Artificial Intelligence

Lecture 26 – COMPSCI 111/111G SS 2018

Super intelligent machines, containment strategies.

What is Artificial Intelligence?

Artificial intelligence is the *computational study of structures and processes that support intelligent behaviour*.

Term first coined in 1956:

- Dartmouth Summer Research Project on Artificial Intelligence

Areas of research include:

- Computer vision
- Natural language processing
- Robotics
- Knowledge-based systems
- Machine learning

Aims of Artificial Intelligence

Three interrelated aims:

- Engineering aim
- Psychological aim
- General/Philosophical aim

Source:

Metaphor and Artificial Intelligence, Why They Matter to Each Other, J.A. Barnden, University of Birmingham
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.136.3416

Engineering Aim

To engineer, or provide computational principles and engineering techniques for, "useful" artefacts that are arguably intelligent.

- Mechanistic similarity to human or animal minds/brains is not necessary.

The artefact may be useful in one of a variety of domains:

- Industry
- Mathematics
- Art
- Everyday life

Psychological Aim

To create computational principles, theories or systems that provide a greater insight on cognition in *human or animal minds/brains*.

General/Philosophical Aim

To create computational principles, theories or systems that provide a greater insight on cognition in *general*.

- Human made artefacts
- Naturally occurring organism
- Cognizant entities yet to be discovered.

Includes looking at philosophical issues like the nature of intelligence, thought, consciousness, etc.

What is Intelligence?

When we say that humans are *intelligent*, we mean they exhibit certain high-level cognitive abilities, including:

- Carrying out complex reasoning
  - E.g., solving physics problems, proving mathematical theorems
- Drawing plausible inferences
  - E.g., diagnosing automobile faults, solving murder cases
- Using natural language
  - E.g., reading stories, carrying out extended conversations
- Solving novel, complex problems
  - E.g., completing puzzles, generating plans, designing artifacts

Does not include:

- Executing motor skills or autonomic activity (breathing, reflexes etc.)

Philosophical View Of Intelligence

Behaviourist/Functionalist approach:

- External behaviour matters
- If it behaves intelligently, then it is intelligent
- Turing test

Cognitive approach:

- What happens internally matters
- We must consider how it thinks, not just look at the behaviour
- Chinese room

The Turing Test

Proposed by Alan Turing in his 1950 paper "Computing Machinery and Intelligence".

- Defines criteria for determining machine intelligence
- "Are there imaginable digital computers which would do well in the imitation game?"

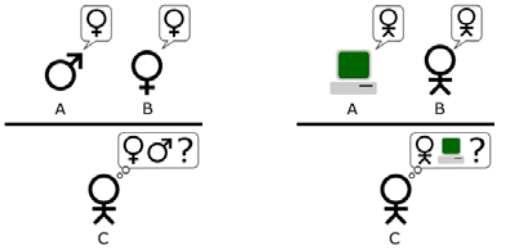
Imitation game:

- Three players – A, B, and C
- A is a man and B is a woman. C, the interrogator is of either gender
- Player C is unable to see either player A or player B
- C asks A and B questions, trying to determine which of the two is a man and which is the woman

Standard Turing test:

- Three players – A, B, and C
- A is a computer and B is a person of either sex. C, the interrogator is also a person of either gender
- Player C is unable to see either player A or player B
- C asks A and B questions, trying to determine which of the two is human and which is the machine

The Turing Test



Imitation game | Turing test

If on completion of the Turing test, C cannot tell A and B apart, then machine A is intelligent.

The Chinese Room

Thought experiment proposed by John Searle in his 1980 paper "Minds, Brains, and Programs".

Refutes functionalist viewpoint:

"The appropriately programmed computer with the right inputs and outputs would thereby have a mind in exactly the same sense human beings have minds"

The Chinese Room

Premise:

- Person in a closed room who has no understanding of Chinese.
- Room contains a manual with instructions detailing the appropriate response, in Chinese characters, to every possible input, also in Chinese characters.
- Person can communicate via written responses with the outside world through a slot in the door.

Scenario:

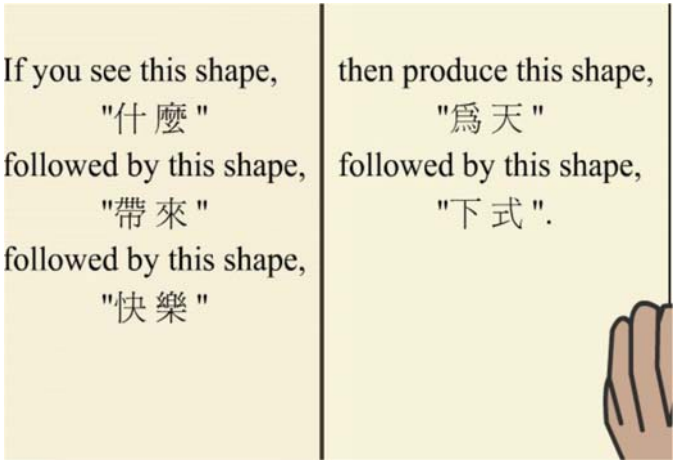
- A Chinese person passes messages written in Chinese, to the person in the Chinese Room.
- Person in the room responds using the manual; they appear to be conversant in Chinese despite not understanding any of the communication.

Argument:

- Without "understanding", a machine's activity cannot be described as "thinking". Since a machine does not think, it does not have a "mind" in the same way you would say a person does.

Source: https://en.wikipedia.org/wiki/Chinese_room

Chinese Room Rulebook



Strong AI versus Weak AI

Strong AI

- The view that a computer could become self-aware and exhibit intelligent behaviour.

Weak AI

- The view that computers could not become self-aware and reason.
- Can be used to solve specific problems in a well-defined domain

Examples of Strong AI

Examples Of Weak AI

IBM Deep Blue

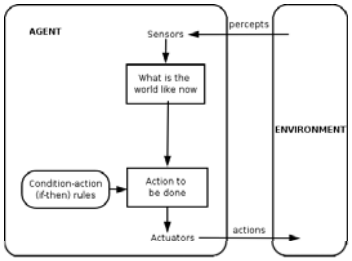
- Chess playing computer
- Won a game against reigning world champion Garry Kasparov in 1996, losing the overall match.
- Won the match against Kasparov in 1997; first computer to do so in a match under standard chess tournament time controls.
- Deep Blue was programmed with history of Kasparov's previous games.
- Programming was modified between games to avoid traps.
- Kasparov was not permitted to study Deep Blue's previous games.

IBM Deep Blue

Examples Of Weak AI

Agents

- Autonomous entity that works in a defined environment.
- Agent achieves goals within environment using:
  - Percepts – observations of the environment obtained through sensors
  - Actions – made on the environment using actuators



Source: https://en.wikipedia.org/wiki/Intelligent_agent

Curiosity Rover

Part of the Mars Exploration Program to study:

- Whether Mars could have ever supported life.
- Role of water on Mars
- Climate and geology of Mars

Curiosity rover navigates surface of Mars autonomously.

Source: http://www.jpl.nasa.gov/news/news.php?release=2013-259

Examples Of Weak AI

Expert System

- Computer system that emulates decision making ability of a human expert.

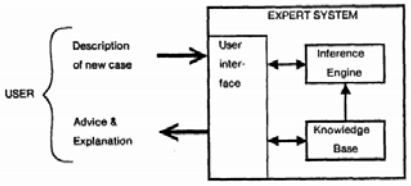
Two components:

- Knowledge base – repository of information/facts about the world as well as rules that can be applied to the facts. Rules usually have an IF-THEN representation.
- Inference engine – applies rules to known facts to deduce new knowledge.

Sources: https://en.wikipedia.org/wiki/Expert_system

MYCIN

- Mycin is an example of an early expert system.
- Initially designed to diagnose bacterial infections.
- List of possible bacterial culprits provided, ranked from high to low based on the probability of each diagnosis.
- Antibiotic treatment regimen, dose adjusted for patient's body weight, was also given.



Sources:
https://en.wikipedia.org/wiki/Mycin
http://people.dbmi.columbia.edu/~ehs7001/Buchanan-Shortliffe-1984/Chapter-01.pdf

Representing Problems As Symbols

- AI programs reduce problems to symbols.
- Problems are solved through the manipulation of these symbols.
- The manipulation of these symbols can seem intelligent.
- The computer does not "know' what the symbols mean.

Example

Scenario:
- A farmer needs to cross a river by boat taking with him his dog, goose, and a sack of corn.

Constraints:
- The boat is small and can only hold one item along with the farmer.
- The dog can't be left alone with the goose. The dog will eat the goose.
- The goose can't be left alone with the corn. The goose will eat the corn.

Problem:
- What is the order in which the farmer transfers his property across the river?

Symbolic Representation

Dog = d

Goose = g

Corn = c

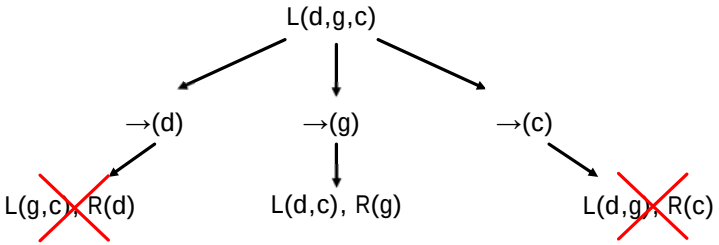
At the start of the problem, all three are on the left bank of the river. The right bank is empty.
- Start state: L(d,g,c), R()

The goal is to get all three across to the right bank:
- Goal state: L(), R(d,g,c)

Operators are used to indicate actions the farmer can take:
- Row dog to right bank = →(d)
- Row corn to left bank = ←(c)

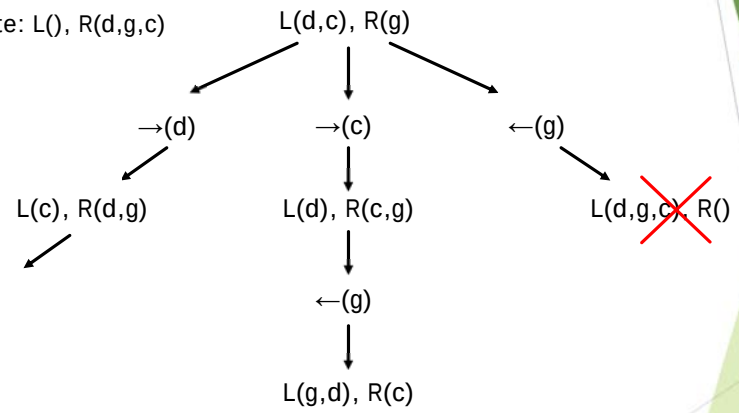
State Space Search

Start state: L(d,g,c), R()
Goal state: L(), R(d,g,c)



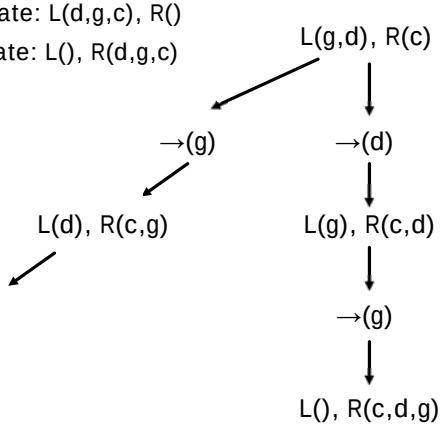
State Space Search

Start state: L(d,g,c), R()
Goal state: L(), R(d,g,c)



State Space Search

Start state: L(d,g,c), R()
Goal state: L(), R(d,g,c)



Problem solution

Start state: L(d,g,c), R()

Goal state: L(), R(d,g,c)

Solution: →(g) →(c) ←(g) →(d) →(g)

Summary

Artificial intelligence is the *computational study of structures and processes that support intelligent behaviour*.

Two philosophical views of intelligence:

- Behaviourist/functionalist and cognitive.

Strong AI versus Weak AI.

- The study of Weak AI has produced many useful applications.

Emphasizes symbolic representations of problems