GRAND CHALLENGE 5: The Architecture of Brain and Mind: Integrating Low-Level Neuronal Brain Processes with High-Level Cognitive Behaviours in a Functioning Robot

Aaron Sloman *
School of Computer Science, University of Birmingham UK
A.Sloman@cs.bham.ac.uk
http://www.cs.bham.ac.uk/~axs/

Introduction

In 2002 the UK Computing Research Committee initiated discussion of possible long term Grand Challenge projects to address deep scientific issues rather than purely practical goals. Seven of the proposals were selected, summarised in a booklet on the UKCRC website (www.ukcrc.org.uk). One of those (GC5), potentially of interest to a wide range of researchers in AI and other disciplines, is about attempting to understand and model mechanisms of brain and mind in the context of an ambitious project to design a robot with a significant subset of the capabilities of a young child. This is a long term project, with no specific deadline, and no earmarked funds, but several UK research proposals have been inspired by the Challenge, and a real start has been made through the closely related EC *Cognitive Systems* initiative which has shorter time-scales and a lot of funding.

This paper is about the *content* of GC5. The challenge assumes that neither brains nor minds can be adequately understood independently of the other. Things minds do, as virtual machines, determine many of the functional requirements for brains. What brains can and cannot do determine possibilities and limitations for human-like minds. We need to understanding their relationships as physical and virtual machines, and as products of biological evolution, social evolution and of course individual development. Moreover those tasks cannot be left to biologists, neuroscientists and psychologists because they lack expert understanding of many aspects of information processing and relations between virtual machines. However all need to be involved.

Explaining the connection between minds and brains was once thought the province of philosophy. We now see this as a scientific and engineering problem, which is inherently about relations between abstract and concrete information processing systems.

Why a child, not an infant?

The focus on emulating a child a few years old was an attempt to maximise generality, since at that stage most of what goes on is still *mostly* due to a combination of general biological features of humans (the genome) and general features of the physical and social environment, whereas adult competence is heavily determined by culture and individual, idiosyncratic history.

Trying to model neonates would be far more problematic than, e.g., a 3 or 4 year old, because infants are so opaque: infant psychology is highly speculative, and very superficial. A good theory of a 3 year old's brain and mind could provide a window through which to peer backwards into what might have produced both, including a large amount of complex and still obscure bootstrapping.

Gaps in Our Knowledge

Many processes in brains and minds are not yet understood, including how we:

- see and understand many kinds of things around us, including 3-D spatial structure, affordances, and causal interactions (see our paper in Cognitive Robotics workshop)
- extend our ontologies and forms of representation
- understand linguistic communications of many kinds
- decide what to do in diverse situations with conflicting pressures and tradeoffs
- plan and control actions, combining tasks where appropriate
- remember things,
- enjoy or dislike things,
- sense the passage of time.
- become aware of our thoughts, emotions, and limitations
- learn about and take account of the mental states of others,
- create and appreciate music, paintings, dance and jokes,

These (and many others) involve both abstract mental processes and concrete physical processes. The aim of GC5 is to explain how both work, and demonstrate this in robots with a significant collection of human-like capabilities, unlike current robots, which are very limited in what they can do. Many current AI systems scale up well in narrow domains, but none 'scale out' by creatively recombining old competences in new circumstances.

The Historical Roots of the Project

Like some other AI projects, GC5 addresses centuries-old quests that have fired the imagination of many inventors, scientists, philosophers and story-tellers: (a) the attempt to understand what we are, and (b) the attempt to make artificial human-like systems, whether entertaining toys, surrogate humans to work in inhospitable environments or intelligent robot helpers of various kinds.

*Thanks to Jim Cunningham, and members of the GC5 Committee: Murray Shanahan (chair), Mike Denham, Steve Furber, and Mark Lee. They have not checked this, however!

The 'strange' relationship between invisible, intangible mental things and our solid flesh and bones has intrigued and mystified mankind for centuries. Despite the efforts of many philosophers and scientists it remained unexplained for centuries. However, around 1842 Ada Lovelace wrote some notes on Babbage's 'analytical engine' in which she recognised what only began to become clear to others more than a century later, namely that the engine

> *might act upon other things besides number, were objects found whose mutual fundamental relations could be expressed by those of the abstract science of operations, and which should be also susceptible of adaptations to the action of the operating notation and mechanism of the engine . . .*

She suggested that using this "abstract science of operations .... *the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent.*" She clearly grasped the idea of a *virtual (abstract) machine* running on a *physical (concrete) machine*, an idea whose generality is still not widely understood.

We now know how to design, make, maintain and sell(!) many kinds of virtual machines including theorem provers, planners, operating systems, teaching packages, plant control systems, and many more. Many of them outperform humans on specific tasks (e.g. playing chess, data-mining), yet there is nothing with the capabilities of three year old child, or a squirrel, or a nest-building bird. The most powerful chess machines cannot discuss or explain their strategies, or be nice to a child player.

How To Do It

This project will require new ways of combining results from researchers in several disciplines, including philosophy, neuroscience, psychology, linguistics, social science, animal cognition, ethology, mechanical and electronic engineering, computer science and AI.

Inspired by advances in neuroscience, the project will attempt to build machines emulating relevant low level brain (and body) mechanisms, and in parallel with that, the project will attempt to implement many kinds of mental processes, initially without requiring biological realism in the implementation mechanisms, but gradually adding more realism as our understanding increases and the technology advances.

Progress will depend on three parallel, mutually-informing, research streams interacting closely:
(1) collection of requirements specifying what needs to be explained — a much harder, much larger, and more important task than generally acknowledged (see poster on this at member's poster session, and the AISB'06 GC5 web site),
(2) bottom-up design and construction of models of brain mechanisms and body parts, constrained both by factual information and by requirements for stream (3)
(3) top-down design and construction of increasingly complex working models combining more and more kinds of competence specified in (1), inspired and constrained by new advances in stream (2).

The systematic collection of requirements (independent of proposed solutions) will constrain both streams (2) and (3), e.g. avoiding arbitrary benchmarks.

Different research groups will study different parts of the problem, but always in the context of the need to put the pieces together in working systems. Stream (3) will at first be mainly a continuation of current AI research but with a new long term research orientation and a commitment to biological realism – even if that means that the machines built are in some respects inferior to current computing systems – e.g. unable to do complex arithmetical or data-sorting tasks that current machines do very fast. Analysing ways in which the models produced in stream 3 succeed and fail, and why, will feed back information to the other two. E.g. it may help to clarify requirements for new low level mechanisms.

A 20 year target could be a child-like robot able to give (limited) help and companionship to a disabled person in a domestic setting – e.g. complementing guide dogs for the blind. There are more and less demanding versions of this target and many of the requirements are far from obvious.

This long-term target needs to be broken down into a large collection of (partially ordered) sub-goals, with different subsets used to define intermediate milestones for judging progress. Some ways of doing this are described in
http://www.cs.bham.ac.uk/research/cogaff/gc/aisb06/gc5.pdf

Achieving all this will require major scientific advances in the component disciplines, and probably new broader education for scientists. Success could provide the foundation for a wide variety of practical applications of new, more flexible and human-like intelligent systems, able to think, learn, communicate and act like humans, and other animals.

So What's Happening?

The UK research councils rightly decided not to earmark special funds for this project, but to invite proposals that would compete with others for general research funds. Several projects referencing this initiative have been funded and new proposals are in preparation. Fortuitously, the EC recently decided to invest significant sums in Cognitive Systems research, with aims overlapping significantly with GC5, though time-scales are much shorter. So several new projects have started. There was also a UK government-inspired 'Foresight' initiative on Cognitive systems. A tutorial at at IJCAI '05 and a symposium AISB'06 were both inspired by GC5. It would be good if research communities from other continents were to join in. Compare the Genome project.

Further information

- http://www.ukcrc.org.uk/
  General information about the UK Grand Challenges initiative.
- http://www.cs.bham.ac.uk/research/cogaff/gc/
  More detail on the "Architecture of Brain and Mind" proposal.
- http://www.cs.bham.ac.uk/research/cogaff/gc/aisb06
  Symposium on GC5 held at AISB'06
- http://cordis.europa.eu/ist/cognition/
  The EC Framework 6 Cognition Unit (see projects page)
- http://www.eucognition.org/
  The euCognition network
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0502
  The Altricial-Precocial Spectrum for Robots