In *Proceedings 7th International Conference on Mathematical Knowledge Management*
Birmingham, UK, 28-30 July 2008
http://events.cs.bham.ac.uk/cicm08/mkm08/

Kantian Philosophy of Mathematics and Young Robots

Aaron Sloman

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

Abstract. A child, or young human-like robot of the future, needs to develop an information-processing architecture, forms of representation, and mechanisms to support perceiving, manipulating, and thinking about the world, especially perceiving and thinking about actual and possible structures and processes in a 3-D environment. The mechanisms for extending those representations and mechanisms, are also the core mechanisms required for developing mathematical competences, especially geometric and topological reasoning competences. Understanding both the natural processes and the requirements for future human-like robots requires AI designers to develop new forms of representation and mechanisms for geometric and topological reasoning to explain a child's (or robot's) development of understanding of affordances, and the proto-affordances that underlie them. A suitable multi-functional self-extending architecture will enable those competences to be developed. Within such a machine, human-like mathematical learning will be possible. It is argued that this can support Kant's philosophy of mathematics, as against Humean philosophies. It also exposes serious limitations in studies of mathematical development by psychologists.
Keywords:
learning mathematics, philosophy of mathematics, robot 3-D vision, self-extending architecture, epigenetic robotics

1 Introduction: Approaches to Mathematics

Some people have a central interest in mathematics, e.g.: mathematicians, whose job is to extend mathematical knowledge and to teach it; scientists, who routinely use mathematics to express data, analyse data, formulate theories, make predictions, construct explanations, etc.; and engineers, who use it to derive requirements for their designs and to check consequences of designs.

Others study aspects of mathematics: e.g. philosophers who discuss the nature of mathematical concepts and knowledge; psychologists who study how and when people acquire various mathematical concepts and kinds of mathematical competence; biologists interested in which animals have any mathematical competence, what genetic capabilities make that possible, how it evolved, and how that is expressed in a genome; and AI researchers who investigate ways of giving machines mathematical capabilities. Finally there are the children of all ages who are required to learn mathematics, including a subset who love playing with and learning about mathematical structures and processes, and many who hate mathematics and make little progress learning it.

My claim is that there are connections between these groups that have not been noticed. In particular, if we can understand how children and other animals learn about, perceive and manipulate objects in the environment and learn to think about what they are doing, we shall discover that the competences they need are closely related to requirements for learning about mathematics and making mathematical discoveries. Moreover, if we make robots that interact with and learn from the environment in the same way, they too will be able to be mathematical learners – a new kind of biologically inspired robot. The insights that we can gain from this link can shed light on old problems in philosophy of mathematics, and psychology of mathematics.

And finally, if we really make progress in this area, we may be able to revolutionise mathematical education for humans, on the basis of a much deeper understanding of what it is to be an intelligent learner of mathematics.[1]

2 Philosophies of Mathematics

Over many centuries, different views of the nature of mathematical knowledge and discovery have been developed. Those include differing philosophical views about the nature of numbers, for example (simplifying enormously):

1. Number concepts and laws are abstractions from operations on perceived groups of objects. (J.S.Mill (1843) and some developmental psychologists. See (Rips, Bloomfield, & Asmuth, 2008) for discussion.)
2. Numbers are mental objects, created by human mental processes. Facts about numbers are discovered by performing mental experiments. (Intuitionist logicians, e.g. Brouwer. Heyting (1956), Kant? (1781))
3. Numbers and their properties are things we can discover by thinking about them in the right way (Kant, and many mathematicians, e.g (Penrose, 1989)).
4. Numbers are sets of sets, or predicates of predicates, definable in purely logical terms. E.g. the number one is the set of all sets capable of being mapped bi-uniquely onto the set containing nothing but the empty set. (Frege (1950), Russell (1903), and other logicists).
5. Numerals are meaningless symbols manipulated according to arbitrary rules. Mathematical discoveries are merely discoveries about the properties of such games with symbols. (Formalists, e.g. Hilbert.)
6. Numbers are implicitly defined by a collection of axioms, such as Peano's axioms. Any collection of things satisfying these axioms can be called a set of numbers. The nature of the elements of the set is irrelevant. Mathematical discoveries about numbers are merely discoveries of logical consequences of the axioms. (Many mathematicians)
7. It doesn't matter what numbers are: we are only interested in which statements about them follow from which others (Russell, (1917)).
8. There is no one correct answer to the question 'what are numbers?' People play a motley of 'games' using number words and other symbols, and a full account of the nature of numbers would simply be an analysis of these games (including the activity of mathematicians) and the roles they play in our lives. (Wittgenstein: *Remarks on the Foundations of Mathematics*)

J.S. Mill claimed that mathematical knowledge was empirical, abstracted from experiences of actions like counting or matching sets, and capable of being falsified by experience. Most thinkers regard mathematical knowledge as non-empirical, and not refutable by experiments on the environment, though interacting with the environment, including making drawings, or doing calculations on paper, help us notice mathematical truths, or help us find counter-examples to mathematical claims (Lakatos (1976)).

Some philosophers who regard mathematical knowledge as non-empirical think it is all essentially empty of content, because it merely expresses definitions or "relations between our ideas" – i.e. such knowledge is "analytic" (defined in (Sloman, 1965)). Hume had this sort of view of mathematics. Kant (1781) reacted by arguing for a third kind of knowledge, which is neither empirical nor analytic but "synthetic": these significantly extend our knowledge.

As a graduate student in Oxford around 1960 I found that something like Hume's view was common among the philosophers I encountered, so I tried, in my DPhil thesis (Sloman, 1962),

[1] See also http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#math-robot (PDF presentation on whether a baby robot can grow up to be a mathematician).

to explain and defend Kant's view, that mathematical knowledge is synthetic and *a priori* (non-empirical), which clearly accorded much better with the experience of doing mathematics.

"*A priori*" does not imply "innate". Discovering or understanding a mathematical proof can be a difficult achievement. Although mathematicians can make mistakes and may have to debug their proofs, their definitions, their algorithms, their axiom-systems, and even their examples, as shown by Lakatos in (Lakatos, 1976), mathematical knowledge is not empirical in the sense in which geological or chemical knowledge is, namely subject to refutation by new physical occurrences. Both flaws and the fixes in mathematics can be discovered merely by thinking.

3 Psychological Theories About Number Concepts

It is often supposed that the visual or auditory ability to distinguish groups with different numbers of elements (subitizing) displays an understanding of number. However this is simply a perceptual capability. A deeper understanding of numbers requires a wider range of abilities, discussed further below.

Rips *et al.* (Rips et al., 2008) give a useful survey of psychological theories about number concepts. They rightly criticise theories that treat number concepts as abstracted from perception of groups of objects, and discuss alternative requirements for a child to have a concept of number, concluding that having a concept of number involves having some understanding (not necessarily consciously) of a logical schema something like Peano's five axioms. They claim that that is what enables a child to work out various properties of numbers, e.g. the commutativity of addition, and the existence of indefinitely larger numbers. This implies that such children have the logical capabilities required to draw conclusions from the axioms, though not necessarily consciously. That immediately raises the question how children can acquire such competences. They conclude that somehow the Peano schema and the logical competences are innately built into the child's "background architecture" (but do not specify how that could work).

They do not consider an alternative possibility presented in (Chappell & Sloman, 2007; Sloman & Chappell, 2007) according to which such competences may be meta-configured, i.e. not determined in the genome, but produced through interactions with the environment that generate layers of meta-competences (competences that enable new competences to be acquired). Some hints about how that might occur are presented below.

Many psychologists and researchers in animal cognition misguidedly search for experimental tests for whether a child or animal does or does not understand what numbers are.[2] Rips *et al.* are not so committed to specifying an experimental test, but they do require a definition that makes a clear distinction between understanding and not understanding what numbers are.

4 Towards an AI Model of Learning About Numbers

As far as I know, no developmental psychologists have considered the alternative view, presented 30 years ago in (Sloman, 1978), chapter 8, that there is no single distinction between having and not having a concept of number, because learning about numbers involves a never-ending process that starts from relatively primitive and general competences that are not specifically mathematical and gradually adds more and more sophistication, in parallel with the development of other competences. In particular, (Sloman, 1978) suggests that learning about numbers involves developing capabilities of the following sorts:

1. performing a repetitive action;
2. memorising an ordered sequence of arbitrary names;
3. performing two repetitive actions together and keeping them in synchrony;

[2] Compare the mistake of striving for a definitive test for whether animals of some species understand causation, criticised here in presentation 3:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac

4. initiating such a process and then being able to use different stopping conditions, depending on the task;
5. doing all this when one of the actions is uttering a learnt sequence of names;
6. learning rules for extending the sequence of names indefinitely;
7. observing various patterns in such processes and storing information about them, e.g. information about successors and predecessors of numerals, or results of counting onwards various amounts from particular numerals;
8. noticing commonalities between static mappings and process mappings (e.g. paired objects vs paired events);
9. finding mappings between components of static structures as well as the temporal mappings between process-elements;
10. noticing that such mappings have features that are independent of their order (e.g. counting a set of objects in two different orders must give the same result);
11. noticing that numbers themselves can be counted, e.g. the numbers between two specified numbers;
12. noticing possibilities of and constraints on rearrangements of groups of objects – e.g. some can be arranged as rectangular arrays, but not all;
13. learning to compare continuous quantities by dividing them into small components of a standard size and counting.

Such competences and knowledge can be extended indefinitely. Some can be internalised, e.g. counting silently. Documenting all the things that can be discovered about such structures and processes in the first few years of life could fill many pages. (Compare (Liebeck, 1984; Sauvy & Suavy, 1974).) The sub-abilities involved in these processes are useful in achieving practical goals by manipulating objects in the environment and learning good ways to plan and control such achievements. An example might be fetching enough cups to give one each to a group of people, or matching heights of two columns made of bricks, to support a horizontal beam, or ensuring that enough water is in a big jug to fill all the glasses on the table.

Gifted teachers understand that any deep mathematical domain is something that has to be explored from multiple directions, gaining structural insights and developing a variety of perceptual and thinking skills of ever increasing power. That includes learning new constructs, new reasoning procedures, learning to detect partial or erroneous understanding, and finding out how to remedy such deficiencies. (Sloman, 1978) presented some conjectures about some of the information-processing mechanisms involved. As far as I know, nobody has tried giving a robot such capabilities. It should be feasible in a suitably simplified context. I had hoped to do this in a robot project, but other objectives were favoured.[3]

5 Internal Construction Competences

The processes described above require the ability to create (a) new internal information structures, including, for example, structures recording predecessors of numbers, so that it is not necessary always to count up to N to find the predecessor of N, and (b) new algorithms for operating on those structures. As these internal information-structures grow, and algorithms for manipulating them are developed, there are increasingly many opportunities to discover more properties of numbers. The more you know, the more you can learn.

Moreover those constructions do not happen instantaneously or in an error-free process. Many steps are required including much self-debugging, as illustrated in (Sussman, 1975). This depends on self-observation during performance of various tasks, including observations of external actions and of thinking. One form of debugging is what Sussman called detecting the need to create new

[3] See http://www.cs.bham.ac.uk/research/projects/cosy/PlayMate-start.html

critics that run in parallel with other activities and interrupt if some pattern is matched, for instance if disguised division by zero occurs.

The ongoing discovery of new invariant patterns in structures and processes produced when counting, arranging, sorting, or aligning sets of objects, leads to successive extensions of the learner's understanding of numbers. Initially this is just empirical exploration, but later a child may realise that the result of counting a fixed set of objects cannot depend on the order of counting. That invariance is intrinsic to the nature of one-to-one mappings and does not depend on properties of the things being counted, or on how fast or how loud one counts, etc. However, some learners may never notice this non-empirical character of mathematical discoveries until they take a philosophy class!

One of the non-empirical discoveries is that the natural numbers form an infinite set. Kant suggested that this requires grasping that a rule can go on being applied indefinitely. This contrasts with the suggestion by Rips *et al.* (Rips et al., 2008) that a child somehow acquires logical axioms which state that every natural number has exactly one successor and at most one predecessor, and that the first number has no predecessor, from which it follows logically that there is no final number and the sequence of numbers never loops. Instead, a child could learn that there are repetitive processes of two kinds: those that start off with a determinate stopping condition that limits the number of repetitions and those that do not, though they can be stopped by an external process. Tapping a surface, walking, making the same noise repeatedly, swaying from side to side, repeatedly lifting an object and dropping it, are all examples of the latter type.

The general notion of something not occurring is clearly required for intelligent action in an environment. E.g. failure of an action to achieve its goal needs to be detectable. So if the learner has the concept of a repetitive process leading to an event that terminates the process, then the general notion of something not happening can be applied to that to generate the notion of something going on indefinitely. From there, depending on the information processing architecture and the forms of representation available, it may be a small step to the representation of two synchronised processes going on indefinitely, one of which is a counting process.

What is more sophisticated is acquiring a notion of a sequence of sounds or marks that can be generated indefinitely without ever repeating a previous mark. An obvious way to do that is to think of marks made up of one or more dots or strokes. Then the sequence could start with a single stroke, followed by two strokes, followed by three strokes, etc., e.g.

```
| || ||| |||| |||||  etc.
```

That has the disadvantage that the patterns grow large very quickly. That can motivate far more compact notations, like arabic numerals, though any infinitely generative notation will ultimately become physically unmanageable.

6 Extending Simple Number Concepts

A different sort of extension is involved in adding *zero* to the natural numbers, which introduces "anomalies", such as that there is no difference between *adding* zero apples and *subtracting* zero apples from a set of apples.

Negative integers add further confusions. This extension is rarely taught properly, and as a result most people cannot give a coherent explanation of why multiplying two negative numbers should give a positive number. It cannot be *proved* on the basis of previous knowledge because what multiplying by a negative number means is undefined initially. For mathematicians, it is defined by the rules for multiplying negative numbers, and the *simplest* way to extend multiplication rules to negative numbers without disruption of previous generalisations, is to stipulate that multiplying two negatives produces a positive. (Similarly with defining what 3^{-1} and 3^0 should mean.)

Some teachers use demonstrations based on the so-called "number line" to introduce notions of negative integers, but this can lead to serious muddles (e.g. about multiplication). Pamela Liebeck (Liebeck, 1990) developed a game called "scores and forfeits" where players have two sets of tokens: addition of a red token is treated as equivalent to removal of a black token, and vice versa.

(Multiplication was not included.) The person with the biggest surplus of black over red wins. Giving a player both a red and a black token, or removing both a red and a black token makes no difference to the total status of the player. Playing, and especially discussing, the game seemed to give children a deeper understanding of negative numbers and subtraction than standard ways of teaching, presumably because the set of pairs of natural numbers can be used to model accurately the set of positive and negative integers.

Cardinality and orderings are properties of *discrete* sets. Extending the notion of number to include *measures* that are continuously variable, e.g. lengths, areas, volumes and time intervals, requires sophisticated extensions to the learner's ontology and forms of representation – leading to deep mathematical and philosophical problems. In humans, an understanding of Euclidean geometry and topology seems to build on reasoning/planning competences combined with visual competences, as illustrated in (Sauvy & Suavy, 1974). This requires different forms of representation from counting and matching groups of entities. Some of these competences are apparently shared with some other animals – those that are capable of planning and executing novel spatial actions.

7 Doing Philosophy by Doing AI

After completing my D.Phil defending Kant in 1962, I gradually realised something was lacking. My arguments consisted mostly of illustrative examples. Something deeper was required, to explain what goes on (a) when people acquire mathematical *concepts* (e.g. number, infinitely thin line, perfectly straight line, infinite set, etc.), and (b) when they acquire mathematical *knowledge* expressed using those concepts. In 1969, Max Clowes introduced me to programming and AI, and I soon realised that by building a working human-like mind (or suitable fragments of one) we could demonstrate the different modes of development of knowledge discussed by Kant. Many mathematical proofs, especially in Euclidean geometry and topology, but also in number theory, seemed to rely on our ability to perceive structures and structural relationships, so I concluded that explaining how mathematical discoveries were made, depended, in part, on showing how *visual* capabilities, or more generally, *spatial perception and reasoning* capabilities, were related to some kinds of mathematical reasoning.

At that time the dominant view in AI, represented by McCarthy and Hayes (1969) was that logical modes of representation and reasoning were all that an intelligent robot would need. In 1971, I submitted a paper to IJCAI (Sloman, 1971), distinguishing "Fregean" from "analogical" forms of representation and arguing that spatial analogical forms of representation and reasoning could be used in *valid* derivations and could also in some cases help with the organisation of search. Fregean representations are those whose mode of syntactic composition and semantic interpretation use only application of functions to arguments, whereas analogical representations allow properties and relations of parts of a representation to refer to properties and relations of parts of a complex whole, though not necessarily the same properties and relations: in general analogical representations are not isomorphic with what they represent. (E.g. a 2D picture can represent a 3D scene without being isomorphic with it.) The paper was accepted, and subsequently reprinted twice. But it was clear that a lot more work needed to be done to demonstrate how the human spatial reasoning capabilities described therein could be replicated in a machine.

Many others (e.g. (Glasgow, Narayanan, & Chandrasekaran, 1995)) have pointed out the need to provide intelligent machines with spatial forms of representation and reasoning, but progress in replicating human abilities has been very slow: we have not yet developed visual mechanisms that come close to matching those produced by evolution. In part, this is because the requirements for human-like visual systems have not been analysed in sufficient depth (as illustrated in (Sloman, 2009, 2008)). E.g. there is a vast amount of research on object recognition that contributes nothing to our understanding of how 3-D spatial structures and processes are seen or how information about spatial structures and processes is used, for instance in reasoning and acting.[4]

[4] Some of the differences between recognition and perception of 3-D structure are illustrated in http://www.cs.bham.ac.uk/research/projects/cogaff/challenge.pdf

8 Requirements For a Mathematician's Visual System

To address this problem, I have been collecting requirements for visual mechanisms since 1971, in parallel with more general explorations of requirements for a complete human-like architecture (e.g., (Sloman, 1978, 1989, 2002, 2001a, 2001b, 2008)). Full understanding of the issues requires us to investigate: (a) trade-offs between alternative sets of requirements and designs, including different biological examples (Sloman, 2000, 2007); (b) different kinds of developmental trajectories (Sloman & Chappell, 2005; Chappell & Sloman, 2007); and (c) requirements for *internal* languages supporting structural variability and compositional semantics in other animals, pre-verbal humans, and future robots (Sloman, 1979; Sloman & Chappell, 2007). I have also tried to show how that analysis can lead to a new view of the evolution of language (http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang). In particular, it allows internal languages with compositional semantics to include *analogical* forms of representation, whose manipulation can play an important role both in visual perception and in reasoning. This underlies human spatial reasoning abilities that are often used in mathematics. Human-like mathematical machines (e.g. robots that reason as humans do) will also need such competences.

9 Affordances, Visual Servoing, and Beyond

Analysis of biological requirements for vision (including human vision) enlarges our view of the functions of vision, requiring goals of AI vision researchers to be substantially expanded. An example is the role of vision in servo-control, including control of continuous motion described in (Sloman, 1982) and Berthoz (2000), as well as discrete condition-checking.

Gibson's work on affordances in his (1979) showed that animal vision provides information not merely about geometrical and physical features of entities in the environment, as in (Barrow & Tenenbaum, 1978; Marr, 1982), nor about recognising or categorising objects (the focus of much recent AI 'vision' research), but about what the perceiver can and cannot do, given its physical capabilities and its goals. I.e. vision needs to provide information not only about actual objects, structures and motion, but also what processes *can and cannot occur in the environment* (Sloman, 1996). In order to do this, the visual system must use an ontology that is only very indirectly related to retinal arrays. But Gibson did not go far enough, as we shall see.

All this shows that a human-like visual sub-architecture must be multi-functional, with subsystems operating concurrently at different levels of abstraction and engaging concurrently with different parts of the rest of the architecture, including central and motor subsystems. For example, painting a curved stripe requires continuous visual control of the movement of the brush, which needs to be done in parallel with checking whether mistakes have been made (e.g. bits not painted, or the wrong bits painted) and whether the task has been completed, or whether the brush needs to be replenished. For these reasons, in (Sloman, 1989), I contrasted (a) "modular" visual architectures, with information flowing from input images (or the optic array) through a pipeline of distinct processing units, as proposed by Marr and others, with (b) "labyrinthine" visual architectures reflecting the multiplicity of functions of vision and the rich connectivity between subsystems of many kinds.

10 Perception of Actual and Possible Processes

Work on an EU-funded cognitive robotics project, CoSy, begun in 2004,[5] included analysis of requirements for a robot capable of manipulating 3-D objects, e.g. grasping them, moving them, and constructing assemblages, possibly while other things were happening. Analysis of the requirements revealed (a) the need for representing scene objects with parts and relationships (as everyone already knew), (b) the need for several ontological layers in scene structures (as

[5] Described in http://www.cs.bham.ac.uk/research/projects/cosy/

in chapter 9 of (Sloman, 1978)), (c) the need to represent "multi-strand relationships" because not only whole objects but also parts of different objects are related in various ways, (d) the need to represent "multi-strand processes", because when things move the multi-strand relationships change, e.g. with metrical, topological, causal, functional, continuous, and discrete changes occurring concurrently, and (e) the need to represent *possible* processes, and constraints on possible processes. I call the latter positive and negative "proto-affordances", because they are the substratum of affordances, but more general.

Not all perceived changes are produced or can be produced by the perceiver. Likewise seeing that a process is *possible*, e.g. an apple falling, or that possibilities are *constrained*, e.g. because a table is below the apple, does not presuppose that the perceiver desires to or can produce the process. So perception of proto-affordances and perception of processes in the environment makes it possible to take account of far more than one's own actions, their consequences and their constraints. As explained in (Sloman, 2009, 2008), that requires an *amodal, exosomatic* form of representation of processes; one that is not tied to the agent's sensorimotor processes. That possibly is ignored by researchers who focus only on sensorimotor learning and representation, and base all semantics on "symbol-grounding".[6]

The ability to perceive a multi-strand process requires the ability to have internal representations of the various concurrently changing relationships. Some will be continuous changes, including those needed for servo-control of actions, while others may be discrete changes as topological relations change or goals become satisfied. Mechanisms used for perceiving multi-strand processes can also be used both to predict outcomes of possible processes that are not currently occurring (e.g. when planning), and to explain how a perceived situation came about. Both may use a partial simulation of the processes.[7] (Cf. Grush (2004).)

11 The Importance of Kinds of Matter

A child, or robot, learning about kinds of process that can occur in the environment needs to be able to extend the ontology she uses indefinitely, and not merely by defining new concepts in terms of old ones: there are also *substantive* ontology extensions (as in the history of physics and other sciences). For example, whereas many perceived processes involve objects that preserve all their metrical relationships, there are also many deviations from such rigidity, and concepts of different kinds of matter are required to explain those deviations: string and wire are flexible, but wire retains its shape after being deformed; an elastic band returns to its original length after being stretched, but does not restore its shape after bending. Some kinds of stuff easily separate into chunks in various ways, if pulled, e.g. mud, porridge, plasticine and paper. A subset of those allow restoration to a single object if separated parts are pressed together. There are also objects that are marks on other objects, like lines on paper, and there are some objects that can be used to produce such marks, like pencils and crayons. Marks produced in different ways and on different materials can have similar structures.

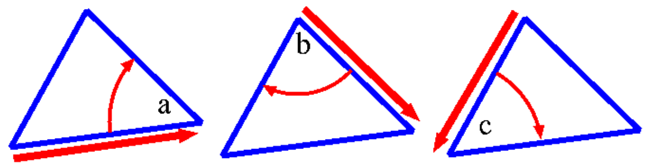
As demonstrated by Sauvy and Sauvy (1974), children, and presumably future robots, can learn to play with and explore strings, elastic bands, pieces of wire, marks on paper and movable objects, thereby learning about many different sorts of process patterns. Some of those are concerned with rigid motions some not. Some examples use patterns in non-rigid motions that can lead to development of topological concepts, e.g. a cup being continuously deformed into a toroid. Robot vision is nowhere near this capability at present.

[6] Reasons for preferring "symbol-tethering" to symbol-grounding theory are given in:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models

[7] Examples are given in the presentation in Note 1, and in this discussion paper on predicting changes in action affordances and epistemic affordances:
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0702

12 Perception and Mathematical Discovery

I have argued that many mathematical discoveries involve noticing an invariant in a class of processes. For example, Mary Pardoe[8], a mathematics teacher, once told me she had found a good way to teach children that the internal angles of a triangle add up to a straight line, demonstrated in the figure.

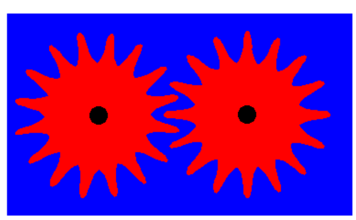


Consider any triangle. Imagine an arrow starting at one corner, pointing along one side. It can be rotated and translated as indicated, going through positions shown in the three successive figures. The successive rotations a, b and c go through the interior angles of the triangle, and because of the final effect they produce, they must add up to a straight line. This discovery may initially be made through empirical exploration with physical objects, but the pattern involved clearly does not depend on what the objects are made of and changing conditions such as colours used, lengths of lines, particular angles in the triangle, temperature, strength of gravitational or magnetic field cannot affect the property of the process. A robot learner should notice that it is not an empirically falsifiable discovery.

However, such discoveries can have “bugs” as Lakatos (1976) demonstrated using Euler’s theorem about polyhedra. That is sometimes wrongly taken to imply that mathematical knowledge is empirical in the same way as knowledge about the physical properties of matter. The discovery of bugs in proofs and good ways to deal with them is an important feature of mathematical learning. For example, the rotating arrow proof breaks down if the triangle is on the surface of a sphere. Noticing this can lead a learner to investigate properties that distinguish planar and non-planar surfaces. But that exploration does not *require* experiments in a physical laboratory, though it may benefit from them. Kant claimed that such discoveries are about the perceiver’s forms of perception, but they are not restricted to any particular perceivers.

13 Humean and Kantian Causation

Adding properties of matter, such as rigidity and impenetrability, to representations of shape and topology allows additional reasoning about and prediction of results of processes. An example is the ability to use the fact that two meshed gear wheels are rigid and impenetrable to work out how rotating one will cause the other to rotate. That kind of reasoning is not always available.



If the wheels are not meshed, but there are hidden connections, then the only basis for predicting the consequence of rotating the wheels is to use a Humean notion of causation: basing predictions of results of actions or events solely on observed correlations. In contrast, where the relevant structure and constraints are known, and mathematical proofs (using geometry and topology) are possible, Kant’s notion of causation, which is structure-based and deterministic, can be used. Causal relationships represented in Bayesian nets are essentially generalisations of Humean causation and based only on statistical evidence. However, a significant subset of the causal understanding of the environment that a child acquires is Kantian

[8] Mary Ensor at the time

because it allows the consequences of novel processes to be *worked out* on the basis of geometric and topological relationships, and kinds of matter involved. For more on Humean vs Kantian causation in robots and animals see the presentations by Sloman and Chappell here: http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac

Many kinds of learning involve strings. If an inelastic but flexible string is attached to a remote movable object, then if the end is pulled away from the object a process can result with two distinct phases: (1) curves in the string are gradually eliminated (as long as there are no knots), and (2) when the string is fully straightened the remote object will start moving in the direction of the pulled end. However, if the string is looped round a fixed pillar, the first sub-process does not produce a single straight string but two straight portions and a portion going round the pillar, and in the second phase the attached object moves toward the pillar, not toward the pulled end.[9]

14 Russell *vs* Feynman on Mathematics

At the beginning of the last century Russell and Whitehead (1910–1913) attempted to demonstrate that all of mathematics could be reduced to logic (Frege had attempted this only for Arithmetic). Despite the logical paradoxes and the difficulty of avoiding them, Russell thought the goals of the project could be or had been achieved, and concluded that mathematics was just *the investigation of implications that are valid in virtue of their logical form, independently of any non-logical subject matter*. He wrote: "Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true" (*Mysticism and Logic* (Russell, 1917)). In some ways Russell seems to have been a philosophical descendant of David Hume, who had claimed that non-empirical propositions were in some sense trivial, e.g. mere statements of the relations between our ideas.

In contrast, the physicist Richard Feynman described mathematics as "the language nature speaks in". He wrote: "To those who do not know Mathematics it is difficult to get across a real feeling as to the beauty, the deepest beauty of nature. ... If you want to learn about nature, to appreciate nature, it is necessary to understand the language that she speaks in" (in (Feynman, 1965)). I believe that Feynman's description is closely related to what I am saying about how a child (or a future robot) can develop powerful, reusable concepts and techniques related to patterns of perception and patterns of thinking that are learnt through interacting with a complex environment, part of which is the information-processing system within the learner. Despite the role of experience in such learning, the results of such learning are not empirical generalisations. Feynman seems to agree with Kant that mathematical knowledge is both non-empirical and deeply significant.

15 Conclusion

We need further investigation of architectures and forms of representation that allow playful exploration by a robot to produce discoveries of patterns in structures and processes that are the basis of deep mathematical concepts and mathematical forms of reasoning. The robot should be able to go through the following stages:

1. Acquiring familiarity with some domain, e.g. through playful exploration;
2. Noticing (empirically) certain generalisations;
3. Discovering a way of thinking about them that shows they are not empirical;
4. Generalising, diversifying, debugging, deploying, that knowledge;
5. Formalising the knowledge, possibly in more than one way.

[9] More examples and their implications are discussed in the presentation in Note 1 and in http://www.cs.bham.ac.uk/research/projects/cosy/papers#dp0601: "Orthogonal recombinable competences acquired by altricial species".

This paper merely reports on a subset of the requirements for working designs. Some more detailed requirements are in (Sloman, 2009). It is clear that AI still has a long way to go before the visual and cognitive mechanisms of robots can match the development of a typical human child going through the earlier steps. There is still a great deal more to be done, and meeting all the requirements will not be easy. If others are interested in this project, perhaps it would be useful to set up an informal network for collaboration on refining the requirements and then producing a working prototype system as a proof of concept, using a simulated robot, perhaps one that manipulates 2-D shapes in a plane surface, discovering properties of various kinds of interactions, involving objects with different shapes made of substances with various properties that determine the consequences of the interactions, e.g. impenetrability, rigidity, elasticity, etc.

Perhaps one day, a team of robot mathematicians, philosophers of mathematics and designers will also be able implement such systems. First we need a deeper understanding of the requirements.

Acknowledgements

Some of the ideas reported in this paper were developed as part of the requirements analysis activities in the EU-funded CoSy project http://www.cognitivesystems.org, and discussions with Jackie Chappell. Work by Jim Franklin referenced in the presentation in Note 1 is also relevant. Some related ideas are in early work by Doug Lenat, e.g. AM and Eurisko, (Lenat & Brown, 1984).

References

Barrow, H., & Tenenbaum, J. (1978). Recovering intrinsic scene characteristics from images. In A. Hanson & E. Riseman (Eds.), *Computer vision systems* (pp. 3–26). New York: Academic Press.

Berthoz, A. (2000). *The brain's sense of movement.* London, UK: Harvard University Press.

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

Feynman, R. (1965). *The Character of Physical Law.* Cambridge MA: MIT Press. (The 1964 Messenger Lectures)

Frege, G. (1950). *The Foundations of Arithmetic: a logico-mathematical enquiry into the concept of number.* Oxford: B.H. Blackwell. ((Tr. J.L. Austin. Original 1884))

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Glasgow, J., Narayanan, H., & Chandrasekaran, B. (Eds.). (1995). *Diagrammatic reasoning: Computational and cognitive perspectives.* Cambridge, MA: MIT Press.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377–442.

Heyting, J. (1956). *Intuitionism, an Introduction.* Amsterdam: North Holland.

Jablonka, E., & Lamb, M. J. (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life.* Cambridge MA: MIT Press.

Kant, I. (1781). *Critique of pure reason.* London: Macmillan. (Translated (1929) by Norman Kemp Smith)

Lakatos, I. (1976). *Proofs and Refutations.* Cambridge, UK: Cambridge University Press.

Lenat, D. B., & Brown, J. S. (1984). Why AM and EURISKO appear to work. *Artificial Intelligence*, *23*(3), 269–294.

Liebeck, P. (1984). *How Children Learn Mathematics: A Guide for Parents and Teachers.* Harmondsworth: Penguin Books.

Liebeck, P. (1990, Jun). Scores and Forfeits: An Intuitive Model for Integer Arithmetic. *Educational Studies in Mathematics*, *21*(3), 221–239.

Marr, D. (1982). *Vision.* San Francisco: W.H.Freeman.

McCarthy, J., & Hayes, P. (1969). Some philosophical problems from the standpoint of AI. In B. Meltzer & D. Michie (Eds.), *Machine Intelligence 4* (pp. 463–502). Edinburgh, Scotland: Edinburgh University Press. (http://www-formal.stanford.edu/jmc/mcchay69/mcchay69.html)

Mill, J. S. (1843). *A System of Logic, Ratiocinative and Inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*. London: John W. Parker.

Penrose, R. (1989). *The Emperor's New Mind: Concerning Computers Minds and the Laws of Physics*. Oxford: Oxford University Press.

Rips, L. J., Bloomfield, A., & Asmuth, J. (2008). From Numerical Concepts to Concepts of Number. *The Behavioral and Brain Sciences*, *31*(6), 623–642. (doi:10.1017/S0140525X08005566)

Russell, B. (1903). *The principles of mathematics*. Cambridge, UK: Cambridge University Press.

Russell, B. (1917). *Mysticism and logic and other essays*. London: Allen & Unwin.

Sauvy, J., & Suavy, S. (1974). *The Child's Discovery of Space: From hopscotch to mazes – an introduction to intuitive topology*. Harmondsworth: Penguin Education. (Translated from the French by Pam Wells)

Sloman, A. (1962). *Knowing and Understanding: Relations between meaning and truth, meaning and necessary truth, meaning and synthetic necessary truth*. Unpublished doctoral dissertation, Oxford University. (http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#706)

Sloman, A. (1965). 'Necessary', 'A Priori' and 'Analytic'. *Analysis*, *26*(1), 12–16. (Now online http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#701)

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

Sloman, A. (1978). *The computer revolution in philosophy*. Hassocks, Sussex: Harvester Press (and Humanities Press). (http://www.cs.bham.ac.uk/research/cogaff/crp)

Sloman, A. (1979). The primacy of non-communicative language. In M. MacCafferty & K. Gray (Eds.), *The analysis of Meaning: Informatics 5 Proceedings ASLIB/BCS Conference, Oxford, March 1979* (pp. 1–15). London: Aslib. (http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#43)

Sloman, A. (1982). Image interpretation: The way ahead? In O. Braddick & A. Sleigh. (Eds.), *Physical and Biological Processing of Images (Proceedings of an international symposium organised by The Rank Prize Funds, London, 1982.)* (pp. 380–401). Berlin: Springer-Verlag. (http://www.cs.bham.ac.uk/research/projects/cogaff/06.html#0604)

Sloman, A. (1989). On designing a visual system (towards a gibsonian computational model of vision). *Journal of Experimental and Theoretical AI*, *1*(4), 289–337. (http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#7)

Sloman, A. (1996). Actual possibilities. In L. Aiello & S. Shapiro (Eds.), *Principles of knowledge representation and reasoning: Proceedings of the fifth international conference (kr '96)* (pp. 627–638). Boston, MA: Morgan Kaufmann Publishers. (http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15)

Sloman, A. (2000). Interacting trajectories in design space and niche space: A philosopher speculates about evolution. In *et al.*. M.Schoenauer (Ed.), *Parallel problem solving from nature – ppsn vi* (pp. 3–16). Berlin: Springer-Verlag.

Sloman, A. (2001a). Beyond shallow models of emotion. *Cognitive Processing: International Quarterly of Cognitive Science*, *2*(1), 177-198.

Sloman, A. (2001b). Evolvable biologically plausible visual architectures. In T. Cootes & C. Taylor (Eds.), *Proceedings of British Machine Vision Conference* (pp. 313–322). Manchester: BMVA.

Sloman, A. (2002). Architecture-based conceptions of mind. In P. Gärdenfors, K. Kijania-Placek, & J. Woleński (Eds.), *In the Scope of Logic, Methodology, and Philosophy of Science (Vol II)* (pp. 403–427). Dordrecht: Kluwer. (http://www.cs.bham.ac.uk/research/projects/cogaff/00-02.html#57)

Sloman, A. (2007). Diversity of Developmental Trajectories in Natural and Artificial Intelligence. In C. T. Morrison & T. T. Oates (Eds.), *Computational Approaches to Representation Change during Learning and Development. AAAI Fall Symposium 2007, Technical Report FS-07-03* (pp. 70–79). Menlo Park, CA: AAAI Press. (http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0704)

Sloman, A. (2008). Putting the Pieces Together Again. In R. Sun (Ed.), *Cambridge Handbook on Computational Psychology* (pp. 684–709). New York: Cambridge University Press. (http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#710)

Sloman, A. (2009). Architectural and representational requirements for seeing processes and affordances. In D. Heinke & E. Mavritsaki (Eds.), *Computational Modelling in Behavioural Neuroscience: Closing the gap between neurophysiology and behaviour.* (pp. 303–331). London: Psychology Press. (http://www.cs.bham.ac.uk/research/projects/cosy/papers#tr0801)

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

Sloman, A., & Chappell, J. (2007). Computational Cognitive Epigenetics (Commentary on (Jablonka & Lamb, 2005)). *Behavioral and Brain Sciences*, *30*(4), 375–6. (http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0703)

Sussman, G. (1975). *A computational model of skill acquisition.* San Francisco, CA: American Elsevier. (http://dspace.mit.edu/handle/1721.1/6894)

Whitehead, A. N., & Russell, B. (1910–1913). *Principia Mathematica Vols I – III.* Cambridge: Cambridge University Press.