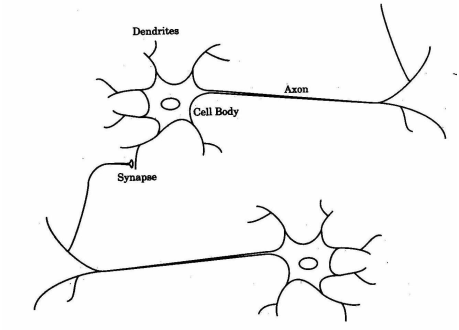


Learning

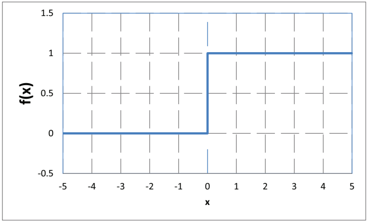
- Supervised vs. unsupervised
- Supervised
  - Learning by provided examples
  - Imitation
  - Parameter tuning
- Unsupervised
  - Learning by experiences
  - Reinforcement
  - Evolutionary (semi-supervised)

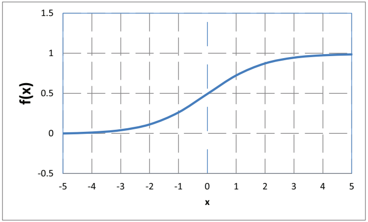
Artificial Neural Networks

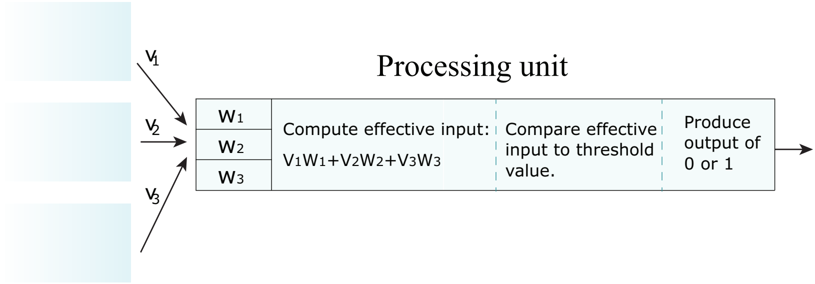
- Human brain
  - 10^{11} neurons
  - 10^{14} synapses



Perceptron



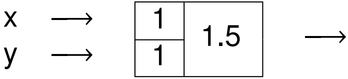




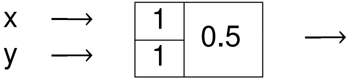
1 *iff* greater than or equal to the threshold

Some Building Blocks

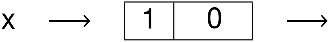
- AND



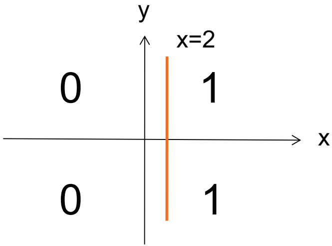
- OR

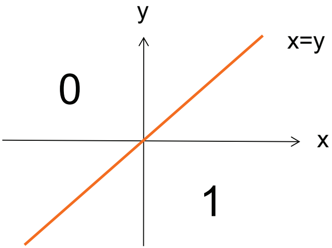


- SIGN



Some Examples

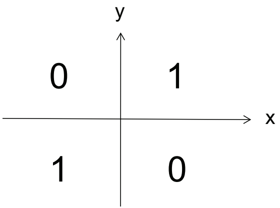


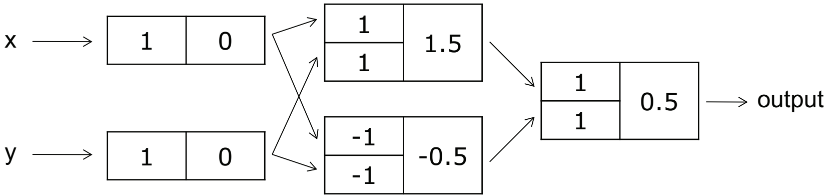


$x \longrightarrow \boxed{1 \mid 2} \longrightarrow$ output

$$\begin{matrix} x \longrightarrow \\ y \longrightarrow \end{matrix} \boxed{\begin{array}{c|c} 1 & \\ \hline -1 & 0 \end{array}} \longrightarrow \text{output}$$

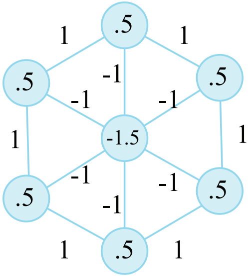
The XOR Problem





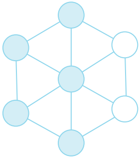
Associative Memory

- Content addressable



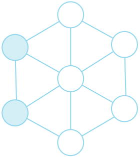
How Does It Work

a.



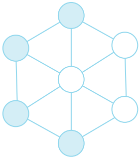
Start: All but the right-most units are excited

b.



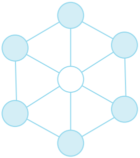
Step 1: Only the leftmost units remain excited

c.



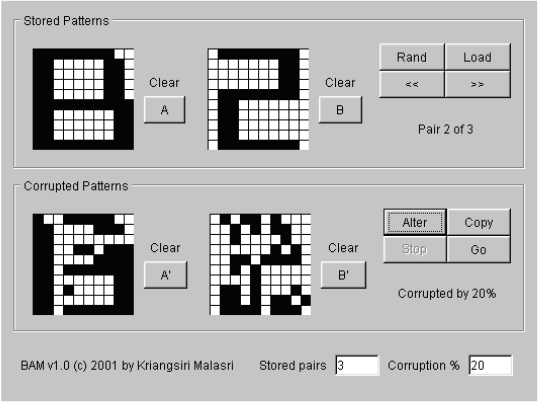
Step 2: The top and bottom units become excited

d.



Final: All the units on the perimeter are excited

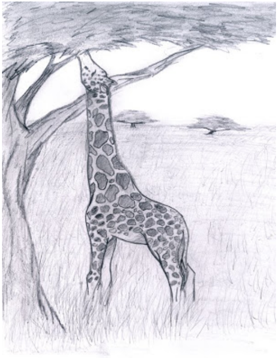
Example



BAM applet: http://www.cbu.edu/~pong/ai/bam/bamapplet.html

Darwin's Theory of Evolution

- Evolution
  - The change in populations of organisms over generations.
- Darwin's idea: Natural selection
  - Struggle to survive
  - Survival of the fittest
  - Genetic variation: inherited traits



Black-Box Optimization

$$x \longrightarrow \boxed{f} \longrightarrow y$$

- Finding the x that yields the highest y with an unknown f
- Evolving the giraffe that is the fittest in an unknown environment.
- Instead of finding a solution, let's evolve a solution.

(1+1) Evolutionary Strategy

- Simplest evolutionary strategy
- One parent: n-dimension real vector, $P = (p_1, ..., p_n)$
- Generate one child by mutation: $C = (c_1, ..., c_n)$
  - $c_i = p_i + N(0, \sigma^2)$
- Replace P by C iff C is better.
- Modify σ according to the replacement rate r.
  - One fifth rule

1/5 Rule Intuition

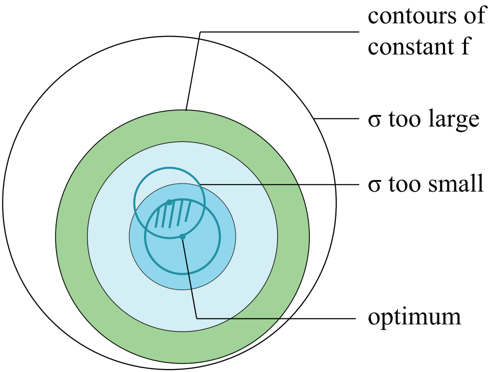
- $\sigma \leftarrow \sigma / C^{1/n}$, if $r > \Theta$
- $\sigma \leftarrow \sigma \cdot C^{1/n}$, if $r < \Theta$

- If replacement rate high, not exploring enough → increase step size.

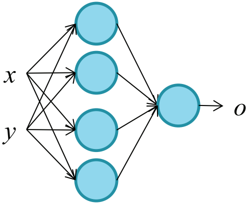
- If replacement rate low, too daring → reduce step size.

- $\Theta = 1/5$ (Guessed by Rechenberg) and $C = 0.817$ (Progress analysis by Schwefel)

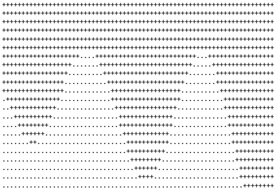
Visualization of 1/5 Rule

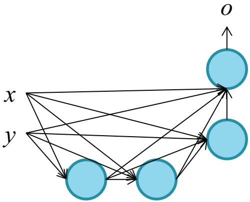


Training NN with (1+1)ES

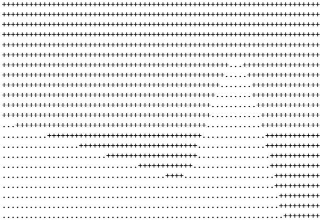


Target:

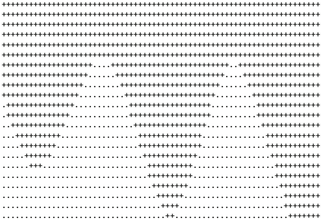




5 neurons, 17 parameters



4 neurons, 18 parameters



Total Differences: 172(123.084600)/1540