COMP9414: Artificial Intelligence

Lecture 1a: Foundations

Wayne Wobcke

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

About Me

- AI paper on Neural Networks (1986)
- Logic and Natural Language Processing (1985–88)
- Logic and Knowledge Representation (1989–97)
- Intelligent Agents Theory (2002–15)
- Personal Assistant Applications
  - Intelligent Desktop Assistant (1998–2000)
  - E-Mail Management Assistant (2002)
  - Smart Personal Assistant, like Siri (2003–2006)
  - Clinical Handover Assistant (2003–2008)
- Recommender Systems (2008–14)
- Text Mining for Event Extraction (2014–19)
- Topic Modelling for Political Sentiment Analysis (2015–19)
- Data Science (2015–)

This Lecture

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

What is Artificial Intelligence?

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

What is Artificial Intelligence?

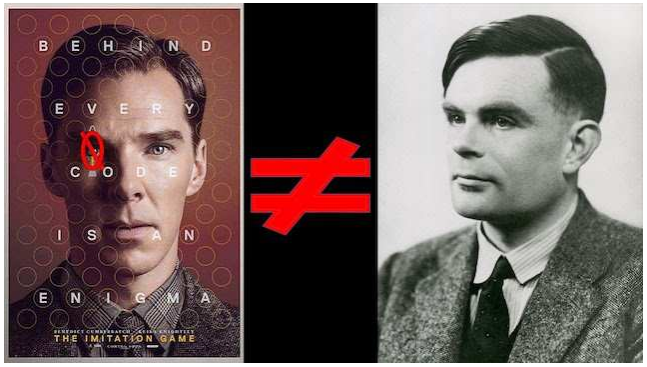
- Science
  - Hypothesis: Intelligence **is** Computation
  - Could be wrong!
- Engineering
  - Techniques for designing/building "intelligent" systems
  - Emphasis on what "works" (for what purposes)

  Analogy between bird flight and aerodynamics

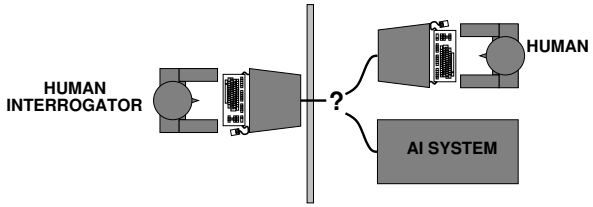
What is Computation?

- 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)?

The Imitation Game



Turing Test

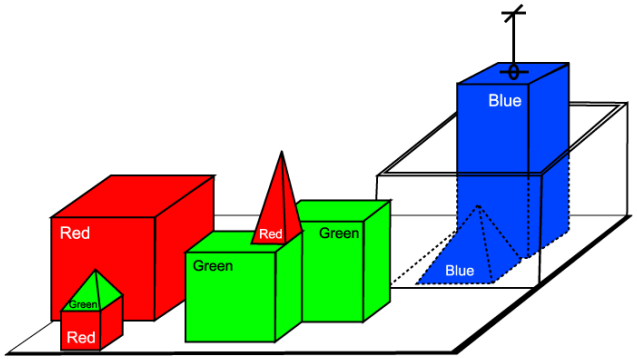


- Aim is for the computer to imitate a human well enough to fool a human interrogator
- Turing's own conclusion (1950)
  - Question "can a machine think" is meaningless
  - In future, speak of machines "thinking" without contradiction

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



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.

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 (McCullogh & Pitts 1943)
- Cybernetics/Control theory (Wiener 1948)
- Game theory (von Neumann & Morgernstern 1947)
- Decision theory (Bellman 1957)
- Formal linguistics (Chomsky 1957)

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)

Robots – Good or Evil?

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)

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

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 – 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 – 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 – 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

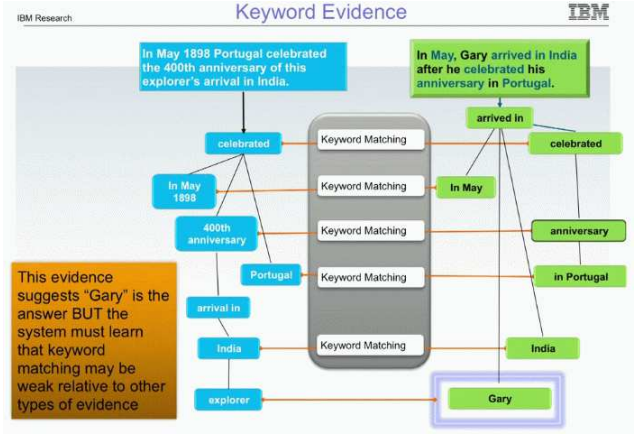
- Which of the following can be done at present?
  - Play a decent game of table tennis (ping-pong)
  - Drive in the centre of Cairo, Egypt
  - Drive along a curving mountain road
  - Play games like Chess, Go, Bridge, Poker
  - Discover and prove a new mathematical theorem
  - Write an intentionally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

Chess, Vision – Easy or Hard?

Recent Advances

- Industrial Robots
  - Loading/unloading ships
  - Warehouse order fulfilment
  - Self-driving cars
- Deep Learning/Hybrid Models
  - Image classification
  - Language processing
  - Game playing

IBM Watson DeepQA



YouTube: "Building Watson - A Brief Overview of the DeepQA Project"

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)

Course Schedule

1	Artificial Intelligence and Agents
3	Problem Solving and Search
4	Constraint Satisfaction Problems
5	Logic and Knowledge Representation
6	Reasoning with Uncertainty
7	Machine Learning
8	Natural Language Processing
9	Knowledge Based Systems
10	Neural Networks and Reinforcement Learning
11	Review

What is Not Covered

- Robotics
- Computer Vision
- Statistical Learning
- Deep Learning
- Game Playing
- Evolutionary Algorithms
- Multi-Agent Systems
- Recommender Systems
- Spoken Dialogue 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 is towards complex software systems
- Science
  - Nowhere close to any general theory of intelligence
  - Nowhere close to human reasoning in many domains
  - Nowhere close to understanding the brain