COMPSCE 111 / 111G

An Introduction to Practical Computing

Artificial Intelligence
AI

Research Areas & Motivations
What is Artificial Intelligence?

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

Term first coined in 1956 at the Dartmouth Summer Research Project on Artificial Intelligence

Many of the world's leaders in 20th C AI were at this meeting.

(John McCarthy, Marvin Minksy, Alan Newell, Herbert Simon, Claude Shannon)
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

Depending on who you talk to, 'intelligence' can also include "bodily skills."

**The execution of motor skills.** e.g. bipedal walking, playing tennis.

**Navigation.** e.g. using sensors (eyes, cameras) to navigate -- e.g. bee and ant navigation.
Areas of AI research

**Computer vision** -- e.g. taking a video or still image and identifying objects, or optical character recognition (reading printed text)

**Natural language processing** -- getting computers to understand natural (rather than formal) languages (written or spoken)

**Robotics** -- e.g. building robots that can **navigate on Mars** without human help, that can **clean the house** for you (e.g. Roomba), that can go into dangerous places for **search and rescue** etc. etc.

**Knowledge-based systems** -- using a database of knowledge to try to emulate / replace a human expert.

**Machine learning** -- e.g. training neural networks to be able to correctly classify images, or predict trends
The goal(s) of AI research

Three interrelated aims:

- engineering aim
- psychological aim
- philosophical aim
Engineering Aim

**Goal:** to build useful "intelligent" artefacts.

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
- etc.
Psychological Aim

Goal: To gain insight into **natural forms of intelligence** (by building **artificial** forms of intelligence).

Typically the target is human intelligence..

..but increasingly, it is becoming clear that there is also a great deal that can be learned from studying animal minds/brains...

...or even the adaptive behaviour of slugs, worms and insects!
Philosophical Aim

To create computational principles, theories or systems that provide a greater insight into the fundamentals of cognition.

To approach questions such as...
What is intelligence?

How can we make something that is conscious?

What is consciousness anyway?

How can we make artificial things that have their own desires?

What is a desire anyway?
AI

Strong AI vs. Weak AI
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 cannot become self-aware and reason.

But they can be used to solve specific problems in a well-defined domain.
Examples of Strong AI
Examples of Weak AI
Mars Rovers

**Curiosity** and **Opportunity**
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

Launched 2011. Designed to operate for a few months. Curiosity is still being used!

Lag of communication between Earth and Mars means that the rovers have to navigate semi-autonomously.
Expert System

Computer system that emulates decision making ability of a human expert. Two components:

**Knowledge base** – a 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.

**MYCIN** -- an early expert system that was designed to diagnose bacterial infections.

Asked physician a series of yes/no questions, and used responses to:
- generate a list of possible bacterial culprits, ranked from high to low based on the probability of each diagnosis
- propose an antibiotic treatment regimen, dose adjusted for patient’s body weight

MYCIN was never actually used in practice but research indicated that it proposed an acceptable therapy in about 69% of cases, which was better than the performance of infectious disease experts who were judged using the same criteria.

Interesting ethical and legal issues to be considered here. What if the prognosis is wrong? Who is responsible? What is a good enough success rate to use this kind of tool?

From https://en.wikipedia.org/wiki/Mycin
IBM Deep Blue

Chess playing computer that in 1997 famously won a match against the World Champion, Garry Kasparov …

The 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.

Kasparov was not permitted to study Deep Blue’s previous games.
Symbolic AI

Many AI programs reduce problems to symbols, and 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 (more on this in a bit).

Let's see an example of a puzzle that has been transformed into symbols.
Puzzle
A farmer needs to cross a river by boat taking with him his dog, goose, and a sack of corn.

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?
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)

Some states must be avoided.

L(d,g), R(c);  L(c), R(d,g);
L(g,c), R(d); ..etc.

Operators are used to indicate actions the farmer can take:

Row dog to right bank = \( \rightarrow(d) \)

Row corn to left bank = \( \leftarrow(c) \)
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)

L(d,c), R(g) →(d) →(c) ←(g)

L(c), R(d,g) →(d)
L(d), R(c,g) →(c)
L(g,d), R(c) ←(g)
L(d,g,c), R() ←(g)

back to where we started!
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)

Computers are fast, but some games branch really fast, and the number of possible moves is too much even for supercomputers.

Our brains are much slower than supercomputers. Yet we still (can) play chess quite well! Do we work the same way?

Some famous arguments concerning Strong AI
Two contrasting philosophical views...

When deciding if something is intelligent...

**Behaviourist/Functionalist Approach**
The external behaviour is all that matters!

If it behaves intelligently, then it is intelligent.

What is going on "behind the scenes" doesn't matter. (Could be very simple).

The Turing Test is a famous method for evaluating intelligence that embraces the functionalist approach (more in a moment).

**Cognitivist Approach**
What happens internally matters!

*We must consider how it thinks, not just look at the behaviour.*

The Chinese Room argument is a famous attack on the functionalist approach. It argues that...

1. It isn't enough to just look at behaviour to detect consciousness.
2. Computationalist approaches to AI will never produce human-like intelligence.
The Turing Test

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

Turing says that the question "Can machines think?" is ambiguous. And proposes a method (a game) that is less ambiguous, but still gets at what is meant by the question.
Imitation game

Three players:
- Player A is a man
- Player B is a woman
- The interrogator

The interrogator is unable to see or hear the other players. All s/he can do is ask them questions (in a way that does not give their sex away, e.g. by receiving answers written on a piece of paper).

A and B both try to convince the interrogator that they are the woman.

The interrogator tries to determine which player is the man and which is the woman.
The Standard Turing Test

Three players:
- Player A is a **machine/computer**
- Player B is a **human**
- The interrogator (a human)

The interrogator questions A and B (in a way that doesn't make it obvious which is human.)

A and B both try to convince the interrogator that they are **human**.

The interrogator tries to figure out which player is the human and which is the machine.

**If the interrogator fails, then the machine is intelligent!**
The Chinese Room...

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

Goal is to refute the Strong AI Claim

"a computer that produces the same outputs as a human has a mind just like the human's mind."

Strong AI Claim (paraphrased from Searle)

Which also argues strongly against the ability of behaviour-based tests* to detect "intelligence" / minds / consciousness.

* like the Turing Test
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.

If you see this shape, "什麼"
followed by this shape, "帶來"
followed by this shape, "快樂"
then produce this shape, "為天"
followed by this shape, "下式".

http://www.mind.ilstu.edu/curriculum/searle_chinese_room/searle_chinese_room.php
The Chinese Room...

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.

http://www.mind.ilstu.edu/curriculum/searle_chinese_room/searle_chinese_room.php
Summary

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

We discussed:

- 5 Different Research Areas
- 3 Motivations behind AI research

Strong AI versus Weak AI
The study of Weak AI has produced many useful applications.

We mentioned behaviourist and cognitive approaches to understanding intelligence.

Behaviourist/functionalist: behaviour is all that matters

Cognitive: how the behaviour is being generated is important.

The most common approach in AI...
- emphasizes **symbolic representations** of problems
- and uses **search** to find the best possible action to take