

# COMP9414: Artificial Intelligence

## Lecture 1a: Foundations

Wayne Wobcke

e-mail:w.wobcke@unsw.edu.au

## About Me

- Logic and Natural Language Processing (1985–1989)
- Logic and Knowledge Representation (1989–1995)
- Intelligent Agent Theory (1996–2007)
- Personal Assistant Applications
  - Intelligent Desktop Assistant (1998–2000)
  - Smart Personal Assistant, like Siri (2002–2006)
  - Clinical Handover Assistant (2003–2007)
- Agent-Based Modelling (2008–2013)
- Recommender Systems (2008–2014)
- Data Science/Computational Social Science (2015–)
  - Text Stream Mining for Event Extraction (2015–2019)
  - Topic Modelling for Political Sentiment Analysis (2015–2019)
  - Machine Learning for Official Statistics (2020–)

## This Lecture

- What is Artificial Intelligence?
- Arguments Against the Possibility of AI
- Analyst Assistant for Event Extraction

## What is Artificial Intelligence?

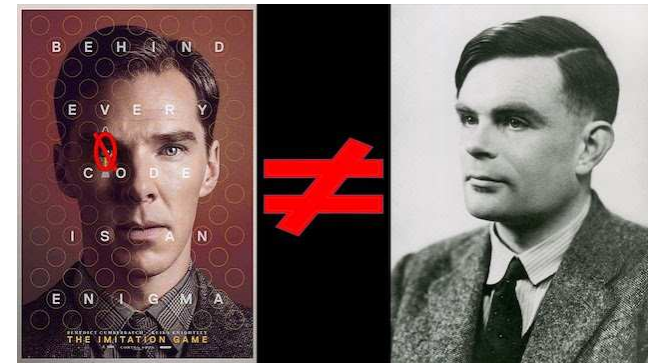
|   |  |
|---|--|
| <b>Thinking Humanly</b><br>“The exciting new effort to make computers think . . . machines with minds, in the full and literal sense.” (Haugeland, 1985)<br>“[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . .” (Bellman, 1978) | <b>Thinking Rationally</b><br>“The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985)<br>“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992) |
| <b>Acting Humanly</b><br>“The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990)<br>“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)   | <b>Acting Rationally</b><br>“Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998)<br>“AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)                           |

## What is Artificial Intelligence?

- Science
  - ▶ **Hypothesis:** Intelligence **is** Computation
  - ▶ Could be wrong!
- Engineering
  - ▶ Techniques for designing/building “intelligent” systems
  - ▶ Emphasis on what “works” (for what purposes)

Analogy between bird flight and aerodynamics

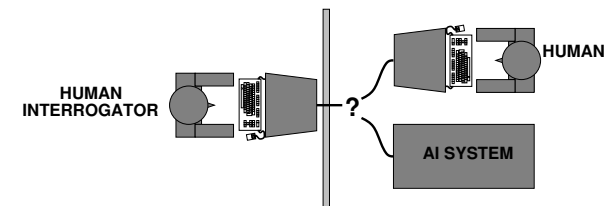
## The Imitation Game



## What is Computation?

- **Church-Turing Thesis:** A function is **computable** if it can be defined (“computed”) using standard programming languages (such as Python), i.e. that contain assignment, conditionals, sequencing and iteration [and this yields the same answer regardless of the language]
- What **isn't** computable (that also might produce “intelligence”)?
  - ▶ Brain processes (face/speech recognition?, thought??)
  - ▶ Evolutionary processes?
  - ▶ Complex adaptive systems (ant colonies, flocking birds)?

## Turing Test



- Aim is for the computer to **imitate** a human well enough to **fool** a human interrogator
- Turing’s own conclusion (1950)
  - ▶ Question “can a machine think” is meaningless
  - ▶ In future, speak of machines “thinking” without contradiction

## Eliza

Men are all alike. (user)

In what way? (computer)

They're always bugging us about something or other.

Can you think of a specific example?

Well, my boyfriend made me come here.

Your boyfriend made you come here?

He says I'm depressed much of the time.

I am sorry to hear you are depressed.

It's true. I am unhappy.

Do you think coming here will help you not to be unhappy?

I need some help.

What would it mean to you if you got some help?

Perhaps I could learn to get along with my mother.

Tell me more about your family.

## SHRDLU

Pick up a big red block. (user)

OK. (computer)

Grasp the pyramid.

I don't understand which pyramid you mean.

Find a block which is taller than the one you are holding and put it into the box.

By *it*, I assume you mean the block which is taller than the one I am holding. OK.

What does the box contain?

The blue pyramid and the blue block.

What is the pyramid supported by?

The box.

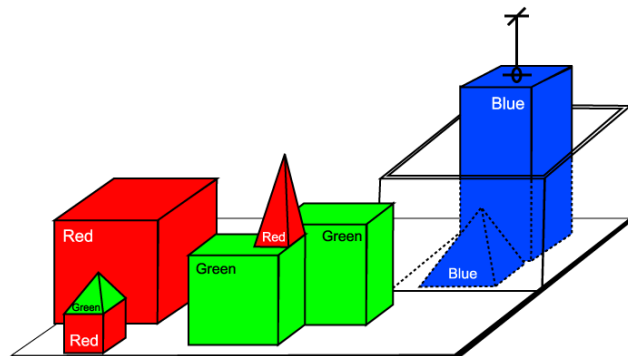
How many blocks are not in the box?

Four of them.

Is at least one of them narrower than the one which I told you to pick up?

Yes, the red cube.

## SHRDLU



## Physical Symbol System Hypothesis

- A **symbol** is meaningful physical pattern that can be manipulated
- A **symbol system** combines and manipulates symbols

Claim: A physical symbol system has the necessary and sufficient means for general intelligent action

Semantics?: An expression designates an object if, given the expression, the system can either affect the object itself or behave in ways dependent on the object” (Newell and Simon 1976)

## Knowledge Level and Symbol Level

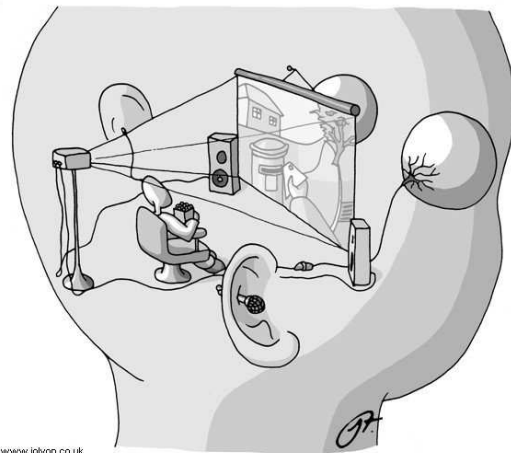
- The **knowledge level** is in terms of what an agent knows and what its goals are (**external** theory of the agent)
- The **symbol level** is a level of description in terms of what reasoning the agent is doing (**internal** description of the agent's operation)

External vs internal, explicit vs implicit (see the frog in Lecture 1b)

## A Scientific Approach – Key Ideas

- Logic (Aristotle c. 350BC, Boole 1848, Frege 1879, Tarski 1935)
- Formal algorithms (Euclid c. 300BC)
- Probability theory (Pascal 17th C, Bayes 18th C)
- Utility theory (Mill 1863)
- Dynamical systems (Poincare 1892)
- Structural linguistics (Saussure 1916, Bloomfield 1933)
- Formal systems (Gödel 1929, Turing 1936)
- Neural networks (McCulloch & Pitts 1943)
- Game theory (von Neumann & Morgenstern 1947)
- Cybernetics/Control theory (Wiener 1948)
- Decision theory (Bellman 1957)
- Formal linguistics (Chomsky 1957)

## Problem: Avoid Homunculus (Little Man)



## Related Disciplines

- Philosophy
  - Mind-Body Problem
  - Nature of Knowledge
  - Nature of Scientific Claims
- Psychology (Cognitive)
  - Results only about very specific models
  - Replicability crisis
- Linguistics
  - Formal grammar (after Chomsky)?
  - Computational Linguistics = Statistical Machine Learning?
- Neuroscience
  - Brain function vs structure from medical imaging
  - 100 billion neurons with up to 10,000 connections

## Science Fiction (not Science)

- Greek Mythology (Pygmalion, Talos)
- 1580 Rabbi Loew (Golem, a clay man brought to life)
- 1818 Mary Shelley (Frankenstein)
- 1883 Carlo Collodi (Pinocchio)
- 1920 Karel Capek (Rossum's Universal Robots)
- 1950 Isaac Asimov (Three Laws of Robotics)
- 1951 Osamu Tezuka (Astro Boy)

## Arguments Against AI

- Misplaced emphasis on abstract reasoning rather than low-level perception and behaviour
  - ▶ “Intelligence Without Reason” (Brooks 1991)
- General intelligence vs specific modules
  - ▶ “How the Mind Works” (Pinker 1997)
- Philosophical objections to AI
  - ▶ Gödel's theorem, undecidability (Lucas 1961, Penrose 1989)
  - ▶ Chinese Room (Searle 1980)
  - ▶ “What Computers (Still) Can't Do” (Dreyfus 1972, 1993)

## Robots – Good or Evil?



## Arguments Against AI – The X Factor

A computer can't be intelligent because it can't . . .

- be creative, generate new insights
- produce poetry, a symphony, a work of art, etc.
- beat the world champion of Chess or Go
- make mistakes
- have emotions, empathy
- be conscious, have free will, possess vital spirit, a soul
- have experiences (qualia), e.g. the taste of ice-cream
- make ethical judgements

## Arguments Against AI – Weak vs Strong AI

“Minds, Brains, and Programs” (Searle 1980)

- **Weak AI:** The claim that computers can be made to act **as if** they are intelligent, providing a tool to study the mind, but only **simulate** intelligence
- **Strong AI:** The claim that machines acting intelligently exhibit **genuine** intelligence: have a mind, have cognitive states, understand language

Does this assume an X factor – intentionality, “aboutness”?

## Arguments Against AI – The Chinese Room

Suppose we have a room with a person who “implements” a program whose data is stored on paper, etc., that can reliably pass the Turing Test conducted in Chinese.

- The human doesn’t understand Chinese
- The human is analogous to a computer program
- So computer programs don’t understand language

But **where** is the understanding in the brain?

## Arguments Against AI – Incompleteness

Gödel showed that any formal system (powerful enough to encode sentences as objects) is **incomplete**, i.e. there is a sentence *G* that cannot be proven that is nevertheless true (if the system is consistent)

- But people can “easily see” the truth of *G* (Lucas 1961)
- *G* is a sentence like “*G* cannot be proven”
- Like the Liar Paradox
  - ▶ “This sentence is false” . . . which can’t be true, or false

So what?

## Arguments Against AI – Brain Prosthesis

Suppose we replace each neuron in a brain by a functionally equivalent electronic device.

- The behaviour of the system is unchanged
- But the consciousness gradually disappears (Searle 1992)
- Or the result is a conscious machine (Moravec 1988)

Is consciousness required for intelligence?

Is consciousness causally connected to behaviour?

Can consciousness be studied scientifically?

Does this even matter for AI?

## Dose of Reality – State of the Practice

- Which of the following can be done at present?
  - ▶ Play a decent game of table tennis (ping-pong)
  - ▶ Drive in the centre of Cairo, Egypt
  - ▶ Drive along a curving mountain road
  - ▶ Play games like Chess, Go, Bridge, Poker
  - ▶ Discover and prove a new mathematical theorem
  - ▶ Write an intentionally funny story
  - ▶ Give competent legal advice in a specialized area of law
  - ▶ Translate spoken English into Swedish (or Chinese) in real time

## Recent Advances

- Industrial Robots
  - ▶ Loading/unloading ships
  - ▶ Warehouse order fulfilment
  - ▶ Self-driving cars
- Deep Learning/Hybrid Models
  - ▶ Image classification
  - ▶ Game playing
  - ▶ Large Language Models (LLMs)

## Chess, Vision – Easy or Hard?



## Analyst Assistant

- Technologies and Techniques
  - ▶ Pipeline architecture (Software Engineering)
  - ▶ “Big data” streaming platform and storage (Databases)
  - ▶ Sentence segmentation (NLP)
  - ▶ Part of speech tagging (Reasoning with Uncertainty)
  - ▶ Sentence parsing (NLP)
  - ▶ Defining and reasoning with domain-specific ontology (KBS)
  - ▶ Rule Induction/Generalization (Machine Learning)
  - ▶ Event ranking (Machine Learning)

**Challenge:** Do all this in one go using an LLM



Analyst Assistant – Demo 1

Foreign tourists attacked in western Afghanistan

Governor of Herat province says eight Britons, three Americans and one German among group ambushed by Taliban

A group of western tourists been ambushed by Taliban least six people wounded.

The foreigners were travell province of Herat with an A convoy. The group were he centre, close to the Iranian

A spokesman for Herat's go US citizens and one Germai Scotland.

Analyst Assistant – Demo 3

A group of western tourists on a sightseeing adventure tour of Afghanistan have been ambushed by Taliban gunmen, Afghan officials said on Thursday, with at least six people wounded.

| New event                     | Link to | Save    | Delete  |
|-------------------------------|---------|---------|---------|
| +Actor                        | Action  | +Target | +Entity |
| A by Taliban gunmen           |         |         | ✗       |
| V ambushed                    |         |         | ✗       |
| T A group of western tourists |         |         | ✗       |
| E by Taliban gunmen           |         |         | ✗       |

Summary: by Taliban gunmen ambushed A group of western tourists

Publication date: Thu Aug 04 22:18:35 AEST 2016

Actors:

by Taliban gunmen (Entity.Organization.Taliban)

Action: ambushed

Targets:

A group of western tourists

Entities:

by Taliban gunmen (Entity.Organization.Taliban)

Event linked to: 35

Parsed sentence: (TOP (S (S (NP (NP (NP (DT A) (NN group)) (PP (IN of) (NP (JJ western) (NNS touris ts)))) (PP (IN on) (NP (NP (DT a) (NN sightseeing) (NN adventure) (NN tour)) (PP (IN of) (NP (NNP Afgh anistan)))))) (VP (VBP have) (VP (VBN been) (VP (VBN ambushed) (PP (IN by) (NP (NNP Taliban) (NNS gunmen)))))) (, .) (NP (NNP Afghan) (NNS officials)) (VP (VBD said) (PP (IN on) (NP (NNP Thursday))) (, .) (PP (IN with) (S (NP (QP (IN at) (JJ least) (CD six)) (NNS people)) (VP (VBN wounded)))) (, .) Rule fired: (S (NP (NP <S1>)) (VP (VP (VBN <V>) (PP <A1>))))

Analyst Assistant – Demo 2

<< Foreign tourists attacked in western Afghanistan | World news (17259)

0 A group of western tourists on a sightseeing adventure tour of Afghanistan have been ambushed by said on Thursday, with at least six people wounded.

1 The foreigners were travelling in the Chesht-e-Sharif district in the western province of Herat with e militants opened fire on the convoy.

2 The group were heading for Herat city, the country's ancient cultural centre, close to the Iranian bo

3 A spokesman for Herat's governor said the tourists included eight Britons, three US citizens and one

4 Two of the British citizens were believed to be from Scotland.

5 A military spokesman said the attack left at least five foreign tourists and their Afghan driver wound

6 Sources with knowledge of the local insurgency said the Taliban's shadow governor in Chesht-e Shz to be behind it.

7 The attack reportedly ended at about 1.30pm local time.

8 About two hours later, the group arrived at a medical clinic in Obeh district, west of the ambush sit

9 Reportedly, none of the injuries were serious.

10 According to sources in Kabul the party had gone to Afghanistan with Hinterland Travel, a small Bi Brighouse, West Yorkshire, and run by veteran tour guide Geoff Hann.

11 The company specialises in adventure tours to war-torn countries such as Iraq and Afghanistan.

12 It has defended travelling to Afghanistan despite foreign office advice to avoid it.

13 Promotional material describes Afghanistan as at the "cutting edge of adventure tourism, which in progress".

14 The firm's latest 21-day tour, advertised on its website, began on 26 July.

Analyst Assistant – Demo 4

Add Search Edit Repeat Del Save

Event

- Action
  - Abduct
  - Airstrike
  - Alliance
- Announce
  - Financial Reform
- Appeal
  - Appeal for aid
    - Appeal for economi
    - Appeal for humanit
    - Appeal for military
  - Appeal for military
- Arrest
- Arson
- Assassinate
- Assault
- Attack
  - Suicide
  - Behead
  - Bomb
  - Border Close
  - Border Open
  - Capture
  - Claim responsibility
- Clash
  - Border Clash

Movement of the Taliban in Pakistan

- aka: TTP; Tehrik-i-Taliban; Tehrik-e-Taliban; Pakistani Taliban

NATO

- Resolute Support Mission
  - aka: North Atlantic Treaty Organization
- National Directorate of Security

Pakistan rebel group

- United Baloch Army
  - aka: UBA
- Quadrilateral Coordination Group

TTP group Jamaat-ul-Ahrar

- aka: Jamaat-ul-Ahrar

TTP group Mujahideen Special Group

- aka: Mujahideen Special Group

Taliban

- aka: Islamic Emirate of Afghanistan; Taleban



## Course Schedule

|   |   |
|---|---|
| 1 | Artificial Intelligence and Agents              |
| 2 | Problem Solving and Search                      |
| 3 | Constraint Satisfaction Problems                |
| 4 | Logic and Knowledge Representation              |
| 5 | Reasoning with Uncertainty                      |
| 6 | Neural Networks and Deep Learning               |
| 7 | Machine Learning                                |
| 8 | Natural Language Processing Toolkits            |
| 9 | Neural Networks for Natural Language Processing |

## Changes for 2024 Term 3

- Lectures
  - ▶ Retain emphasis on solid foundations (Wayne Wobcke)
  - ▶ Revise material on Machine Learning (Gustavo Batista)
  - ▶ New approach to Natural Language Processing (Aditya Joshi)
- Tutorials
  - ▶ 2 hour tutorials, more emphasis on AI Python programming
  - ▶ Written exercises (exam) and Jupyter notebooks (assignments)
  - ▶ **Do not use ChatGPT for written exercises!**
- Assessment
  - ▶ Assignment 1: Less focus on programming (25%)
  - ▶ Exam: 50% rather than 60% – **Hurdle 40%**

## What is Not Covered

- Robotics
- Motion Planning
- Computer Vision
- Statistical Learning
- Reinforcement Learning
- Game Playing
- Evolutionary Computing
- Multi-Agent Systems
- Recommender Systems

## Summary: State of the Art

- Engineering
  - ▶ Many recent advances on subproblems
  - ▶ Mostly derived using large data sets
  - ▶ Models use human expertise and learning
  - ▶ Trend towards complex software systems
  - ▶ Trend towards (proprietary) **very** large neural networks
- Science
  - ▶ Nowhere close to any general theory of intelligence
  - ▶ Nowhere close to human reasoning in many domains
  - ▶ Nowhere close to understanding the brain