School of Computer Science
University of Birmingham UK

CN - CR

Every intelligent ghost must contain a machine

an information-processing machine

HBP/NeuroRobotics: A slightly skeptical look from the standpoint of the Meta-Morphogenesis project

(DRAFT: Liable to change)

Aaron Sloman
School of Computer Science, University of Birmingham.
(Philosopher in a Computer Science department)

GENERIC ABSTRACT for talk at workshop 8-9 Jan 2014
A (possibly) new way to approach AI/Robotics/Cognitive Science

(Some subset of the following will be presented.)

Background: The Human Brain Project is one of two very large long term 'flagship' projects recently selected for funding by the European Commission, summarised here:
https://www.humanbrainproject.eu
One of the sub-projects (SP10) is Neurorobotics, described very briefly here:
https://www.humanbrainproject.eu/neurorobotics-platform
It aims to develop one of the six platforms to be produced by the HBP
https://www.humanbrainproject.eu/discover/the-project/platforms
https://www.humanbrainproject.eu/neurorobotics-platform1

[This is too long, but shortening it would take more time than I have available. Sorry.]

My presentation will use the standpoint of the Meta-Morphogenesis project to draw attention to problems of understanding requirements for systems to be developed in such an ambitious project, illustrated by some of the achievements of biological evolution that cannot easily be identified using current research methods in neuroscience, psychology, cognitive science, linguistics, AI, Robotics, ethology, philosophy etc.
This approach was inspired by the challenge of combining Turing's early work on digital computation (on Turing machines) with the work he published shortly before his death on chemical morphogenesis. I suspect that if he had lived he might have tried to use the combination of ideas to answer one of the great unanswered questions of science: how could a lifeless planet with no information available about forms of life, their requirements, their possible designs, produce the diversity of life forms found on our planet including many highly intelligent animals, among them human mathematicians.
Doing the kind of mathematics that led to Euclid's elements is closely connected with being able to perceive, reason about, and make use of what Gibson called *affordances* in the environment, though I think there were more types of affordance than Gibson recognised, because he was focusing on relatively primitive forms of behaviour.
Perception of the full range of affordances (e.g. affordances for gaining information, affordances for changing information available to others, affordances confronting one's offspring who may need help, and many more) seems to require information-processing mechanisms whose capabilities are very different from current AI systems and robots. Bridging that gap seems to be one of the implicit aims of the HBP even if it hasn't been mentioned explicitly in the project proposal, as far as I know. Some of the problems are presented in this discussion:

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html
Hidden Depths of Triangle Qualia

Depending on what I learn from other participants at the meeting, I shall provide reasons (some of them in the documents listed below) for thinking that there are many aspects of the ways in which brains of humans and many other species work that cannot be identified by physical or other measurements of brain activity or experiments on humans and other animals. Such 'probes' merely produce tiny samples from a vast store of required information about problems solved by evolution over many millions of years, and additional problems solved by epigenetic mechanisms to which evolution delegated important functions.

For similar reasons it would be very difficult for alien scientists to work out what's going on in the World Wide Web by sending teams of researchers to take measurements all over the planet, including setting up experiments in the vicinity of devices connected to the internet. They might be able to make progress if they had independently developed a similar system and understood such topics as the need for machine languages, compilers and interpreters, a variety of programming languages for different purposes, a host of types of virtual machinery (including platform VMs such as operating systems, and application VMs such as word-processors, email handlers, chess and other programs, etc.), various types of concurrency, various types of inter-process communication, various kinds of interrupts, many layers of protocols of various sorts, the problems of security and mechanisms that might be used to address security issues, mechanisms allowing the system to change and grow, and many more.

Brains have additional problems of control because different parts of the internet can have different physical locations and perform quite unrelated parts, whereas animal sensors and effectors are far more constrained. TV cameras connected to the same network can be scattered over wide terrain or attached to multiple mobile devices, whereas eyes, hands, tongue and other sensors are constrained by body size, shape and location on the same body. So animals face a recurring need to decide where to look, what to touch, where to go next, etc.

For these reasons, values, preferences, policies, desires, plans, intentions and related control mechanisms are needed for dealing with competing needs on various time-scales, including simultaneous control of foveal fixations and body parts such as grippers that can interact with parts of the environment. This requires an architecture in which components can at any time be influenced in quite detailed ways by information from other components.

As far as I know there are no artificial working visual systems or language understanding systems that meet such requirements (not least because designers tend to work on isolated subsystems to be assembled later), and nobody knows how brain mechanisms support such tightly integrated, mutually influencing interfaces and mechanisms.

One of the consequences of having a rich repertoire of actions and a variety of sensors in a very rich and changing environment -- often presenting new locations, new spatial configurations, new processes in which multiple objects interact -- is that the space of possibilities is too vast to be covered by current forms of learning, e.g. using pre-labelled images. Somehow organisms confronted with such variety have to develop **generative** theories that enable novel configurations to be parsed, interpreted and related to current goals, plans, preferences, needs, etc. (This is a generalisation of the point

Chomsky made in the 1960s about the need to be able to cope with novel sentences, such as many of the sentences in this document.)
The assumption that all sensor contents, motor signals, and current internal states can usefully be represented as vectors of scalar measures (with fixed dimensionality) as required by many current learning systems is an assumption that just does not fit the changing complexity of actions and environments of many animals. We do not seem to have good theories about the forms of representation used by brains or minds to cope with this diversity, although verbal descriptions, parse trees, collections of logical formulae, networks, graphs, and various kinds of dynamical systems may provide hints.
In particular, the fact that we cannot get current computers to replicate the kinds of geometrical discoveries leading up to Euclid's elements, seems to be closely related to our failure so far to give machines the ability to perceive and understand the rich variety of types of affordances (collections of possibilities, constraints on possibilities, invariants across process types) required for intelligent perception and action. J.J.Gibson introduced the notion of 'affordance' but explored only a small subset of cases. I've tried to point out the need to go far beyond Gibson in this presentation on the functions of vision:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#gibson
"What's vision for, and how does it work?
From Marr (and earlier) to Gibson and Beyond"

These issues are not merely relevant to the task of trying to understand, model or replicate human brain function. They are also relevant to the problems of designing useful future robots, for instance personal assistants, or robot carers for the ill or elderly. I have presented some of the problems in this paper (published in a book on Artificial Companions edited by Yorick Wilks):
http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#oii
"Requirements for Digital Companions: It's harder than you think"

A vast amount has been written about consciousness, with different authors presenting very partial views of what the problems are, what the possible answers might be like and what sorts of mechanisms could be involved. If instead we try to understand how the phenomena that we are interested in could have resulted from biological needs and solutions provided by evolution, building on the mechanisms that were previously available, this could lead us to much better theories than are currently available. In particular, instead of assuming that the noun 'consciousness' refers to one thing, so that we can ask what 'it' is, how 'it' evolves, what brain mechanisms enable 'it', etc. we should focus on the adjective, in contexts of the form 'X is conscious of Y', allowing X and Y to vary as widely as possible. This can lead to a theory of consciousness as a highly polymorphic phenomenon with many different functions in different organisms or in different problem situations, with different supporting mechanisms required. When we have a good theory we can try to see how it maps on to what is known about brain mechanisms (and the vast array of information about different sorts of consciousness and influences on consciousness, including drugs, exhaustion, sensor damage, brain damage, and 'software' problems of control in various kinds of psychological disorder). These are not merely esoteric matters to be left to philosophers and medical practitioners: they are required for understanding many aspects of natural information processing and

for designing versatile and effective robots.

Architectures

My impression gained at the workshop is that some members of the project tend to think about architectures in terms that are much too simplistic -- e.g. as if brains were mostly concerrned with managing ‘‘sensori-motor loops’’. I have argued over many years that evolution produced a succession of co-existing architectural layers performing different sorts of functions that could be subdivided in many ways, e.g. ‘‘horizontally’’ in terms of the kinds of environments, tasks, modes of learning, modes of perception, modes of action, modes of interaction with different sorts of things in the environment and, vertically in terms of three overlapping ‘‘pillars’’ of perception, action and more central functioning (e.g. learning, managing of motivation, resolving conflicts, self-observation, etc.).

I have sometimes divided the three layers using the labels ‘‘reactive’’, ‘‘deliberative’’ and ‘‘meta-management’’ (partly based on work by Luc Beaudoin’s PhD thesis (1994). These layers need to be supplemented with ‘‘alarm’’ mechanisms. Some subsytems have to straddle all the layers, e.g. the mechanisms involved in human linguistic competences.

Some of these ideas are presented in connection with the Cognition and Affect project, here
http://www.cs.bham.ac.uk/research/projects/cogaff/#overview
Different kinds of ‘‘functionalist’’ models of mind, related to this, are summarised here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html

Some of Marvin Minsky’s architectural ideas in The Emotion Machine (2006) are closely related.

Background information for the presentation

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html
  The project

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html
  Introduction to the idea that pre-verbal children (like some non-human animals) make mathematical discoveries, and use the results, without realising what they are doing.

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/beyond-modularity.html
  Introduction to some of the ideas presented by Neuro-developmental psychologist Annette Karmiloff-Smith in her 1992 book (Beyond Modularity), including ideas that I feel were not understood by her critics in the BBS ’treatment’ of the book in 1994.

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html
  Discussion of some aspects of the discoveries leading to Euclidean geometry, attempting to identify the modes of reasoning that resist implementation using current automated theorem-proving techniques. Complex hidden features of triangle qualia are not yet grasped by machines.

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/bio-math-phil.html
  Evolution as a blind theorem-prover. The success of evolution depends on the existence of a huge variety (an infinite variety) of mathematical domains many of them discovered and used by evolution -- unwittingly. However some of its products have a deeper understanding. How?

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/evolution-info-transitions.html
  Partial sketch of most of the work to be done.

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vision
  Some gaps in current theories and models of animal/human vision.

- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/austen-info.html
  Jane Austen vs Claude Shannon on "information".

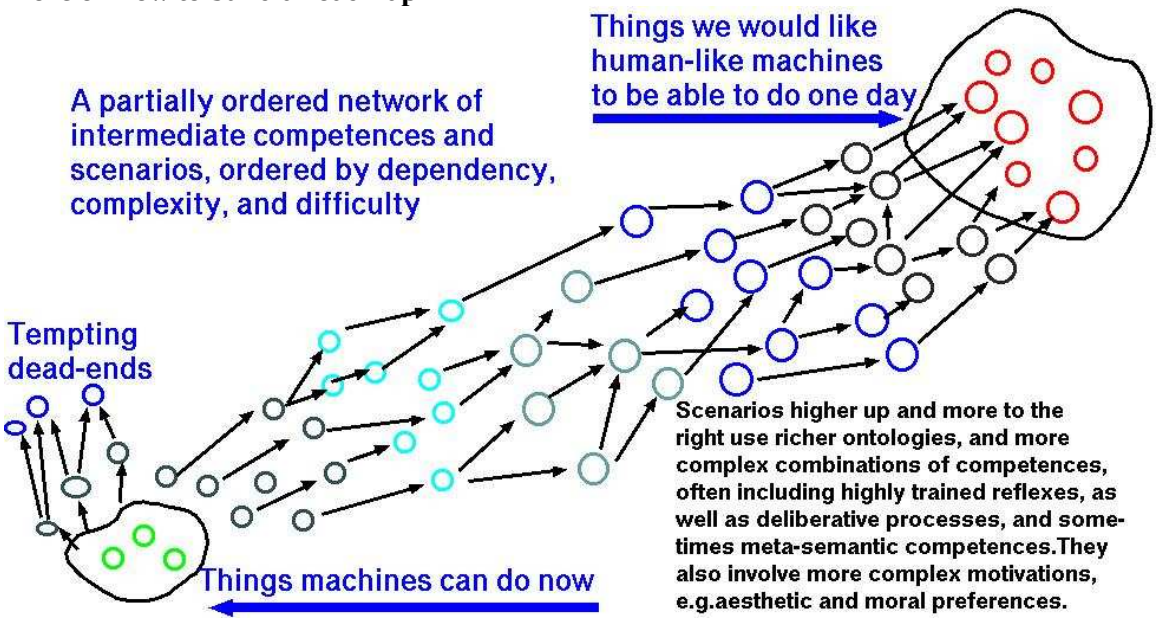
- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/fully-deliberative.html
  Requirements for a Fully-deliberative Architecture
  Contents

```
--  Background (Updated 4 Jan 2014)
--  Different uses of 'reactive'
--  Different interpretations of 'deliberative'
--  This is not a debate about definitions
--  Proto-deliberative vs fully-deliberative
--  Fully-deliberative systems
--  Criteria for fully-deliberative (constructive-deliberative) competence
--  A difference between depth first and breadth first search
--  The need for temporal competence
--  The need for modal competence
--  The need for Affective/Evaluative mechanisms/competences
--  Further requirements related to using results of deliberation
--  Some implications
```

- http://www.cs.stir.ac.uk/gc5/
  UKCRC Grand Challenge 5: Architecture of Brain and Mind
  This was one of a group of long term 'grand challenge' projects identified at a conference organised by Robin Milner and Tony Hoare, on behalf of the UK Computing Research Council (UKCRC) in 2002. GC5: Architecture of Brain and Mind proposed a long term investigating combining top-down bottom-up and middle-out research strategies aiming both to understand natural intelligent systems and eventually to produce artificial systems with similar competences. The work of the challenge has now been taken over by The Human Brain Project, but it may be that some of the materials, including conference and workshop reports remain relevant. The older materials can be found here: http://www.cs.bham.ac.uk/research/projects/cogaff/gc/gc5web.html

- http://www.cs.bham.ac.uk/research/projects/cosy/papers/meta-requirements.html
  A First Draft Analysis of Some Meta-Requirements for Cognitive Systems in Robots
  (An exercise in logical topography analysis.)
  (Meta-functional-requirements)
  Aaron Sloman and David Vernon
  This is a contribution to construction of a research roadmap for future cognitive systems, including intelligent robots.

- **More on how to build a roadmap**



See also: http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk42
Talk 42: COSY-PR-0701 (Also on slideshare: www.slideshare.net/asloman/)
What's a Research Roadmap For? Why do we need one? How can we produce one? (PDF)
A much expanded version of a presentation at the euCognition Munich Workshop
on Research Roadmaps for cognitive systems research on 12 Jan 2007.

Installed: 3 Jan 2014
Last updated: 4 Jan 2014;9 Jan 2014

This document is
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/hbp-robotics.html
A PDF version is also available, though it may not be fully up to date:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/hbp-robotics.pdf

A partial index of discussion notes is in
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/AREADME.html

Maintained by Aaron Sloman
School of Computer Science
The University of Birmingham rcol 80