

Introduction to AI

COMP2002

Lauren Ansell

Introduction

Today's topics:

- Module introduction
- Introduction to artificial intelligence

Session learning outcomes - by the end of today's lecture you will be able to:

- Describe the principal tasks in AI.
- Give examples of applications of AI.
- Summarise ethical considerations arising from the use of AI.

Aims

To familiarize students with the **underlying principles of AI**, expose them to topics such as **optimization, machine learning, knowledge representation and reasoning** and **natural language processing**, and to instill an appreciation of the importance of **ethical and social considerations** behind the use of AI

Learning outcomes

At the end of the module the learner will be expected to:

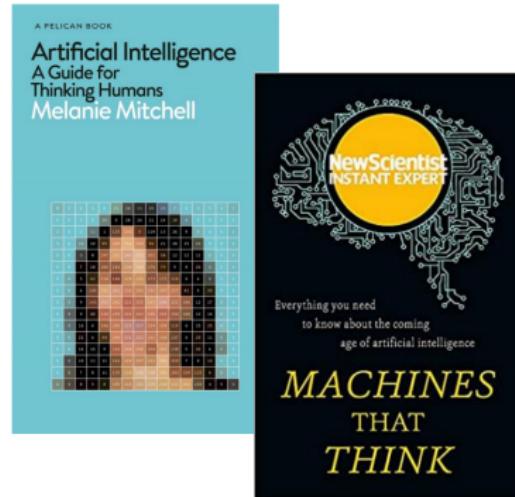
- ① Describe and analyse a range of **AI methods** and their **applications**.
- ② Compare AI paradigms and **evaluate the appropriateness** of a particular paradigm for **specific application domains**.
- ③ Choose and apply **appropriate** AI methods to a **chosen sample domain**.

Suggested Reading And Support Material

Artificial Intelligence –
A Guide for Thinking Humans
Melanie
Mitchell, Pelican Books, 2019

Machines that Think
New Scientist, 2017

Additional
suggested/recommended
reading that will help you
improve your understanding
of the material,



GitHub repository:
<https://github.com/laurenansell/COMP2002>

Schedule

Week 1: Introduction to AI

Week 2: Machine Learning

Week 3: Neural networks

Week 4: Deep Learning

Week 5: Evolutionary Computation

Week 6: Multi-objective Optimisation

Week 7: Knowledge Representation

Week 8: Agents

Week 9: Natural Language Processing

Week 10: Reinforcement Learning

Week 11: Evaluating AI and XAI

What Is “Artificial Intelligence”?

“Artificial intelligence is the study of how to make computers do things which, at the moment, humans are better” - Elaine Rich, 1983

The story so far...

1943: McCulloch and Pitts propose basis of neural networks

1950: The Turing test is proposed

1955: Computer Scientist John McCarthy coins the term “AI”

1964: MIT-developed chatbot ELIZA holds its first conversations

What Is “Artificial Intelligence”?

1997: IBM’s Deep Blue defeats Garry Kasparov

2011: Siri launched and IBM Watson wins “Jeopardy”

2014: Eugene Goostman “passes” the Turing Test and Alexa launched

2017: AlphaGo beats world champion Ka Jie



Applications Of AI



NETFLIX



facebook



Different Types Of Computers

The human brain

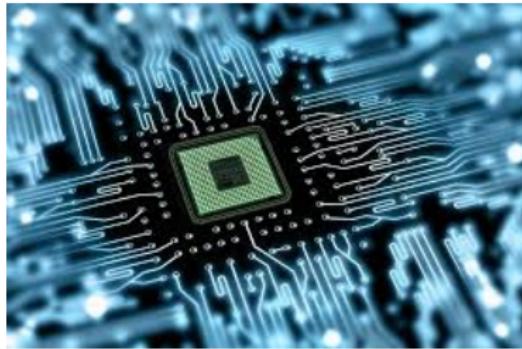
- Receives inputs from our senses
- Processes this information and makes decisions based on it
- Learns from experience
- Powerful memory



Different Types Of Computers

Computers

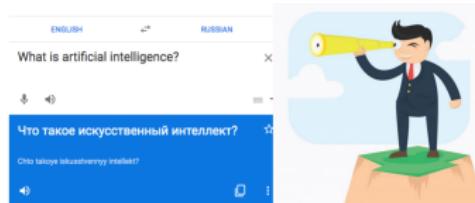
- Use programs to process data in a pre-determined way
- They can use data from a variety of inputs – cameras, microphones, the internet
- They cannot think for themselves
- **How can we make computers “think”?**



An Example: Natural Language Processing

“John saw Mary with a telescope” could have multiple meanings:

- John saw Mary through a telescope
- John saw Mary carrying a telescope



Sentiment

- The true meaning behind our words – difficult for a computer to understand
- e.g., jokes and sarcasm

Another Example: Computer Vision

- The brain uses **information from your eyes** to process visual information
- **Identify** and **track** objects
- Our task is to build software that takes the **information from cameras** and performs the same task using **machine learning**

Applications of computer vision

- Image analysis - e.g., healthcare
- Autonomous vehicles – navigation
- Fraud detection (e.g., face recognition)

Another Example: Self-driving Cars

Cars that can get from A to B with **no human control**

Need to be able to:

- Navigate
- Understand their environment – spot hazards, recognize and understand traffic signs...

Could be **safer**

Benefit society – more transport options for **children and the elderly**

How Do We Know If We've Achieved AI?

The Turing Test

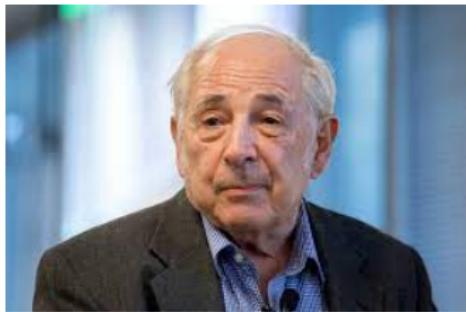
- A person has a conversation with an **AI and another human**
- They pass the same message to each entity **without knowing which is which (human and AI)**
- They must judge which is which **based on their responses alone**



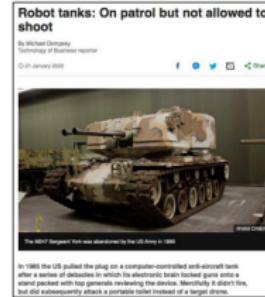
How Do We Know If We've Achieved AI?

Searle's Chinese Room

- An AI is said to **understand Chinese** – it passes the Turing Test
- Does the AI really **understand** Chinese or has it simulated understanding?
- A person could follow rules and generate responses to phrases **without understanding**



Ethics In AI



- Ethical discussions occur around any new technology
- Decision making vs support
- Challenging to define AI itself – so hard to define ethical issues
- Who is responsible if an AI harms someone?
- Can AI discriminate?

Epistemic strategies

- Reduce the **uncertainty** around AI – e.g. monitor likely future developments
- Transparency or **explainable** AI

Technical strategies

- Make AI **safer** or **error free**
- Align AI to **human values** and ensure **human control**

Moral strategies

- Allay fears that there are ethical issues with AI

The State Of The Union: Deep Fakes



Deep fakes use deep learning techniques to construct novel footage of a person talking
Barack Obama voices an opinion about Donald Trump in a convincing video constructed using deep learning

Large Language Models - ChatGPT

Support the Guardian
Available for everyone, funded by readers
[Support us →](#)

[Print subscriptions](#) [Sign in](#) [Search jobs](#) [Search](#) [UK edit](#)

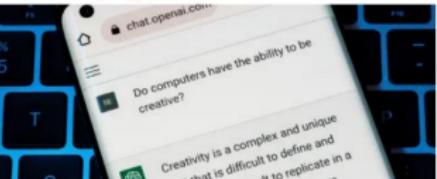
News **Opinion** **Sport** **Culture** **Lifestyle** **More** ▾

Education ▶ Schools Teachers Universities Students

Chatbots

AI-assisted plagiarism? ChatGPT bot says it has an answer for that

Silicon Valley firm insists its new text generator, which writes human-sounding essays, can overcome fears over cheating



Alex Hern
[@alexhern](#)
Sat 31 Dec 2022 09.00 GMT

[Home](#) | [Cost of Living](#) | [War in Ukraine](#) | [Coronavirus](#) | [Climate](#) | [UK](#) | [World](#) | [Business](#) | [Politics](#)

Technology

ChatGPT: New AI chatbot has everyone talking to it

7 December 2022



GOTHAM

By Chris Vallance

Technology reporter

A new chatbot has passed one million users in less than a week, the project behind it says.

ChatGPT was publicly released on Wednesday by OpenAI, an artificial intelligence research firm whose founders included Elon Musk.

But the company warns it can produce problematic answers and exhibit biased behaviour.

Open AI says it's 'eager to collect user feedback to aid our ongoing work to

Red Flags

People could be warned that they are interacting with AI – **red flags**

AI should be **explainable** so that it can be understood by non-expert humans



Clever Hans – an unexplainable black box

“Clever Hans” was
a **horse that could count**
– scoring high accuracy

He was
deriving the answer **from the**
person asking the question

AI can reach the right answer for the wrong reasons
e.g., classify boat images because of water, huskies by the presence of snow...



“Explainable” AI



A classifier outputs “beagle” given the left-hand image
A second system produces the right-hand image to illustrate
which regions of the image led to that classification

Explainable AI improves **trust and verifiability**, enables new **insight into algorithm operation** and is necessary to **support legislation** (e.g., GDPR)

Uses explanations to...

- ① Explain how the AI has produced a model
- ② Explain individual predictions
- ③ Explain model behaviour
- ④ Explain with representative examples

Is transparency always helpful?

- A provider may use an explanation to instill a false sense of security
- IP concerns and “gaming the system”
- Privacy issues
- Is transparency actually needed? Or can we rely on safety tests?

Artificial intelligence

- Has many (many) applications in everyday life
- Though we're familiar with its use there are many (many) open problems

Ethics and society

- Considerable ethical considerations from the use of AI
- Warnings and explainable AI can help reduce potentially harmful societal impact