From "Baby Stuff" to the World of Adult Science: Developmental AI from a Kantian viewpoint.

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

NOTE

This is an incomplete set of notes prepared for the workshop on Matching and Meaning, on 9th April 2009 http://dream.inf.ed.ac.uk/events/wmm-2009/ at the AISB'09 Conference in Edinburgh. It is one of a large collection of papers, discussion notes and presentations related to an attempt to understand what problems about information processing, representation and control were solved by evolution over millions of years, how they were solved, and how such problems might be solved in artificial systems – not because I want to build and use them but more because I want to use the design and implementation process to test the ideas. Some other relevant material is on my "Talks" web page, e.g. including the slides for the workshop
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mm09
and a set of slides on some largely unnoticed aspects of child development related to mathematical competences
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler

Abstract. In contrast with ontology developers concerned with a symbolic or digital environment (e.g. the internet), I draw attention to some features of our 3-D spatio-temporal environment that challenge young humans and other intelligent animals and will also challenge future robots. Evolution provides most animals with an ontology that suffices for life, whereas some animals, including humans also have mechanisms for *substantive* ontology extension based on results of interacting with the environment. Future human-like robots will also need this. Since pre-verbal human children and many intelligent non-human animals, including hunting mammals, nest-building birds and primates can interact, often creatively, with complex structures and processes in a 3-D environment, that suggests (a) that they use ontologies that include kinds of material (stuff), kinds of structure, kinds of relationship, kinds of process and kinds of causal interaction and (b) since they don't use a human communicative language they must use information encoded in some form that existed prior to human communicative languages both in our evolutionary history and in individual development. Since evolution could not have anticipated the ontologies required for all human cultures, including advanced scientific cultures, individuals must have ways of achieving substantive ontology extension. The research reported here aims mainly to develop *requirements* for explanatory designs. Developing forms of representation, mechanisms and architectures that meet those requirements will have to come later.

1 INTRODUCTION: THE PROBLEM

Current machine perceptual and manipulative abilities are extremely limited compared with what humans and many other animals can do, despite vast amounts of effort that have gone into work on vision, learning, planning, reasoning and robotics. I try to show that the reasons for the inadequacies are not at all obvious and to indicate some research directions that may be worth pursuing in order to bridge the gaps. A major part of the research task is identifying requirements to be met.

Any information processing system that interacts with some portion of reality by acquiring and making use of information about that reality, needs an ontology if it is to be able to acquire and use new information. It needs an *extendable* ontology if it is to be able to acquire and use new *kinds* of information. What this means is not easy to explain.

An ontology can be explicit or implicit. It is explicit if the contents are specified in some formalism that can be manipulated, stored, transmitted, used to make inferences, etc. (Intermediate cases would have only a subset of these capabilities.)

The ontology is implicit if there is no such formal specification, only a set of mechanisms that deal with instances of the ontology. For example, a common thermostat uses an implicit ontology in which there are temperatures that can vary continuously in one dimension and a control circuit that is either on or off. A washing machine controller has an implicit ontology which allows different washing programmes to be selected, a programme to be started, running or finished, and to go through a sequence of states while running. The designers and human users will make use of an explicit ontology, but the machine has no idea what users are doing or thinking, or what it has done or could do.

It is very likely that most of the ontologies used in most animal brains are implicit. Humans are an exception, at least for some parts of the ontology, and there may be other animals with explicit ontologies. Future human-like robots will probably need explicit ontologies for some of their activities. The designers of such robots will almost certainly need explicit ontologies (and meta-ontologies) even if the things they design use only implicit ontologies.

Biological evolution is a designer with only implicit ontologies and meta-ontologies – unless something is encoded in genomes that nobody has discovered.

An (explicit or implicit) ontology is an (explicit or implicit) specification of what sorts of things can be referred to or represented, what kinds of larger configurations they can be part of, what kinds of things can be part of them, what sorts of properties and relationships can categorise them, in what sorts of ways they can change (i.e. what processes can occur) what kinds of things can be known about them, what kinds of information can be missing about them, what sorts of inferences can be made, and how information about them can used in interacting with them.

An ontology need not be discrete, so it need not be representable with a tree-structured taxonomy: types of existent may be spread out in continuous spaces. Entities may be capable of existing in different sets of relationships, e.g. temporal, spatial, causal, functional, social, and economic, and therefore an ontology can be a tangled high-dimensional network – or more likely it will have a structure for which the notion of dimension is not relevant since its complexity instead of everywhere being factored into N independently variable components may be different in different parts of the network. (This is true of spaces defined by grammars, for example: the set of sentences in English does not have a dimension. Neither does the set of biological organisms.)

If the ontology is extendable, allowing for the discovery, or conjecture, of previously unknown types of reality, then it will not have a fixed structure. A system that can extend its ontology may have an explicit meta-ontology, specifying the ways in which the ontology can change or develop. Alternatively the meta-ontology may be implicit in mechanisms that perform those changes.

Investigating what sort of ontology, fixed or changeable, with or without an explicit meta-ontology, is required for an animal, human, or robot of a particular sort to function in our world is a hard research problem. At present I cannot specify the form of ontology in any detail, so this paper will be very vague in some respects. It is hoped that it will not be too vague to drive some directed research towards removing the vagueness.

2 OTHER KINDS OF ONTOLOGY RESEARCH

Many researchers, some of them labelled as "ontology engineers" are attempting to develop ontologies, or mechanisms for automatically generating ontologies, for use in connection with symbolic or digital environments, such as the internet or a set of corporate databases in a company or government organisation. (E.g. see[6] [8]) In contrast, this paper is not about the development, management, or use of artificial ontologies, but about ontologies used in certain animals. If we can understand the requirements to be met by such ontologies and can find good ways of creating and using them in artificial systems, then perhaps, at some future date, we shall be able to create robots that develop and use them, thereby overcoming one of the serious obstacles to producing robots with human-like intelligence. It could also inform educationalists, and reduce the obstacles schools and other institutions present to development of humans with human-like intelligence.

It may also turn out that biologically inspired ontologies and ontology-related mechanisms are also needed if machines are to understand many of the contents of the internet and other information stores in the same ways as humans do – since most of the contents of the internet are created by humans and are understood by humans, using their biological information-processing systems. So there is no apriori reason why any machine should be able to develop an appropriate understanding without using ontologies and mechanisms that are similar (at some level of abstraction) to those used by humans. That is left as an open question in this paper.

I should make it clear that I do not yet know how to design systems that meet the requirements discussed in this paper, and I don't think anyone else does either. However, in the long run, the common practice of starting from things we know how to do and then looking for minor variants may not, on its own, lead to significant progress, any more than looking for your lost keys where the lamplight is. However, combining that with research on more detailed requirements specifications may eventually reveal new ways to make progress.

3 WHAT ARE ONTOLOGIES FOR?

In many systems for which ontologies are being developed by software engineers, the purpose of the ontology is to facilitate a collection of symbolic operations, e.g. translating documents from one formalism or format to another, or to support data-mining operations in documents and databases from different sources. However, those applications are designed to interact only with symbolic structures, addressing problems that are very different from the problems that confront some organisms (or future robots) interacting with a 3-D environment. The software may produce graphical displays to help human users, e.g. showing a tree or graph on a screen, but typically the machine does not perceive the display, and works only on a symbolic description of the display. Even if it could see and reason about the visual display, that would not be the same as seeing, reasoning about and manipulating 3-D structures and processes in a physical environment.

There are some AI vision and robotic systems that perform impressively in very restricted 2-D or 3-D physical task domains, requiring little understanding of what they are doing or why it works. Examples include balancing a pole, repeatedly welding identical car bodies, or assembling components on a production line.

Some mobile robots are very impressive as hardware+software engineering products, e.g. BigDog – the Boston dynamics robot `http://www.bostondynamics.com/content/sec.php?section=BigDog` and some other mobile robots that are able to keep moving in fairly rough terrain, including, in some cases, moving up stairs or over very irregular obstacles, and, in the case of BigDog, recovering automatically from being pushed off balance, by sticking out a leg to prevent a fall.

However, like a river that very successfully gets water from mountains to the sea, they lack understanding of what they are doing, what they have done, what they could have done, what goals they could achieve in different circumstances, why some goals should be abandoned, etc. though they can sometimes react *as if they* understood, either because a programmer designed the control software to act appropriately, or because a training regime caused a learning mechanism to adjust the control parameters to perform as needed. It is possible to combine a variety of programmed or trained condition-action rules that together give the appearance of understanding, but hide underlying rigidity – like a paint spraying robot that uses only previously learnt movements to spray a piece of furniture even if confronted with an item whose shape is different and needs different spray movements.

Such machines cannot, and have no need to, explain or think about what they are doing, why they are doing it, whether there is any other way of doing it, why they are not doing that, etc. They cannot watch someone else doing similar things and make suggestions for avoiding errors or improving performance. They cannot hypothesise that something is happening that they have never previously encountered that may require their ontology to be extended.

Existing robots that manipulate objects can be triggered to perform an action, but cannot perceive processes, notice new possibilities, or reason about what the result would be if something were to happen, except in very simple cases.

Neither can they reason about why something is not possible. I.e. they lack abilities to perceive and reason about positive and negative affordances.

They cannot wonder why an action failed; wonder what would have happened if they had done something differently; notice that their action might have failed if so and so had occurred part way through, etc.; or realise after the event that some information was available that they did not notice at the time.

Those robots have implicit ontologies. An agent with an implicit ontology has mechanisms that are driven by what happens and in what circumstances, which causes chains of reactions that lead to behaviours responding to those happenings in those situations. In the case of many organisms, e.g. microbes, insects and many others, this type of restricted competence based on an implicit ontology is adequate for the species to continue existing – even if many individuals die "prematurely" because of the rigidity of their responses.

But not all organisms are like that. Some of them (certainly humans, and arguably some others) have the ability to generate information structures representing things that do not exist, including processes that could occur, situations that could arise, things that might have happened in the past, and things that the agent hypothesises might exist, but cannot be perceived either because they are too far away, or obscured by intervening matter, or because they cannot be detected by the agent's sensors, even though when present they can, under some conditions, have effects that are perceived – like the invisible molecular structure that causes sugar, but not sand, to dissolve in water when stirred.

The ability to generate such hypothetical information structures involves having a store of meanings (concepts) that can be combined in various ways to represent possible objects or object parts, states of affairs, events or processes – without having to be driven by sensory inputs. This requires having some *form of representation*, namely something like a *medium* that can be interpreted as expressing information, and which allows meaningful structures to be constructed, manipulated, stored, combined, disassembled, and used for various purposes.

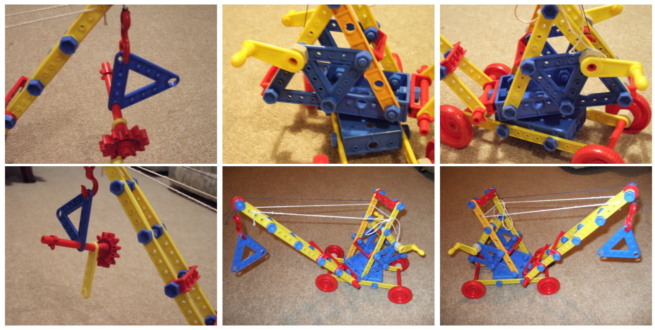
Some of the flexibility of explicit ontologies comes from the fact that the expressive medium need not be physical: in many cases it is more convenient to use symbols or representations composed of entities in virtual machines [17] [14] [12] [9]. Putting such pieces of meaning together is a kind of creativity, which provides the ability to deal with novel situations as well as the ability to represent non-existent situations. (Compare Boden [1] and [11, Chap 2].)

4 ONTOLOGIES FOR SEEING

It is not always noticed that perception can involve similar creativity, in unfamiliar situations. The possession of a suitable ontology, with appropriate generative forms of representation able to express possibilities within the ontology, is required for seeing a novel situation, insofar as such perception involves creating a new usable information structure, either transiently or, if the situation is remembered, in a medium or long term information store. (Compare understanding a sentence you have never heard or read previously, like some of the sentences here.)

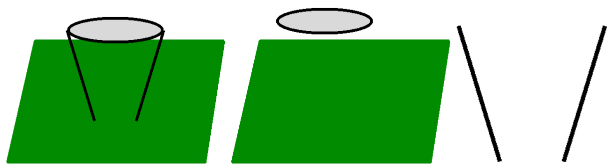
Familiarity with roles of low level pictorial cues in representing 3-D edges, orientation, curvature of surfaces, joins between two objects or surfaces, etc., allows you to use compositional capabilities to see 3-D structure, and some causal and functional relationships, in pictures (even static, monocular pictures) never previously seen.

How many features, relationships (topological, semi-metrical, metrical, causal) can you see in these pictures, taken from `http://www.cs.bham.ac.uk/research/projects/cosy/photos/crane`?

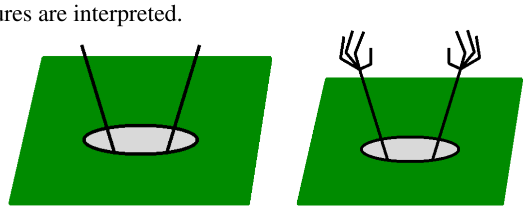


No AI vision system comes close to being able to see these.

The combinatoric creativity involved in perceiving and understanding spatial structures is very different from that involved in doing algebraic or logical operations, because whereas individual components of an algebraic or logical formula retain their syntactic and semantic functions when the elements are rearranged, or when the formula is embedded in a larger formula, that is not true of "analogical" representations of spatial configurations, as pointed out in [10] and [11, Chapter 7].



For example what you see in the above pictures? Only 2-D configurations? Or do you see them as involving 3-D structures and relationships. Do you interpret the same things always in the same way, or are the 3-D relations you see between two parts of the scene dependent on what else is in the image? Notice how context can influence interpretation of parts. Perceptual compositional semantics is highly context-sensitive, as shown by comparing the interpretation of the image components in the previous pictures with the interpretations below. What does the picture on the right suggest to you? Obviously that will depend on how the two multi-pronged structures are interpreted.



Words can add more context that influences how the image is interpreted, by activating different parts of your ontology of 3-D structures, processes, relationships. The figure on the right was based on a "droodle" found on the internet with the caption "Strong worm catches early bird". Another possible caption is "Shark-infested sewer". Different people asked to invent captions will produce different phrases, depending on which ontologies they have and also how fragments of their ontologies are activated by the fragments and relationships of fragments in the picture.

5 SEEING POSSIBILITIES

Some of the interpretations of the droodle above depend on hypothesising physical structures that are not visible but are connected to the structures that are visible. Interpreting droodles, however, would not normally be regarded as a typical illustration of biological functions of vision. In contrast, the images below, which are poor quality, low resolution photographs of 3-D configurations, can be interpreted as scenes in which various kinds of actions are possible, where the possibilities are constrained by the perceived structures – taken from this presentation: `http://www.cs.bham.ac.uk/research/projects/cogaff/challenge.pdf`

If you attend to various locations on the surfaces of these objects and if you see their 3-D shapes, even with low precision and much noise, you will be able to work out roughly how two fingers need to be oriented to grasp at those locations. For instance, try attending to, or getting someone to point at various parts of the rim of the cup, or of the edge of the saucer, or of the spoon or cup handle. In each case you should be able to work out roughly how your finger and thumb will need to be oriented if you grasp at that point. In doing that you are making use of an ontology of spatial positions, orientations and relationships that can hold between surfaces or between objects with surfaces. This shows that perception of action affordances makes uses of an ontology allowing a whole range of spatial arrangements between surfaces of fingers and other objects, among other things. The word "roughly", earlier, indicates that the locations and orientations are not represented with great precision – an important feature of the ontology.

You can probably go beyond thinking about static grasping, to construct a representation of a *process* in which you transform a configuration of one sort into the other sort – which could be done in various different ways, depending where things are held, where they are placed in intermediate configurations, etc. That is, you can make a plan in your head for transforming the configuration seen in one picture of the cup saucer and spoon the configuration in the other picture.

In doing that visualisation, you use an ontology of process fragments that can be "attached" to the various portions of the visible surfaces of the objects involved. Doing that is working out a plan: this need not be a very metrically precise plan in order to be the basis of intelligent action, or intelligent advice to someone else about how to rearrange the objects. An important question that I shall not try to answer in any detail is how you manage to steer through the explosive (infinite) space of continuously varying possible movements at various levels of abstraction so as to find a particular set of movements (a plan) that achieves the desire result. It is often assumed that this requires a search through a discrete space of possible combinations of actions, whose discreteness results from the ability to chunk continuous spaces into sub-spaces (often with slightly fuzzy boundaries). Exactly how that space is related to the ability to see is a topic for another time.

It is not necessary to recognize any of the object categories for the purpose of constructing such a plan, as long as you can see their shapes, namely how the various parts of their surfaces are arranged in space. I could have used pictures of objects that you did not recognise at all. You can do many things with something you see but do not recognize, including planning things to do to it.

6 EXTENDABLE ONTOLOGIES

Extendable, and at least partly explicit, ontologies are needed by animals that have to acquire and use information about, reason about, and interact with rich and complex 3-D structures and processes in the physical environment. Even perception of 2-D images of representing abstract structures, such as written words, requires use of ontologies with multiple layers using different sub-ontologies, as illustrated in [11, Chap. 9]. Impressive recent work in machine vision has shown how computers can use statistical methods to automatically *induce* ontologies as a result of being exposed to many pictures, e.g. the work by Fidler and colleagues in [4, 3]. However at present such ontologies are concerned (a) only with static structures and (b) levels of 2-D organisation in images, as opposed to being able to cope with processes in 3-D space with changing 3-D structures and relationships. Perhaps the methods used will generalise to those cases (using considerably more computing power).

However, not only do the ontologies used in visual systems need to be extendable to cope with new types of entity, they also need to be usable for more tasks than recognition or description of externally presented configurations. As illustrated earlier, the ontology needs also to be usable for representing and reasoning about non-existent but *possible* entities and processes, including possible future sequences of events, but also for representing things in the past, or out of sight, or invisible but capable of having visible effects.

Similar requirements are relevant to future machines doing automated design, inspection and repair of complex machinery; automated rescue systems; domestic aids for disabled people; and robots performing tasks in remote and humanly uninhabitable environments, e.g. on space platforms and other planets. For such systems, required ontologies will not refer only to abstract structures (e.g. web pages and their contents, collections of scientific data, or business information systems concerned with financial transactions) but also to some of the sorts of things many animals can deal with, including spatial structures and processes, causal interactions, assembly or disassembly of objects of varying degrees and kinds of complexity, including changes of

- material properties (e.g. becoming brittle),
- spatial relations (including shape changes),
- causal relations (e.g. producing obstructions, or loosening a grip)
- functional relations (e.g. modifying a structure to serve a new purpose)

7 WITH WHAT STARTING POINT?

A newborn human infant cannot see or do all those things. Why not? – And what has to change to produce those competences? It seems that newborn humans start off with a limited ontology provided by evolution [7] along with the ability to extend the ontology by interacting with the environment.

Some newborn animals can do very sophisticated things very soon after birth (e.g. deer, chicks) so evolution **can** produce innate sophisticated competences, with whatever ontology is required. However in many cases an implicit ontology will suffice.

Infant humans, orangutans, corvids, ... lack behavioural competences some other species have at birth or hatching, even though the other species do not develop so far in their lifetime. Perhaps that is because humans are born with something more powerful than the competences picked up by other animals. This is the hypothesis under development in the collaboration reported in [13, 2].

Many researchers assume *learning* is that more powerful something: but what sort of learning? And from what starting point?

A common assumption is that the initial learning is of a general kind, that can learn anything, provided that enough training data can be provided.

The designers of such systems don't bother to study the environment: they expect to leave that to their future learning systems – but that may not work, for the reason given by McCarthy in [7]:

> "Evolution solved a different problem than that of starting a baby with no a priori assumptions.
>
> ....
>
> Instead of building babies as Cartesian philosophers taking nothing but their sensations for granted, evolution produced babies with innate prejudices that correspond to facts about the world and babies' positions in it. Learning starts from these prejudices. What is the world like, and what are these instinctive prejudices?"

A logicist roboticist might think all the required innate prejudices can be expressed as axioms and deployed through a logic engine. However, studying the environment animals interact with, and learn in, suggests that we need a much richer theory, involving what McCarthy describes, and also

- An initial architecture, that can extend itself in certain ways, including ontology extension.
- Initial (still unknown) forms or representation adequate for encoding specific sorts of information (including information about processes in which 3-D surfaces change their shapes and spatial relations), and which support specific forms of information manipulation.
- Initial sensory, motor, and internal processing mechanisms, including mechanisms for constructing new goals, for goal conflict resolution, and for detecting opportunities to learn.
- Initial behavioural dispositions that drive learning tailored to perceiving and producing 3-D structures and processes.
- An initial, mostly implicit, "framework theory" determining the type of ontology that is assumed and ways in which it can be used and extended. Compare Kant's [5].

  E.g. implicit assumptions about the topology of space/time, kinds of stuff able to occupy and move around in space, modes of composition of structures and processes, kinds of process that can occur involving the stuff, kinds of causation, the differences between doing and passive sensing, ...

- Delayed activation of an architectural layer that uses the combination of the environment and the early architecture as a new developmental "playground" in order to drive ever more sophisticated testing, debugging, and extensions as conjectured in [2].

8 WHAT SORT OF INITIAL ONTOLOGY?

Many theorists assume that the initial ontology includes only sensory and motor contents and patterns relating them, a somatic, multi-modal, ontology) – I claim that will not suffice for children, chimps, or crows. Instead, I conjecture that from the start many learners will use, and attempt to extend, an *exosomatic*, amodal ontology (about what's going on outside – not just the shadows on Plato's cave wall), including:

- bits of stuff (of various kinds) that can occur in the environment
- bits of surface of bits of stuff, in various shapes, locations, orientations
- bits of process (of various kinds) that can occur in the environment
- ways of combining them to construct larger structures and processes in the environment (not necessarily with global consistency)
- at various levels of abstraction: metrical, semi-metrical, topological, causal, functional....

Semi-metrical representations include things like: "W is further from X than Y is from Z", orderings with gap descriptions, symmetries and partial symmetries. (And other things, still to be determined.)

Semi-metrical distance and angle measures could include comparisons between distances and angles instead of use of global units, like 'cm' or 'degrees'.

Instead of items in the environment being located relative to a single global coordinate frame, they could be embedded in (changing) networks of more or less local relations of the above types.

9 HOW CAN ALL THIS WORK ?

Powerful multi-layer, extendable constraint-propagation mechanisms will need to be available for vision, haptic perception, reasoning, planning, predicting, etc. to work. For more on this see, for example, [18]. The main unsolved problem seems to be: what forms of representation are required to support these processes?

It is argued in [16] that in our pre-linguistic evolutionary history, our pre-verbal individual development and in some other non-verbal animals, there are "languages" that are not used for communication, but are used **internally** for perception, reasoning, goal formation, planning, plan execution, question formation, prediction, explanation, causal understanding, as described above, and those languages include (a) structural variability (for dealing with novelty), (b) compositional semantics (modified by context sensitivity) (c) manipulability (for reasoning, planning, hypothesising, etc.).

Suggestions for making progress

Instead of the normal AI strategy of thinking about how to extend our existing mechanisms, or how to deploy them in new ways, perhaps we should spend more time engaged in a deep study of features of the environment and ways of interacting with it, looking at examples of children and other animals doing that, and altering their competences as a result. On that basis we can try to derive constraints on the forms of representation and ontologies that can explain the detailed phenomena observed at different stages of development (which in children are *partially*, not *totally* ordered).

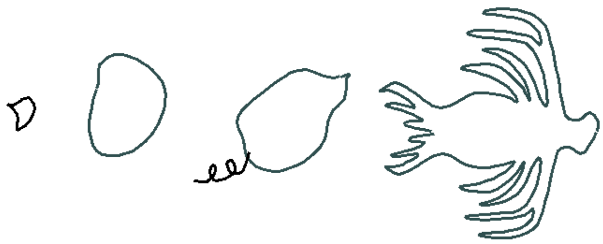
In the light of all that, we should try to design and test mechanisms, architectures, robots that illustrate the theories.

The problems will be different for different sorts of organisms and robots, e.g. depending on the complexity of their sensors and manipulators, the kinds of terrain they inhabit and the kinds of things they need to acquire and avoid. See:[15]

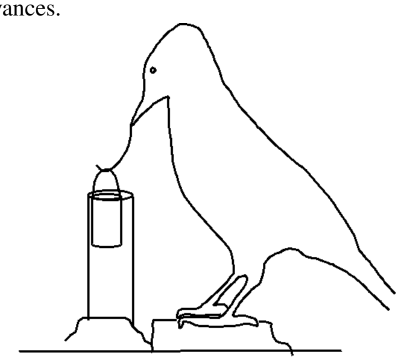
Composition/binding
These different aspects of reality can be composed/combined in many different ways. Long before there was algebraic/functional/logical composition there was spatio-temporal composition. Also auditory/temporal composition – in music and many natural sounds. We need to distinguish composition in the spatio-temporal environment (e.g. combining actions and things acted on, or sounds) from composition in internal representations of things that can be spatio-temporally combined (e.g. composition in representations in virtual machines).

At present we have only a relatively small number of forms of information-composition that we can implement and use in computers. Perhaps by studying the environments of various sorts of intelligent systems very carefully we can derive new requirements for forms of representation and forms of composition and manipulation. This may lead to the creation of new kinds of artificial information-processing systems.

10 LIFE IS INFORMATION PROCESSING



The world contains matter, energy, and information. Organisms acquire and use information, in order to control how they use matter and energy – in order to acquire more matter, energy and information, and also reproduce, repair, defend against intruders, dispose of waste products... Somehow evolution produced more and more sophisticated information processors, driven in part by changes in the environment, which led to changes in morphology which provided more opportunities in the environment requiring more sophisticated information processing (for example, when organisms acquired manipulators that could move independently of eyes), as conjectured in [15]. It seems that evolutionary advances driven/selected by particular challenges often produced opportunities for new more complex advances.



Betty, the New Caledonian Crow made hooks from straight pieces of wire, in several different ways, in order to get a bucket of food out of a tall transparent tube.

All this poses great challenges for science and engineering, namely, to understand that process, to understand the products, and to design working systems that replicate various aspects of the products. In order to do this we need a better understanding of
- the structure of design space
- the structure of niche space
- the many design tradeoffs linking them
- the possible trajectories in design space,
- the possible trajectories in niche space,
- the many complex feedback loops linking both.

11 DEVELOPMENT OF ENVIRONMENT AND COGNITION

The cognitive system, including sensory mechanisms, motor control systems, learning systems, motivational mechanisms, memory, forms of representation, forms of reasoning, etc. that an organism (or robot) needs will depend both on what is in the environment and also what the physical structure and capabilities of the organism are. The current fashion for emphasising the role of embodiment in cognition mostly leads to claims that a particular form of embodiment *solves* or *eliminates* cognitive problems. My claim, on the contrary is that added complexity of animal bodies provides new more complex problems of cognition and control, as explained in [18].

For a micro-organism swimming in an ever changing chemical soup it may suffice to have hill-climbing mechanisms that sense and follow chemical gradients, perhaps choosing different chemical gradients according to the current needs of the organism.

As the environment becomes more structured, more differentiated with more enduring objects and features (e.g. obstacles, food sources, dangers, shelters, manipulable entities) and the organisms become more articulated, with more complex changing needs, the information-processing requirements become increasingly more demanding.

As more complex information processing capabilities develop, the opportunities to observe, modify and combine them in new ways also develop.

The cognitive system, including sensory mechanisms, motor control systems, learning systems, motivational mechanisms, memory, forms of representation, forms of reasoning, etc. that an organism (or robot) needs will depend both on
- what is in the environment

  and
- what the physical structure and capabilities of the organism are.

Many researchers who emphasise the importance of embodiment of animals and robots make a mistaken assumption:

> they claim that embodiment and physical morphology solve the problems and reduce the burdens on cognition, by producing required results "for free" when movements occur.

However, the point I am making is that as bodies become more complex, with more parts that can be moved independently to cooperate with one another in performing complex actions on complex, changeable structures in the environment, the cognitive demands (for perception, learning, planning, reasoning, and motor control, and the ontologies involved) increase substantially, requiring more powerful forms of representation and more complex information-processing architectures.

12 TURING'S MISTAKE?

A major challenge for such an investigation is to understand the variety of possible starting points for an individual born or hatched in a particular sort of environment, after millions of years of evolution of the species. In [19] Turing wrote:

> "Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child brain is something like a notebook as one buys it from the stationer's. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child brain that something like it can be easily programmed."

On this point (little mechanism and much space), Turing was uncharacteristically badly wrong, like all the AI researchers who try to find a small number (some hope **one** will suffice) of powerful, general, learning mechanisms that can learn from arbitrary data: *Evolution did not produce general-purpose data-miners.*

Most species produced by evolution start off with almost all the information they will ever need, leaving only scope for minor adjustments of parameters, e.g. for calibration and minor adaptations. A few species learn a lot using mechanisms that evolved to learn in a 3-D world of static and changing configurations of objects, including other intelligent agents: they start with powerful special-purpose mechanisms. In short: *Evolution itself is a general-purpose data-miner, changing what it mines.* But it needs something like a planet-sized laboratory, and millions of years, to produce things like humans

13 MCCARTHY DISAGREES WITH TURING

As indicated earlier, John McCarthy, in [7], emphasises an important point missed by Turing (and by many AI researchers). In the same article he wrote:

> "Animal behavior, including human intelligence, evolved to survive and succeed in this complex, partially observable and very slightly controllable world. The main features of this world have existed for several billion years and should not have to be learned anew by each person or animal."

McCarthy's own theories about requirements for a neonate are tempered by his goal of attempting to see how much could be achieved using logic. We need to keep an open mind as to which forms of representation and modes of syntactic composition and transformation may be required, or may be useful at times. (As argued in 1971 in [10], and Chapter 7 of [11].)

I am not arguing **against** the use of logic, but **for** a search for additional (new) forms of representation.

14 DEVELOPMENTAL PSYCHOLOGISTS vs DESIGNERS

Many developmental psychologists investigate what is and is not innate in newborn humans, and other animals. Examples studying humans include (among many more): E. Spelke, P. Rochat, E. Gibson & D. Pick, A. Karmiloff-Smith, and much earlier J. Piaget,
and studying animals: N. Tinbergen, K. Lorenz, J. Goodall, W. Köhler, E.C. Tolman, I. Pepperberg, M. Hauser, A. Kacelnik (and colleagues), N. Clayton, S. Healey, F. Warneken, M. Tomasello. Unfortunately not enough of these researchers have learnt to look at something done by a child, chimp, or chick and ask

- How could **that** work? What else can the mechanisms do?
- **How** do they do it?

Instead many researchers ask questions like:

- Under what conditions does this happen?
- How can the task be made easier or more difficult for species X?
- Is this innate or learnt?
- If it is learnt what triggers the learning?
- Which other animals can and cannot do it?
- How early does it happen?
- Which additional tests can I perform to detect these and similar competences?

They don't adopt what McCarthy calls "the designer stance". That is a very difficult thing to do.

Acknowledgements

Several of these ideas resulted from interactions with members of the EU-funded CoSy robotic project `http://www.cognitivesystems.org` and its sequel the CogX project `http://cogx.eu` at the University of Birmingham and in partner organisations. Some of the points here referring to non-human animals were influenced by colleagues in the School of Biosciences working on animal cognition, Jackie Chappell and Susannah Thorpe. My thinking on these topics was inspired in part by reading Kant and Piaget many years ago.

Apologies for this unfinished/unpolished paper. Later versions will be available here:
`http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-mm09.pdf`

REFERENCES

[1] M. A. Boden, *The Creative Mind: Myths and Mechanisms*, Weidenfeld & Nicolson, London, 1990.

[2] Jackie Chappell and Aaron Sloman, 'Natural and artificial meta-configured altricial information-processing systems', *International Journal of Unconventional Computing*, **3**(3), 211–239, (2007). http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609.

[3] S. Fidler, M. Boben, and A. Leonardis, 'Similarity-based cross-layered hierarchical representation for object categorization', in *Proceedings Conference on Computer Vision and Pattern Recognition,*, pp. 1–8. IEEE Press, (2008). http://sanja.fri.uni-lj.si/wp-content/uploads/fidler08cvpr.pdf.

[4] S. Fidler and A. Leonardis, 'Towards Scalable Representations of Object Categories: Learning a Hierarchy of Parts', in *Proceedings Conference on Computer Vision and Pattern Recognition,*, pp. 1–8. IEEE Press, (2007). http://vicos.fri.uni-lj.si/data/alesl/cvpr07fidler.pdf.

[5] I. Kant, *Critique of Pure Reason*, Macmillan, London, 1781. Translated (1929) by Norman Kemp Smith.

[6] Catherine Legg, 'Ontologies on the Semantic Web', **41**, 407–452, (2007,).

[7] J. McCarthy, 'The well-designed child', *Artificial Intelligence*, **172**(18), 2003–2014, (2008). http://www-formal.stanford.edu/jmc/child.html.

[8] Olena Medelyan and Catherine Legg, 'Integrating Cyc and Wikipedia: Folksonomy meets rigorously defined common-sense', in *Proceedings, Workshop on Wikipedia and AI AAAI'08*, eds., R. Bunescu, E. Gabrilovich, and R. Mihalcea, pp. 13–18, Menlo Park, CA, (2008). AAAI.

[9] J. L. Pollock, 'What Am I? Virtual machines and the mind/body problem', *Philosophy and Phenomenological Research.*, **76**(2), 237–309, (2008). http://philsci-archive.pitt.edu/archive/00003341.

[10] A. Sloman, 'Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence', in *Proc 2nd IJCAI*, pp. 209–226, London, (1971). William Kaufmann. http://www.cs.bham.ac.uk/research/cogaff/04.html#200407.

[11] A. Sloman, *The Computer Revolution in Philosophy*, Harvester Press (and Humanities Press), Hassocks, Sussex, 1978. http://www.cs.bham.ac.uk/research/cogaff/crp.

[12] A. Sloman, 'The Well-Designed Young Mathematician', *Artificial Intelligence*, **172**(18), 2015–2034, (2008). http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0807.

[13] A. Sloman and J. Chappell, 'The Altricial-Precocial Spectrum for Robots', in *Proceedings IJCAI'05*, pp. 1187–1192, Edinburgh, (2005). IJCAI. http://www.cs.bham.ac.uk/research/cogaff/05.html#200502.

[14] A. Sloman and R.L. Chrisley, 'Virtual machines and consciousness', *Journal of Consciousness Studies*, **10**(4-5), 113–172, (2003).

[15] Aaron Sloman, 'Diversity of Developmental Trajectories in Natural and Artificial Intelligence', in *Computational Approaches to Representation Change during Learning and Development. AAAI Fall Symposium 2007, Technical Report FS-07-03*, eds., C. T. Morrison and T. Tim Oates, pp. 70–79, Menlo Park, CA, (2007). AAAI Press. http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0704.

[16] Aaron Sloman. Evolution of minds and languages. What evolved first and develops first in children: Languages for communicating, or languages for thinking (Generalised Languages: GLs)?, 2008. http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0702.

[17] Aaron Sloman, 'Virtual Machines in Philosophy, Engineering & Biology', in *Proceedings Workshop on Philosophy & Engineering WPE-2008*, eds., Natasha McCarthy and David Goldberg, Royal Academy of Engineering, London, (2008). http://www.cs.bham.ac.uk/research/projects/cogaff/08.html#803.

[18] Aaron Sloman, 'Some Requirements for Human-like Robots: Why the recent over-emphasis on embodiment has held up progress', in *Creating Brain-like Intelligence*, eds., B. Sendhoff, E. Koerner, O. Sporns, H. Ritter, and K. Doya, 248–277, Springer-Verlag, Berlin, (2009). http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0804.

[19] A.M. Turing, 'Computing machinery and intelligence', *Mind*, **59**, 433–460, (1950). (reprinted in E.A. Feigenbaum and J. Feldman (eds) *Computers and Thought* McGraw-Hill, New York, 1963, 11–35).