Future Human-Like Robots Requirements vs. Designs

Understand problems before you try to solve them (Using iterated implementation if necessary)

Aaron Sloman
University of Birmingham
http://www.cs.bham.ac.uk/~axs/

These slides are in my 'talks' directory:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#fet09

Some related talks

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mm09
Ontologies for baby animals and robots – From "baby stuff" to the world of adult science: Developmental AI from a Kantian viewpoint.

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler
A New Approach to Philosophy of Mathematics: Design a young explorer, able to discover "toddler theorems"

I am not trying to build a robot

I am trying to understand what evolution did.

One way of doing that is trying to build things, to find out what's wrong with our theories and what the problems are.

But also keep looking at products of evolution to compare them with what you have achieved so far....

On making predictions and producing the future

Arthur C. Clarke's Three Laws of Prediction:

1. When a distinguished but elderly scientist states that something is possible, he is almost certainly right.

   When he states that something is impossible, he is very probably wrong.

AI/Robotics is an exception: the predictions of elderly and not so elderly scientists are always wrong, i.e.

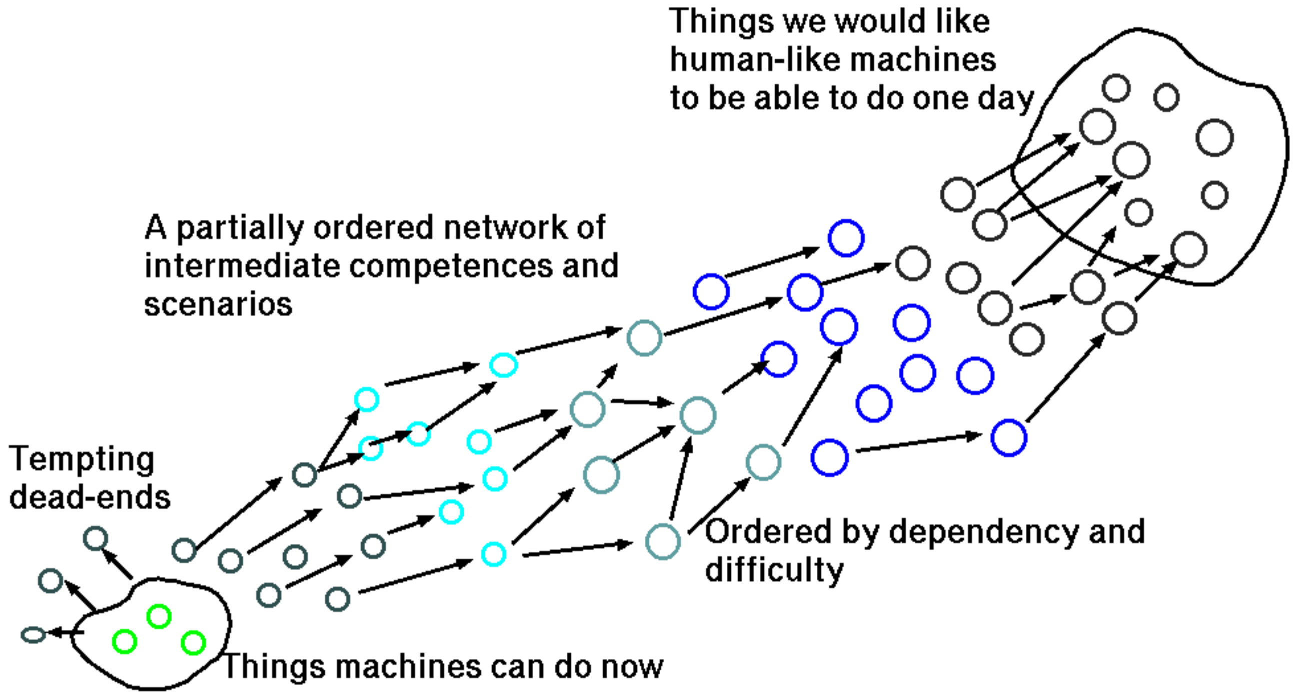
In the next few years we shall have machines that can ... blah blah blah...

The other two laws:

2. The only way of discovering the limits of the possible is to venture a little way past them into the impossible.
3. Any sufficiently advanced technology is indistinguishable from magic.

Why are people too optimistic?

- It is not because they are dishonest
- It is because they think they know what the requirements for intelligent machines are, i.e. what sorts of competences need to be displayed.
- So they assume that finding mechanisms to produce those competences will solve the problem.



- They (unwittingly) pursue tempting dead-ends because they aim for local improvements instead of trying to understand the long term requirements.

  E.g. trying to do better on some benchmark tests used by everyone else.

Wrong moves often based on poor concepts

- Take some widely used term that characterises allegedly important features of human intelligence
- Assume you know what sorts of behaviours would constitute meriting that description.

  Often citing some philosopher or psychologist as an authority.
- Try to implement those behaviours

EXAMPLES:

- Emotion – Having emotions; Recognising emotions; very bad definitions accepted

  Need a theory of control systems
- Motivation – Reward based; Architecture basted
- Seeing – confused with recognition, model building, structure modelling
- Learning
  - Driven by reward based motivation instead of architecture-based motivation.
  - Language learning misconstrued: it's really collaborative language construction
- Communicating – ignoring most aspects of human communication
- Embodiment – ignoring what's most important about embodiment: the environment
- Trying to make the concept of metaphor do impossible work.
- Having consciousness
- Understanding X for various values of X
  - Numbers. Causation. Affordances. Various object categories (e.g. 'car', 'cow')

Using probabilities instead of improved information

How should noisy data and uncertainty be handled.

- Many assume the answer is to use probabilistic reasoning

  Requiring very complex apparatus for manipulating probability distributions.

  With good results in a certain subset of cases and very poor results in others.

- Why do that if there are better alternatives, e.g.

  - Getting better information
    - change viewpoint
    - move an obstacle
    - move the partly hidden object (translate, rotate, ...)
  - Altering a goal so that uncertainty is removed
  - Select an alternative means

- This requires developing a theory of epistemic affordances

Starting with the wrong ontology

E.g. some assumptions about what a robot needs to know about, perceive, think about, reason about....

- Objects

  instead of kinds of stuff, kinds of object fragment, kinds of relationships, kinds of process

- Affordances

  instead of processes and proto affordances (possibilities and constraints on processes)

- Assuming a global metric, e.g. coordinate systems

  Instead of building up problem-specific semi-metrical spatial structures and later unifying them.

Broadening research minds

- Need to explore more varieties of representation

  Not just logic, vectors of measures, property lists.

  E.g.

    multi-stable networks of dynamical systems – could they be good for representing perceived or thought about processes?

    Use of semi-metrical representations of spatial and temporal relations.

- Need to explore more varieties of architecture

  Not just pipelines and layers but perhaps different subsystems corresponding to different animal ancestors.

- Need to explore more varieties of development

  Including substantive ontology extension

  Developing new forms of representation

  Growing new sub-architectures or new architectural linkages

  Developing understanding of why some things have to be true

    (the basis of mathematical competences).
    see the on-the-fly session at 17.30 on toddler mathematics.

- Need to explore space of sets of requirements (niche space) and space of sets of designs (design space)

  Exploring just one design is like trying to do physics only in Prague

So we need to interact more with biologists

The essence of life is information-processing, including information-based control

construction

repair

digestion

defence (predators, diseases)

reproduction, ...

– probably more kinds than we know about yet.

I spend a lot of time now talking to not only psychologists and neuroscientists but also biologists studying animal cognition:

bird

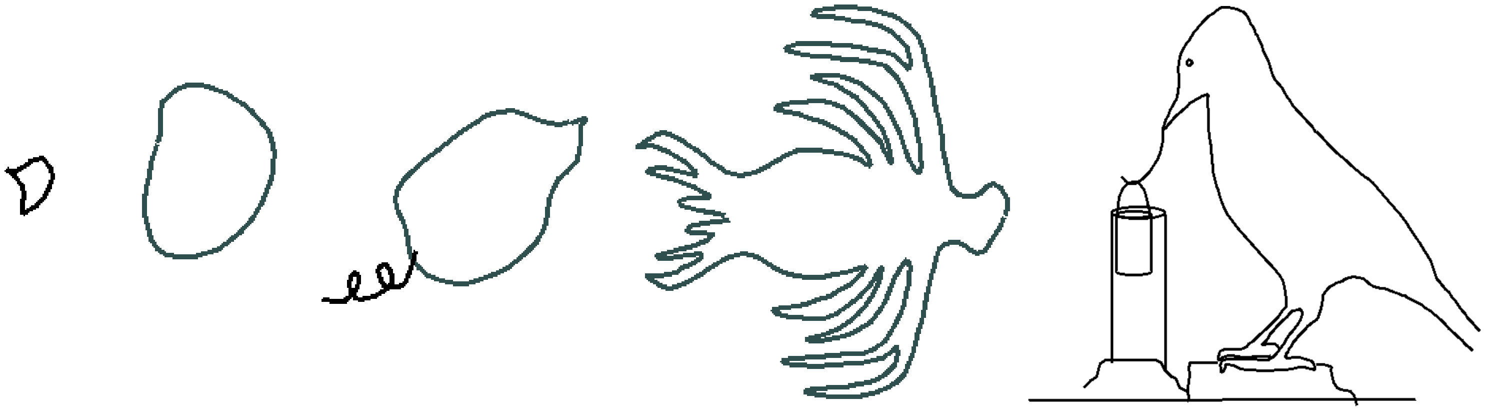
primate

octopus

Is there convergent cognitive evolution –with very different hardware?

Driven by deep features of the environment...

All organisms are information-processors but the information to be processed has changed and so have the means



Types of environment with different information-processing requirements

- Chemical soup
- Soup with detectable gradients
- Soup plus some stable structures (places with good stuff, bad stuff, obstacles, supports, shelters)
- Things that have to be manipulated to be eaten (e.g. disassembled)
- Controllable manipulators
- Things that try to eat you
- Food that tries to escape
- Mates with preferences
- Competitors for food and mates
- Collaborators that need, or can supply, information.

Can we do better?

Can we do better?

YES WE CAN!

But we shall need to

- alter our educational system to provide graduates with a broader outlook
- depend less on young and inexperienced research students to do our research
- change career structures to allow bright researchers to go on learning new disciplines

  (Interdisciplinarity has to be in brains not buildings or departments or projects)
- Make sure more people learn to do analytical philosophy (very hard to teach).
- Broaden types of computational thinking

  Maybe there are kinds of computation we don't yet imagine?

  Maybe biological evolution has already discovered more than we have?
  (possibly requiring new kinds of mathematics).