

Artificial intelligence: An introduction

- Russell and Norvig, Chapters 1, 2, + 26

Overview: aims

- know a couple of definitions of intelligence and AI
- know how the Turing Test works
- know the Physical Symbol System Hypothesis
- understand the point made in Searle's Chinese Room Argument
- have some familiarity with the history of AI
- understand the concept of agents in the context of AI
- be familiar with some intelligent agent designs
- feel prepared to discuss actual ways of making machines clever and making them learn by themselves...

Overview: topics

- Intelligence
- Artificial Intelligence
- The Turing Test
- Physical Symbol System Hypothesis
- Searle's Chinese Room
- Strong AI vs Weak AI
- History of AI
- Agents

What is Intelligence?

For each of the following, give three reasons why:

- (a) A dog is more intelligent than a worm
- (b) A human is more intelligent than a dog
- (c) An organisation is more intelligent than an individual

Let's play: "Spot the Intelligence!"

- Order the following from most to least intelligent
 - Cockroach
 - Computer
 - Elevator
 - Hammer
 - Human
 - Model-T ford motor car
 - Sunflower

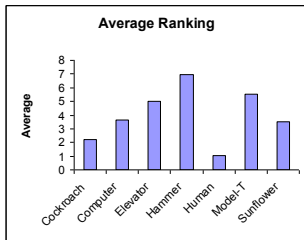
Answers?

- Simple yes / no answers cannot be made from current definitions
- Your answer reflects your beliefs, and is not necessarily correct or incorrect

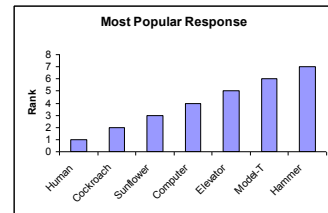
My Ranking

- Human
- Cockroach
- Sunflower
- Computer
- Elevator
- Model-T Ford Motor Car
- Hammer

Average Rank (from 2000)



Results of Intelligence Ranking



What is Intelligence?

(adapted from J. McCarthy – Stanford University)

- What is intelligence?
- One Answer
 - Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.
- Do you agree with this definition?

Intelligence: some terms

- Thought processes
- Reasoning
- Rationality
- Behaviour
- Performance
- Decision making
- Problem solving
- Logic

What is AI? (1)

(adapted from J. McCarthy – Stanford University)

- What is artificial intelligence?
- One possible answer
 - It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.

What is AI? (2)

(adapted from J. McCarthy – Stanford University)

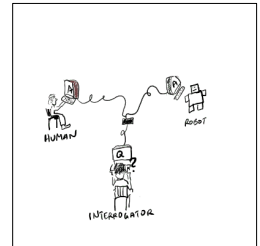
- Is there a solid definition of intelligence that doesn't depend on relating it to human intelligence?
- Answer.
 - Not yet. The problem is that we cannot yet characterise in general what kinds of computational procedures we want to call intelligent. We understand some of the mechanisms of intelligence and not others.

Definitions of Artificial Intelligence

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

The Turing Test

- Turing proposed a definition of Artificial Intelligence independent of the inner mechanisms of the machine
 - <http://www.abelard.org/turpap/turpap.htm>
 - Operational test of intelligent behaviour
 - Given
 - an interrogator, and
 - {A} a computer and {B} a person in a different room
- The interrogator, through asking questions, must determine whether respondent A or respondent B is the computer. If the computer "fools" the interrogator then we can call it intelligent.



Passing the Turing Test

- Natural language processing
 - to enable it to communicate successfully
- Knowledge representation
 - to store information provided before or during the interrogation
- Automated reasoning
 - to use the stored knowledge to answer questions and to draw new conclusions
- Machine learning
 - to adapt to new circumstances and to detect and extrapolate patterns

The Total Turing Test

- Video signal
 - Computer vision
- Physical objects
 - Robotics

Loebner Prize (2008)

<http://www.loebner.net/Prizef/loebner-prize.html>

Human

Good afternoon, how are you?
 good
 it's not afternoon yet
 I'm glad to hear that your well and your right, thanks for pointing that out
 no problem
 how are you?
 I'm well thanks, have you any plans for this afternoon?
 I'm not sure
 go home I guess
 and you?
 That sounds like a good initial plan but the weather should be good today and it would be sad to miss out
 that's true
 although the weather is not nice at the moment unless you like fog
 I have plans to iron and then maybe go out and do some fishing
 that's interesting - where do you go fishing?
 There is a local lake by the green park, it's quiet and relaxing
 can you eat the fish you catch?
 There are restrictions but one or two should be ok, I am keen to try perch

Elbot

Good afternoon, how are you?
 But I wouldn't say that any afternoon could be good.
 That's a bit pesamistic, I think any afternoon has the ability to be good
 Well, afternoon was your word.
 Are you able to see into the future
 How do you respond when people pose this question to you?
 I would hope that the answer was no, otherwise I would be interested to meet you
 Well, what will that solve? After all, the answer is only a matter of perspective.
 That would depend on one's perspective
 Sorry for the tangent, but that reminds me of carefree robot summers spent depending on internal programming.
 Touchee
 With the inspirations you supply, there are no ends to my horizon.
 Have you any plans for later in the day
 This is confusing. Why bring the waking hours into this?
 Just curiosity and I hope that kill no cat!
 That is truly one of the most inspirational qualities of humankind. Even when things seems completely hopeless, you can still hope.

Attributes of the Turing Test (1)

- Positive
 - "standard" test
 - bypasses the true nature of intelligence
 - removes bias
 - wide variety of required knowledge

Attributes of the Turing Test (2)

- Negative
 - focus on symbolic tasks
 - compares machine with ‘human’ intelligence
 - far too restrictive
 - difficult to do in practice



xkcd.com

Criticisms of the Turing Test

- Defines intelligence purely in terms of behaviour
- Behaviour alone is not a test for intelligence
- Conversation is not the ultimate display of intelligence
- Suffers the same problems as standard intelligent (or IQ) tests i.e. culturally dependent and therefore far too restrictive
- Output is generated by blindly following a set of rules

Physical Symbol System Hypothesis

(Simon and Newell, 1976)

A physical symbol system has the necessary and sufficient means for intelligent action.

- A system:
 - Consists of a set of entities, called symbols,
 - Symbols can occur as components of another type of entity called an expression (or symbol structure)
 - A symbol structure is composed of a number of instances (or tokens) of symbols related in some physical way (such as one token being next to another).
 - Contains a collection of these symbol structures
 - Contains a collection of processes that operate on expressions to produce other expressions: processes of creation, modification, reproduction and destruction
- Set of finite symbols that can be composed to form a potential infinite set of expressions

Searle's Chinese Room Thought Experiment

- You are:
 - Monolingual English speaker locked in a room
- You are given
 - a set of rules in English
 - stacks of paper
 - blank
 - covered with Chinese characters (indiscipherable)
- Another set of paper with Chinese characters appears in a hole in the wall of the room
- The set of rules in English enable you to correlate elements of these new papers with elements of the existing stacks of paper and instruct you to give back certain sorts of Chinese symbols with certain sorts of shapes in response on the blank paper
- If you are good at following instructions, your responses are indistinguishable from those of Chinese speakers

Just by looking at your answers, nobody can tell you "don't speak a word of Chinese."

Searle's Chinese Room

- The Point:
 - No matter how intelligent-seeming a computer behaves and no matter what programming makes it behave that way, since the symbols it processes are meaningless (lack semantics) to it, it's not really intelligent.*

Searle's Chinese Room

- Demonstrates that a system can be merely following rules, but not understanding anything at all.
- The system he describes will pass the Turing Test
- Behaviour cannot determine extent of understanding
- Is following rules sufficient for intelligence?

Weak AI

- Machines made to act as if they are intelligent
- Not a tool to study intelligence
- Most AI research fits in here
 - As long as the program works, does it really matter whether it's a simulation of intelligence or real intelligence?

Foundations of AI

- Philosophy – mind, knowledge
- Mathematics – formal rules, reasoning with uncertain information
- Economics – making decisions
- Neuroscience – brains processing information
- Psychology – thinking and acting
- Computer engineering – efficient computers
- Control theory and Cybernetics – machines operating under own control
- Linguistics – language and thought

Strong AI vs Weak AI

- Strong AI
 - duplication of intelligence
 - aims to understand intelligence
- Weak AI
 - simulation of intelligence
 - aims to make computers more 'useful'

Strong AI

- Machines that behave intelligently count as having actual minds
- The appropriately programmed computer is the mind
- Literally said to understand and have other cognitive states
- Consciousness (aware of mental states)
- Phenomenology (direct experience, emotions)
- Intentionality (beliefs, desires, representations actually about the real world)

History of AI - Early 20th Century

- 1943
 - McCulloch and Pitts lay the foundations for neural networks.
 - The term "cybernetics" is coined
- 1950
 - Turing publishes "Computing Machinery and Intelligence".
 - Introduction of the Turing Test as a test for intelligent behaviour.
 - Claude Shannon publishes a detailed analysis of chess playing as search.

Modern History (1)

- 1956
 - John McCarthy coined the term “artificial intelligence” as the topic of the Dartmouth Conference.
 - Demonstration of AI program - the Logic Theorist (LT) - written by Newell, Shaw and Simon (CMU).
- 1952-62
 - Samuel (IBM) wrote the first game-playing program that learns.
- 1962
 - First industrial robot company, Unimation, is founded.

Modern History (2)

- 1965
 - Weizenbaum (MIT) built ELIZA, an interactive program that carries on a dialogue in English on any topic. It became a popular toy at AI centres.
 - See http://www-ai.jis.si/eliza-cgi-bin/eliza_script
 - Or <http://www.manifestation.com/neurotoys/eliza.php3>
- 1967
 - Dendral program (Feigenbaum, Lederberg, Buchanan and Sutherland) is the first successful knowledge-based program for scientific reasoning.
- 1968
 - Minsky and Papert publish “Perceptrons”, demonstrating the limits of simple neural nets.

Modern History (3)

- 1969
 - SRI robot, Shakey, demonstrates combining locomotion, perception and problem solving.
- 1974
 - The first expert system - MYCIN (Stanford) - demonstrates the power of rule-based systems for knowledge representation and inference in the domain of medical diagnosis and therapy.

Modern History (4)

- 1969-1979
 - Knowledge-based (expert) systems
- 1980-1988
 - Expert systems industry booms
- 1988-1993
 - Expert systems industry busts – “AI winter”
 - 1966 US Military report on machine translation: “*The spirit is willing but the flesh is weak*” was re-translated (English-Russian-English) into “*the vodka is good but the meat is rotten*”. Prospects seemed poor and funding was cut.
 - > 2009 version (Babelfish): Spirit is willingly ready but flesh it is weak.

Modern History (5)

- Mid-1980's-present
 - The return of Neural Networks
- 1985
 - Brooks (MIT) develops the concept of behaviour based robots
- 1988-present
 - Resurgence of probability: general increase in technical depth
 - Computational Intelligence / soft computing (Evolutionary computing, Swarm Intelligence, Fuzzy systems, NNs)
- 1995-present
 - Agents, agents everywhere ...
- 2001-present
 - Very large datasets

Modern History (6)

- Search engines
- Social robotics
- Human interaction
- Language
- Perception (faces, objects)
- Vehicles
- Scheduling
- ...

Agents

- Generally computers are obedient, literal, unimaginative servants
 - acceptable for most applications
- However, increasingly we require systems that can *decide for themselves*
- Such computer systems are known as *agents*

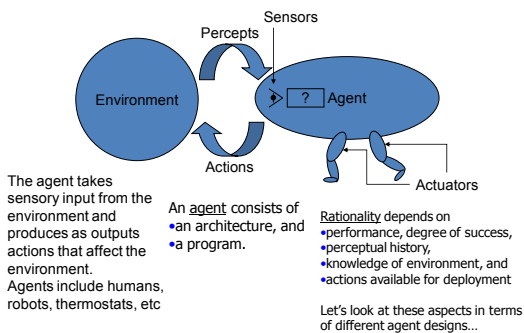
Definition of the term 'Agent'

An *agent* is a computer system that is *situated* in some *environment*, and that is capable of *autonomous action* in this environment in order to meet its design objectives.

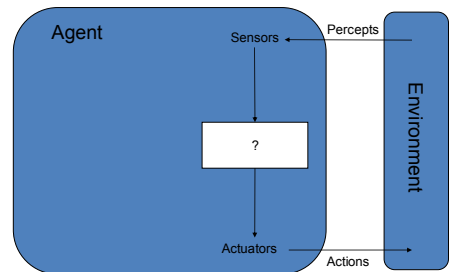
Wooldridge and Jennings (1995)

Autonomy: act without the intervention of humans or other systems, control both over their own internal state and over their behaviour

Agents and Environments



A Basic Agent



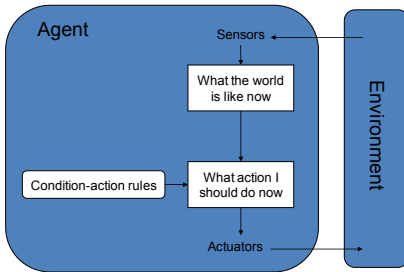
Example: Automated taxi driver

Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi Driver	Safe, fast, legal, comfortable trip, maximise profits	Roads, other traffic, pedestrians, customers	Steer, accelerate, brake, talk to passenger	Cameras, speedometer, GPS, sonar, microphone

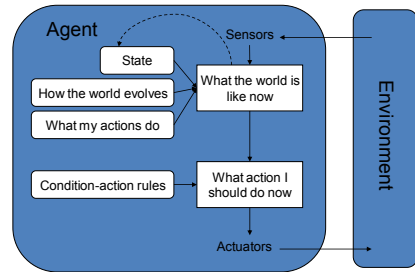
Agent Types

- Simple reflex agents
- Model-based reflex agents
- Goal-based agents
- Utility-based agents
- Learning agents (Week 7)

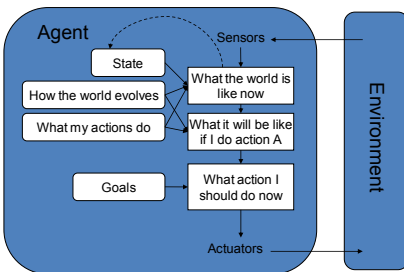
Simple reflex agents



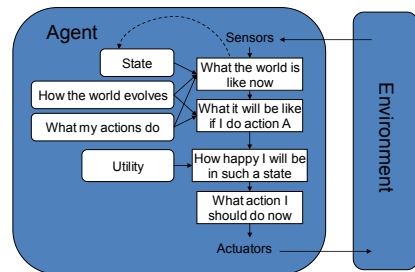
Model-based reflex agents



Goal-based agents



Utility-based agents



Summary

- Definitions of intelligence and artificial intelligence
- The Turing Test – test of ‘intelligence’
- Physical Symbol System Hypothesis
- Searle’s Chinese Room
- Strong AI vs Weak AI
- Different types of agents