WHAT IS ARTIFICIAL INTELLIGENCE?

Aaron Sloman

http://www.cs.bham.ac.uk/~axs

School of Computer Science

The University of Birmingham

Is AI getting machines to behave like the ones in the movies?

Including having emotions?
Including human-like robots?

Is AI engineering?
Is AI science?
Is AI technology?

Is AI psychology?
Is AI philosophy?
Is AI entertainment?

AI can be all those things.

Internet resources for this presentation

These slides can be found here

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk10`

A lot more information about AI and opportunities to study AI is in this document, originally written for careers advisers in UK schools:

`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/aiforschools.html`

Try communicating with the Pop-11 Eliza ‘Chatbot’, here, and see whether you can guess the rules used. You can also see the source code, if you are interested.

`http://www.cs.bham.ac.uk/research/projects/cogaff/eliza`

Some non-interactive movies showing the programs demonstrated in the talk are available here.

`http://www.cs.bham.ac.uk/research/projects/poplog/figs/simagent`

NOTE:

The only operating systems I use are Linux and occasionally Solaris.

My slides are developed and presented using LaTex on Linux.

WHAT IS ARTIFICIAL INTELLIGENCE?

A better answer:

AI is a (relatively) new approach to some very old questions:

WHAT ARE MINDS AND HOW ARE THEY RELATED TO BODIES?

WHAT IS INTELLIGENCE, AND CAN MACHINES HAVE IT?

WHAT ARE FEELINGS, EMOTIONS, BELIEFS, DESIRES...ETC?

AI combines with and contributes to other disciplines, including:

– psychology,
– philosophy,
– linguistics,
– biology,
– anthropology,
– logic,
– mathematics,
– computer science & software engineering,

and other subjects that study humans and other animals.

CAN MACHINES HAVE MINDS?

ONE (PARTIAL) ANSWER:

It is obvious that machines can have minds and be intelligent, because

– Humans are machines and they have minds and intelligence

– The same can be said of other animals, e.g. chimpanzees, elephants, dolphins, dogs, crows, and others.

So that leads us to rephrase our questions:

What sorts of machines are there, and what sorts of minds and what sorts of intelligence can the different sorts have?

Likewise

What sorts of emotions can different sorts of machines have, e.g. humans, elephants, dolphins, mice, fleas, robots of various kinds?

- One of the roles of AI is to investigate these questions.
- Another is to help us build new kinds of useful machines.

AI has two main kinds of goals

- **Science** i.e. studying things that already exist or might exist, explaining how they work, searching for general principles relevant to understanding
  - people,
  - other animals,
  - intelligent machines of the future,
  - and perhaps creatures from other parts of the universe.

- **Engineering**
  using that knowledge to solve practical problems, including
  - making new useful kinds of machines,
  - producing new forms of entertainment
  - perhaps helping us manage ourselves better, e.g. in education, therapy, decision-making, ...
    - because we understand ourselves better
    - because we have new tools.

AI and computer science

AI uses computer science, just as physics uses mathematics, but AI is not computer science, just as physics is not mathematics.

AI uses computers because they are the best available tool, not because they are the object of study.

If we knew how to make brains we could use those instead of computers, for some of the sub-tasks of modelling or explaining natural intelligence.

Already there are attempts to build neural networks that are partly like brains, and do not work like ordinary computers.

But brain science is not computer science.

ARE BRAINS COMPUTERS?

Brains certainly process information, of many different kinds in many different ways.

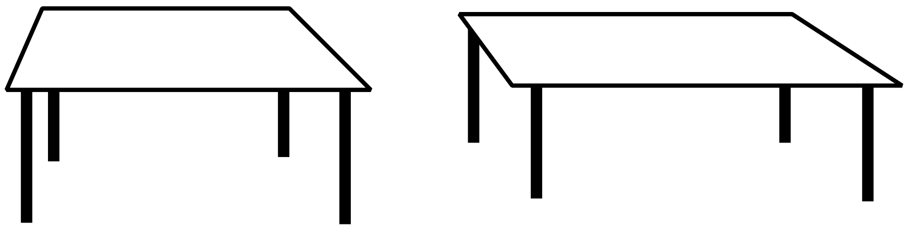
We do not know yet whether there are some mechanisms they use that cannot work on computers as we currently know them, or whether future computers with different basic mechanisms (e.g. chemical mechanisms?) will be needed to replicate some brain functions.

Other brain functions, e.g. playing chess, making plans, solving equations can already be replicated on computers.

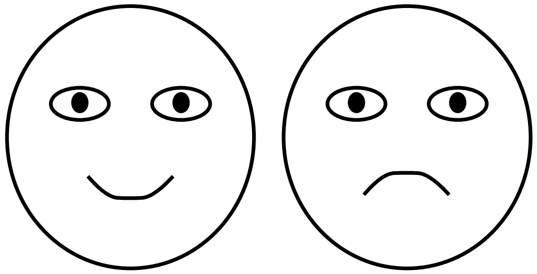
Examples of research problems in AI

Vision – perhaps the hardest problem in AI

How do we get from 2-D patterns of illumination on our retinas to percepts of a 3-D world:

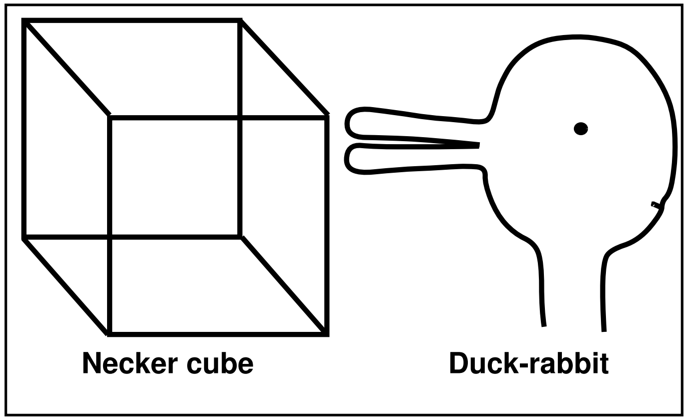


How do we see expressions of emotion in faces?



How can we see the same 2-D visual input in different ways?

E.g. the ambiguous cube. Can you see it flip?
Can you see the duck and the rabbit alternate?
What exactly changes when it flips? In both cases there is no change in the visual stimulus.



We also need to explain perception of motion, visual pleasure, seeing how things work, and much else.

Vision is much, much, more than recognition

Simply recognising something will not tell you how to grasp it.

What competences are required in a visual system to enable a child (or a robot) to get from the first configuration to the second:

- in many different ways,
- with different variations of the first configuration,
- with different variations of the second configuration,
- using the right hand,
- using the left hand,
- using both hands,
- using no hands, only mouth,
- picking up items using a pair of pliers ...?

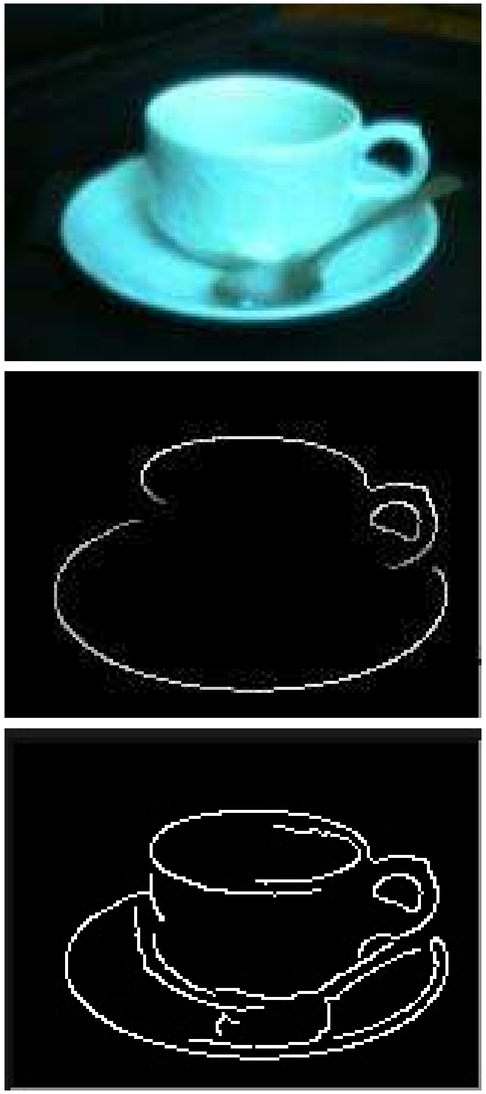
Can you visualise such processes – including interacting curved surfaces?

For more on this see

`http://www.cs.bham.ac.uk/research/projects/cogaff/challenge.pdf`

Vision: beneath the surface

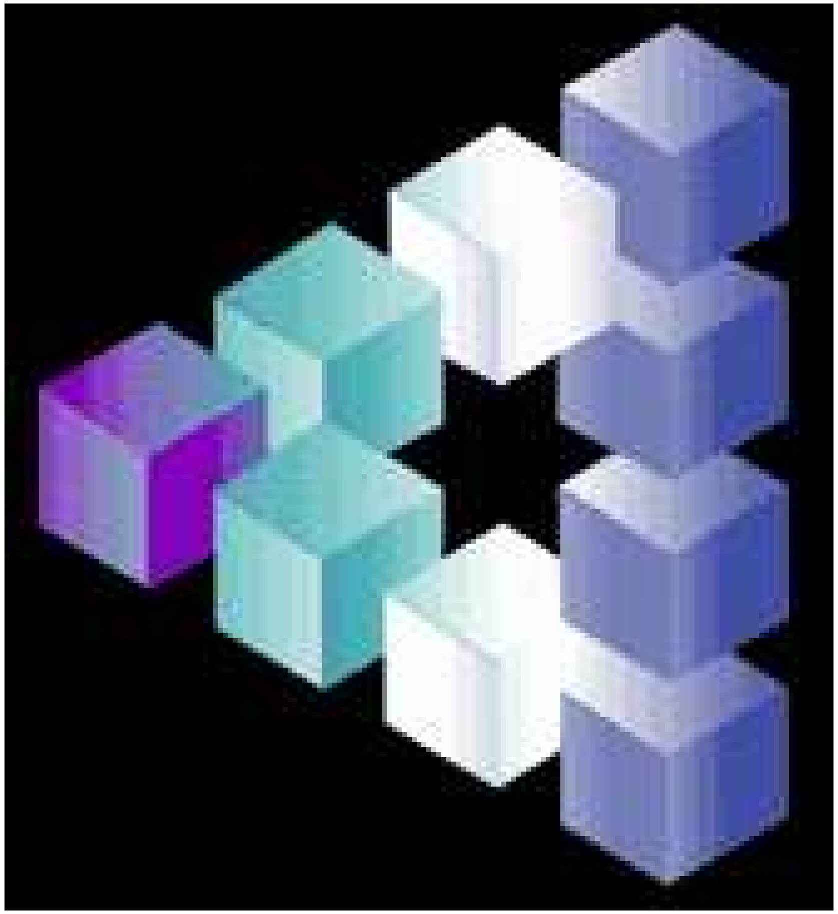
GIVEN AN IMAGE LIKE THE FIRST THERE ARE DIFFERENT SORTS OF 'FEATURE DETECTORS' THAT CAN BE USED, TO FIND RELEVANT STRUCTURE IN THE IMAGE.



- Feature-detectors can be tuned to be more or less sensitive.
- What details are 'visible' in an image depends on how it is processed (Top picture, processed different ways gives middle or bottom picture).
- A visual system should use intelligent decision making about how to process details in different ways in different places.
- This will include deciding what features to detect, deciding how to group them into larger structures, and deciding how to compare them with previously stored information about the world.
- It will need good ways to represent things like curvature, orientation, thickness, weight, since it will not necessarily be able to measure them precisely using rulers, calipers, scales, etc. How do you represent them in your mind?
- What is adequate will depend on what the information is to be used for (e.g. recognition, controlling movement, planning future movements).

Should a robot be able to see impossible objects?

Impossible triangle by Reutersvard – Swedish artist 1934



At first this might simply look like a configuration of cubes in the form of a triangle.

However you should be able to convince yourself that the 3-D configuration cannot exist.

If you remove any three adjacent cubes at one of the corners, the remaining six cubes form a consistent configuration.

But if you put back all the cubes, the relative distances from you become impossible, like

$\mathbf{A} > \mathbf{B} > \mathbf{C} > \mathbf{D}..... > \mathbf{A}$

Should a robot looking at something like this notice the impossibility?

How are you able to see a lot of 3-D fragments that do not add up to a possible 3-D whole? What does that tell us about how a robot should see a 3-D environment?

There are many other AI research areas

How does human language work?

– *Why is the Eliza chatbot not a good model? What's missing?*
– *How do we understand sentences we have never heard before?*
– *How do we produce sentences we have never heard before?*
– *How can we think about and talk about things that do not exist (e.g. Father Christmas, Harry Potter, Darth Vader, heaven, life after death?)*

How do we do mathematical reasoning

– *including thinking about infinite sets (like the set 2, 4, 6, 8, 10, 12, ... or the set of all prime numbers)?*

How do we select and control our actions?

How do we make plans

How do we learn concepts, theories, languages, skills

e.g. programming, musical sight-reading, self-control, thinking ...

Some of these questions arise for other animals also.

And for AI as engineering they arise for robots and intelligent software systems.

AI is inherently multi-disciplinary

Many of the questions are also studied, though in different ways, by other disciplines. E.g.

- Psychology and brain science

  study perception, learning, language use, reasoning.
- Linguistics

  studies the detailed structures of different languages, and how they developed.
- Biology

  studies how various competences evolved, and the interactions between what's in the genes and what is learned.
- Philosophy

  investigates the nature of thought and language and reasoning, the relations between mind and body, the nature of science, and much else.

AI needs to interact with these other disciplines in order to benefit from their theories and in order to take account of the full range of phenomena to be explained.

E.g. from psychology and brain science we can learn some of the very strange things that can happen if a bit of your brain is damaged: you may have some *parts* of a previous ability still working while other parts don't work.

Not all AI researchers focus on the same aims

Some are primarily interested in the Science:

e.g. trying to understand and model how human or animal intelligence works (vision, learning, reasoning, language use, motor control etc.)

Some are primarily interested in the Engineering:

e.g. trying to build useful new kinds of machines, including perhaps machines that are more intelligent than humans

(as calculators already surpass human arithmetic capabilities).

There is more money available to support the engineering aims,
and they are often easier to work on,
so more people work on those aims,
but they indirectly contribute to the first aim (AI as science)
by developing tools and techniques that can be used for both.
just as the science contributes to the engineering.

Common to science and engineering aims: Building WORKING Models

These scientific and engineering objectives of AI are both served by building working models (computer programs, robots, etc.)

– to test our theories

or

– to perform useful tasks

Our degree course includes all three aspects

- Science (especially trying to model how humans work)
- Engineering
- Programming (serving both of the other two)

SOME SIMPLE DEMOS SHOWING SOME AI TECHNIQUES

Some Common Themes

To enable machines to perceive complex scenes, learn about how the world works, communicate in natural language, solve problems, make plans, appreciate works of art, and so on, we must study at least three kinds of things that will have to exist within them:

– Structures that somehow encode useful information, including sentences, maps, diagrams, and probably many other things

– Mechanisms for constructing, manipulating, storing, comparing and using those structures: e.g. how do you use a map? How do you use a book of recipes?
  (Many of those mechanisms will be programs, or algorithms.)

– Architectures in which those structures and mechanisms are combined to form complex wholes with many cooperating parts
  (e.g. perceptual mechanisms, storage systems, motivational mechanisms, motor control subsystems.)

Thus: much of the study of AI is concerned with structures, mechanisms and architectures for combing things.

A very common problem: often the best answer to a question, the best plan for action, the best interpretation of an image, cannot be constructed simply by going through an fixed sequence of steps. Instead an intelligent system has to explore alternatives, looking for a good candidate.

Searching is pervasive in AI – and in much human reasoning.
But searching intelligently can sometimes produce better solutions, faster.

AI programming, whether for science or for engineering, uses many different techniques

Including:

- Symbol-manipulating programs
- Logical reasoning systems
- Neural nets
- Genetic algorithms (evolutionary computation)
- Dynamical systems (based on models from physics)
- Where appropriate, new sorts of hardware

Do not believe anyone who tells you that only one kind of technique works: some ignorant people have narrow-minded views of AI

To support the development of these and other techniques, AI researchers have designed especially powerful and flexible programming languages, e.g. Lisp, Prolog, Pop-11, Scheme, Rule-based languages, constraint languages, ...

For more on this see this presentation on AI development environments, such as the SimAgent toolkit:

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk11

SUB-FIELDS OF AI

AI has many sub-areas studying different problems.

Examples are

- Natural language processing
- Vision
- Learning
- Automated discovery (induction)
- Memory
- Problem solving
- Theorem proving
- Planning
- Creativity
- Modelling motivation and emotions (e.g. in entertainments)
- Expert systems
- Robotics

  Including robots that move around exploring the environment, robots that manipulate 3-D objects, robots that interact with people, robots that learn, ...

- Motivation and emotions
- Architectures for integrated multi-functional minds
- Evolution of intelligence
- (... and many more ...)

AI as SCIENCE

Trying to understand humans and other animals:
(How do spiders build their webs?)

This includes trying to understand and model the following kinds of phenomena as they occur in nature:

- Visual perception
- Language understanding
- Learning
- Problem solving
- Planning
- Motivation
- Emotions
- Creativity
- etc.

But AI as science also attempts to understand what is possible for artificial systems, e.g.

– robots of the future

– disembodied intelligent systems existing only as software agents.

NOTE

**Most of this work is still at a very early stage.
We have a great deal still to learn.**

DIFFERENT VIEWS OF MIND

OLDER APPROACHES:

- A ghost in a machine (dualism)
- Social/political models of mind
- Mechanical models (e.g. levers, steam engines)
- Electrical models (old telephone exchanges)

PROBLEMS WITH OLDER APPROACHES

- Some lack explanatory power (ghost in the machine)
- Some are circular (Social/Political models of mind)
- Some offer explanations that are too crude to explain fine detail and do not generalise
  (e.g. mechanical and electrical models)

NEWER INFORMATION-PROCESSING APPROACHES

These tend to be richer in explanatory power.

They are able to explain wider ranges of phenomena and also account for more fine-grained detail (e.g. explaining perception of subtle differences in a scene).

These use new forms of mechanism:

- Programs
- Neural nets
- Evolutionary processes
- Others ??
- Combinations

Changes in science and engineering

In the old days scientists and engineers studied

– machines that manipulate matter

– machines that manipulate energy

We are now beginning to understand

– machines that manipulate information.

ARE MORE NEW KINDS OF MACHINES NEEDED ?

NOTE:

Many scientists and engineers in many fields now study and talk about “information processing”, whether in computers, computer networks, other machines, or animals, ecosystems, etc.

However, it is hard to say exactly what is meant by “information”, without giving circular definitions.

Explaining this clearly is in part a philosophical problem.

The same is true of “matter”, “energy” and other deep scientific concepts.
You can find some notes about the concept of “information” here
`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/whats-information.html`

And finally

We are only now beginning to understand what information-processing machines are what intelligence is and what we are.

WHY NOT JOIN IN THE ADVENTURE?

Some sources of information

`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/aiforschools.html`

`http://www.aaai.org/aitopics/`

(There are also many textbooks on AI referred to at those web sites.)

These slides are online at

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk10`