Contents

(Table of contents for pre-publication version only.)

DRAFT: Revised February 16, 2014

NOTE TO COPY EDITORS

Please make NO changes unless you have first checked with me (by email to a.sloman@cs.bham.ac.uk), including spelling and punctuation changes. If such changes are made without prior consent I may withdraw this paper. I can't spare the time to search for and fix copy-editing mistakes – and in any case am not good at finding them all.

For my reasons, please see: `http://www.cs.bham.ac.uk/~axs/publishing.html`

Evolution of Geometrical Reasoning

Aaron Sloman

Invited contribution to: The Incomputable
Eds Mariya Soskova and S Barry Cooper
Previous title: "Incomputability of Geometrical Reasoning"

Abstract Revised 28 Dec 2013:
The Turing-inspired Meta-Morphogenesis project links ideas about the nature of mathematics, as containing infinitely many domains of possible structures and processes, with ideas about how evolution works: by (a) making mathematical discoveries implicitly in solutions to biological design problems, then later (b) providing new species with mechanisms for (implicitly) making such discoveries themselves, then later (c) adding meta-cognitive mechanisms that allow individual organisms to think about, and then later on discuss and share, what they have learnt, or to have their learning enhanced by what others have learnt. There are many intermediate steps, still mostly unknown. Philosophy of mathematics normally focuses only on a few aspects of very late stages in the process, whereas evolution and mathematical structures are deeply intertwined from the earliest stages. Explaining how that is possible seems to require a combination of Turing's ideas about discrete computation and his ideas about chemistry-based morphogenesis. The outcomes include competences produced by both biological evolution and individual learning and development, that are important for perceiving, understanding and using affordances in the environment (a point made by James Gibson, though he focused on a narrow subset). I conjecture that such competences were essential precursors to the mathematical discoveries systematised in Euclidean geometry (a point Gibson did not notice, as far as I know). A consequence is that normal humans are far more mathematically sophisticated than they know or show, as are many other species. Some of the competences have requirements that so far have not been met in AI models, and it is not obvious how they could be implemented using current models of computation, even though there are superficially similar achievements. Is there something about those pre-Euclidean geometrical and topological reasoning competences that Turing-equivalent machines cannot support, and if so could chemistry-based computation do so? Or have we merely failed so far to devise appropriate computational architectures, with the right mixture of spatial reasoning and meta-cognition?

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

1 Turing's largely unnoticed observation

"In the nervous system chemical phenomena are at least as important as electrical." So wrote Alan Turing, in one of his most widely cited papers (Turing, 1950), though the comment has largely gone unnoticed. In a paper published soon after that, (Turing, 1952), he began to illustrate possible functions for chemical control mechanisms in the formation of physical patterns and structures in organisms. I suspect that if he had lived longer he would have developed theories about forms of computation combining the resources of discrete and continuous mechanisms. Chemical processes provide an enormous variety of mixtures of discrete and continuous interaction. Perhaps such mixtures were essential for the evolution of mathematical and proto-mathematical competences in living organisms. And perhaps our future intelligent robots will need such mixtures. This paper presents some of the groundwork for future research on such possibilities, focusing mainly on the nature of mathematics and information-processing requirements for the modes of discovery of mathematical truths by our ancestors and simpler versions found in other animals and young children. The question whether Turing machines provide an adequate basis for modelling and replicating animal and human intelligence, and if not whether chemical information processing machines would suffice, as suggested by the history of the "bootstrapping" of life and mind by evolution, is raised but not answered. My arguments do not depend on incompleteness theorems, only some peculiarities of geometrical reasoning, for example in Euclid's *Elements*. But the investigation depends in part on answering the question "What is mathematics?"

2 What is mathematics?

Many answers have been offered, by philosophers and others. Wittgenstein wrote: "For mathematics is after all an anthropological phenomenon" (Wittgenstein (1978) Part VII section 33), and there seem to be many who agree with him. I shall try to show that the implied claim that mathematics is a creation of human minds (either as individuals or as members of a culture) is false, mainly because there are many mathematical domains whose existence has nothing to do with the existence of humans.[1] There are different sorts of existence illustrated by electrons, New York, the last day of my life, complex geometric shapes, shape-deformation trajectories, prime numbers, infinite subsets of the natural numbers, infinite dimensional vector spaces, proofs, conjectures, and logical contradictions, among many others. At most three of those depend on the existence of humans, or mathematicians of any sort.

The dependence is mainly the other way round: the existence of mathematical domains is required for thinkers to formulate conjectures, discover theorems, create or modify proofs, and so on. For some that's an uncontroversial claim. But I shall go

[1] What "exists" means is a topic for another occasion.

much further and try to show that long before humans existed biological evolution (unwittingly) made use of mathematical domains that were (unwittingly) discovered and used either by natural selection or by its products (e.g. individual organisms, or groups of organisms). For example, homeostatic control mechanisms, found in all organisms, used for many control functions, exemplify the mathematical structure of a negative feedback loop (which admits many mathematically distinct variants), evolved long before human engineers discovered their importance, and human mathematicians studied their properties. While writing this paper I discovered that Popper had expressed related ideas about biological organisms and evolutionary processes as problem solvers (Popper, 1984) though he used different terminology and made no mention of computer models of mathematical discovery and reasoning.

Later, discovery and use of various subsets of mathematics by individuals (illustrated below) must have preceded the social activities leading to explicit systematisation and published proofs of such mathematical discoveries, spectacularly demonstrated in Euclid's *Elements*, written over two millennia ago. There was also much mathematical work done in other ancient cultures.[2]

There are some mathematical domains that are intrinsically related to human activities, for example the domain of currency conversions between British pounds and US dollars at different times. Such domains, however, are special cases of more abstract domains of mappings between sets of numbers, whose existence has nothing to do with humans. The same can be said of translational mappings between sentences in human languages. The rest of this paper is about mathematical domains whose existence does not depend on human activities, though their discovery and use by humans does, of course. There must be infinitely many that have not yet been discovered or used by humans, most of which never will be. Some of them were used by evolution before humans existed.

3 Patterns of information and control

It is generally acknowledged that evolution and its products must use information to control reproduction, including: information about physical forms of individual organisms; information about physical structures underlying those forms (bones, tendons, blood vessels, organs, etc.); information about initiation and control of physical behaviours of whole organisms (exercising limbs, crawling, walking, running, etc.); information about physical and chemical behaviours of internal mechanisms (digestion, respiration, repair of damage, etc.); and information about features of the environment with which individuals must interact (e.g. sensory information used in young mammals to trigger and control sucking, very soon after birth). In all those cases there are mathematical domains with biological instances. In many cases there are different specialisations of the same domain, e.g. with

[2] http://www.math.tamu.edu/~dallen/history/euclid/euclid.html,
http://en.wikipedia.org/wiki/Chinese_mathematics,
http://en.wikipedia.org/wiki/Indian_mathematics

variations across individuals in the same and different species, and variations across developmental processes of individuals, an example being patterns of control of movements during changes of size, weight, shape, strength, and types of motion as an individual develops. Evolution clearly "discovered" mechanisms that can adjust themselves to accommodate such variations, mechanisms studied both in genetics and in burgeoning research on epigenetic processes where genes, environment, and developing structures interact.

What has not received so much attention in the past is the importance of information *about information*, e.g. how to represent information, how to acquire it, how to store it, how to derive new information from old, and now to use it in a variety of processes including control of behaviour and control of information-processing. These competences address a new class of domains, domains of information structures, domains of processes of information manipulation, domains of control processes, and many more. There are now many research communities working on different subsets of biological information processing. But biological evolution has to deal with all of them, and re-use of information about mathematical commonalities is a requirement for avoiding an intractable explosion of problems to be solved. The requirement to be able to re-use generic design information with variation in details has long been understood by engineers, especially software engineers, and played a major role in the development of programming languages that support re-use, especially languages misleadingly labelled "object oriented".[3]

In making use of biologically important mathematical domains, evolution developed increasingly sophisticated forms of representation (i.e. structures for encoding information), mechanisms for using information, and architectures for combining information-processing mechanisms and functions within an organism. Examples include sensing mechanisms, mechanisms for storing sensed information for varying time-scales for varying purposes, mechanisms for comparing information acquired at different times or different places (needed for feedback control, for executing plans, for formulating and answering questions or testing hypotheses), mechanisms for selecting between possible actions on the basis of comparisons, and mechanisms for controlling the conversion of stored energy (e.g. chemical energy) into other forms of energy (e.g. mechanical energy) in ways that depend on current information structures (motives, beliefs, constraints, etc). See Ganti (2003) for more details concerning such processes in the simplest organisms.

Biological information processing involves complex controlled interactions, as opposed to passive storage. The ability of mud to record a footprint illustrates very primitive passive information-processing, but the mud includes no mechanisms that systematically relate a variety of information contents to a variety of needs, goals, constraints, and learning processes. Mud has no needs or goals, nor such mechanisms. So although the mud can be considered as a sort of limiting case (as a circle is a limiting case of an ellipse), the processes are not rich and varied enough to be describable as being related non-trivially to mathematical domains, unlike the use

[3] See `http://www.cs.bham.ac.uk/research/projects/poplog/teach/oop`

of information by an insect, a plant or even a microbe that feeds, excretes, moves and reproduces. (Contrast the uses of mud records by geologists and palaeontologists.)

The vast majority of such evolved organisms had no way of acquiring information about what information-processing they were doing, how they were doing it, what the alternatives were, etc. That required the development of meta-cognitive sub-architectures capable of acquiring, transforming, interpreting, storing and using information *about* information processing. For reasons discussed in (Sloman and Chrisley, 2003; Sloman, 2010), it seems that discoveries made by human engineers in the twentieth century concerning the use of *virtual* machines as well as *physical* machinery were *implicitly* made millions of years earlier by evolution, insofar as organisms began to evolve subsystems in which non-physical processes, such as inference, were *implemented* in physical processes, and then later on self-monitoring and self-modifying meta-cognitive mechanisms in new virtual machines acquired the ability to observe and modify some of the processes in pre-existing virtual machines, for instance noticing that a planning procedure constantly produces plans that don't work unless modified, which might enable the planning procedure to be modified (as demonstrated in (Sussman, 1975) and other AI experiments in the early 1970s).

The hypothesis of evolution of interacting virtual machines within organisms was used in (Sloman, 2010), to explain (in outline) the existence of qualia, and the partial introspectability of qualia, in ways that contradict many philosophical and non-philosophical theories about consciousness, requirements for its existence, and its evolution. For example, Sloman (2009a) demonstrates the possibility of having visual qualia that are stored yet temporarily inaccessible, until an appropriate probing question re-directs attention to the stored information.

These information processing mechanisms and architectures must have evolved and been used long before any humans thought about them, and in some cases before humans existed, e.g. information-processing mechanisms enabling now long extinct carnivores to see, catch, dismember, and eat portions of their prey.

Although the mechanisms *made use of* properties of mathematical domains there is no reason to believe that anyone or anything ever described or discussed such domains before humans did – at least on this planet. There seem to be partly analogous situations in pre-verbal humans discovering and using what can be called "toddler theorems"[4] about what is and is not possible in their world – for example: many children seem to discover and understand the differences between grasps that do, and grasps that do not, have painful effects when pushing drawers or doors shut; and most sighted children discover and understand relationships between availability of information and unobstructed line of sight, which they use in moving themselves or other objects in order to gain information, and at a later stage in order to make information accessible or inaccessible to others. This involves a domain of types of visual information transfer, where the information travels in straight lines between surfaces of objects and eyes of perceivers. For more on this see Section 24 and Figure 4.

---

[4] `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html`

4 Toddler theorems

In Sauvy and Sauvy (1974) games and exploratory activities are described in which young children play with different sorts of objects and materials in the environment, implicitly discovering their mathematical properties as well as making empirical discoveries, e.g. that some materials are rigid some not, and that some non-rigid objects are elastic, others not. A topic worth investigating, if it has not yet been investigated, is whether any pre-verbal children are able to discover and understand differences between contexts where pulling two ends of a string apart do and do not produce a knot. Such competences may be developed and demonstrated in various games and competitions.[5] Two books by Piaget explore differences in abilities, in children of a variety of ages, to reason about what is or is not possible and why (Piaget, 1981, 1983).

Some of the discoveries made by children (and possibly some other animals) go beyond empirically learnt generalisations, and include understanding why some things *must* be the case and why others *cannot* be the case. If your fingers are curled round the edge of a door you *cannot* push the door shut without causing fingers to be compressed between door edge and door jamb. A child may discover that the effect is inevitable unless the hand pushing the door is re-positioned, e.g. pushing with a flat palm. Noticing the necessary connection allows the learner to avoid wasting time trying minor variants of the curled grip, which might be a consequence of purely empirical learning. Discovering the invariant (the necessary connection) requires a meta-cognitive ability to reflect on what has been learnt and what it depends on: which might be described as *proto-mathematical* a competence. At present, I don't think anyone knows what the mechanisms making such competences possible are. (Compare discussion of "representational re-description" in Karmiloff-Smith (1992).)

Getting beyond the "proto" stage requires being able to think about or talk about differences between ways in which relationships between features of a situation do or do not have *invariant* connections. This requires a sort of *meta-semantic* competence: the ability to detect and use relationships between semantic contents. Such a meta-semantic ability can be present and be used without any meta-cognitive ability to detect, think about, talk about, the fact that it is used. That later meta-meta-cognitive ability is required for detecting and correcting flaws in one's reasoning, which may develop in parallel with abilities to detect and comment on flaws in the reasoning of other individuals. All of these processes seem to start, at least in some children, before anyone attempts to teach them mathematics, though teaching done well can build on and enrich the mechanisms. Only much later do most children learn (usually at school) to think and talk about their mathematical thinking.

Unfortunately few teachers (or developmental psychologists, or parents) seem to be aware of these processes. Moreover, the mechanisms are so complex, with so many different possible developmental trajectories, that we must not expect any

[5] `http://www.pitara.com/activities/craft/online.asp?story=45` is an example concerning making knots in an unusual way.

of the most interesting facts about human minds to take the form of universal generalisations ("laws"), e.g. about what all children do, or about what happens at a particular age, or about the order in which competences develop. Rather: the deepest psychology is about the variability in what *can* occur and what sorts of *generative* mechanisms make all those trajectories possible. The Meta-Morphogenesis project includes the long term goal of relating these processes and mechanisms to earlier stages in evolution and to the information-processing competences and mechanisms of other species. This will help to fill the gaps in the theories about meta-configured vs pre-configured competences sketched in Chappell and Sloman (2007), and perhaps explain the different genetic contributions to achievements of great mathematicians, poets, artists, scientists, musical performers, Olympic skiers and sumo wrestlers.

For reasons discussed below, Artificial Intelligence is nowhere near this, despite some spectacular successes. Perhaps one day intelligent robots will be able to follow human-like developmental trajectories. At present even the most impressive computer-based mathematical reasoners and the most widely used learning mechanisms in AI systems, do not do what developing human mathematicians do, namely extend their processing powers, driven in part by types of complexity found in the environment, in part by what they have previously learnt, and in part by needs that they may encounter later, as opposed to learning to get immediate rewards or avoid immediate punishments. In particular all the powerful automated theorem provers that I know of are designed to be tools for humans to use, rather than products of a learning, developing machine with complex and changing needs and goals getting to grips with a complex and changing environment, in which many mathematical domains are instantiated, including geometric domains.

5 Human engineering lags behind and jumps ahead of evolution

Human engineers began to understand these information processing requirements (still largely ignored by most philosophers) only in the 20th Century, millions of years after the implicit discoveries made by evolution.[6] These included "discoveries" concerning new forms of self-monitoring, using meta-cognitive, mechanisms, without which the need for proofs and the proofs themselves could not have been discovered – and were not discovered by non-human animals (on this planet). Unless we understand all this we shall not be able to build intelligent machines (especially autonomous robots) with a wide range of human capabilities; and research in neuroscience, developmental and other branches of psychology, educational theory and philosophy, including philosophy of mind, language and mathematics, will continue to miss key problems and possible explanations.

---

[6] I am not using "information" in the sense of Shannon, but the old concept of "information" understood much earlier by Jane Austen, for example, illustrated in `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/austen-info.html`

5.1 *Kant on mathematics*

Immanuel Kant (1781) argued that mathematical knowledge, including knowledge of geometry and arithmetic, is non-empirical, is synthetic (i.e. not only logical consequences of definitions), and is about necessary truths. Gottlob Frege (1950) agreed with Kant about geometry but claimed to have shown that arithmetical truths are all analytic – they are all logical truths. There have been many refinements of the debate since the time of Frege, but it seems to be widely believed that Kant was shown to be wrong about arithmetic because Frege, Russell and others demonstrated that arithmetic (or at least the decidable subset) is based only on logic and definitions, and wrong about geometry, either because Euclidean geometry was shown by Einstein to be empirical and false, or because geometrical knowledge is only knowledge about which theorems follow from various sets of geometrical axioms. (I have presented a very short summary of long and complicated story.)

However, I think there remains a core of truth in Kant's ideas, especially in the context of trying to explain how humans first made mathematical discoveries and found ways of proving results so as to make empirical tests unnecessary, long before anyone had developed formal logical theories or investigated axiomatic systems. If the ideas presented here are correct then we should be able to create "baby" robots that learn about their environment in something like the way humans and other animals do, then discover new deep structures in their knowledge as only humans seem able to do on this planet. I am sure that it must be possible to replicate those processes, but it may be necessary to extend current forms of computation, perhaps by using chemical information processing, as organisms do. The need for that seems to have been anticipated by Turing in the extract from Turing (1950) at the beginning of this paper. However, this is still a tentative conjecture supported by inconclusive arguments, below.

I shall later offer reasons for thinking that all of these considerations provide reasons for thinking that Kant was right about both arithmetic and geometry as being mathematical domains in which synthetic necessary truths can be discovered using non-empirical mechanisms. However, I shall first have to say more about the nature of mathematics.

Many details of Kant's thinking may need to be re-assessed, in the light of recently understood requirements for information-processing mechanisms able to support or replicate human mathematical competences. These topics are very deep and complex, combining strands from several disciplines that are not normally brought together, and there are many unsolved problems.

5.2 *The biological significance of biological philosophy of mathematics*

I am trying to assemble a collection of observations and hypotheses about evolution, life, mind, and mathematics, that are conjectured to have great potential in combination, though each may be considered flimsy and unclear on its own.

Whether this turns out to be a progressive or a degenerating research programme (in the sense described by (Lakatos, 1980) and elaborated in (Sloman, 1978a, Chap 2)) remains to be seen.

I first tried to present and defend Kant's ideas about mathematics in (Sloman, 1962) but at that time knew nothing about computation or artificial intelligence, and the work suffered as a consequence (like much philosophical writing since then). Nine years later, I began to learn about programming and AI but progress remained very slow and piecemeal. Reading Turing (1952) in 2011 stimulated a revised approach, within the Turing-inspired "Meta-Morphogenesis" project first proposed in Sloman (2013b). This chapter focuses on a small part of one strand of that project, concerned with the nature of mathematics, the variety of forms of mathematical competence, and the evolution of such competences. But the implications go far beyond the nature of mathematics.

Relocate

In particular, it seems that despite the spectacular power of Turing machinery and the technology inspired by the theory of such machines, there may be something missing, which Turing began to explore in (Turing, 1952) shortly before his tragic death, although I do not know whether he made, or was about to make, the connections with philosophy of mathematics sketched in this paper.

This work seems to be contradicting a thesis expressed by Philip Welch in a Turing centenary lecture in 2012: "Anything that is humanly calculable is computable by a Turing machine", especially if "humanly calculable" includes geometrical reasoning as found in Euclid's Elements.

But before addressing that I would like to present some of Frege's ideas.

6 Mixed messages from Frege

Gottlob Frege's comment *"There is nothing more objective than the laws of arithmetic"* (Frege, 1950, § 105), is more accurate than Wittgenstein's suggestion, though I'll be disagreeing with some of his claims. Moreover, to arithmetic I would add logic, topology, geometry, set theory, and many other mathematical domains. I think Kant (1781) was right to claim that there are branches of mathematics, including geometry, that are not reducible to logic and definitions. The truths of such mathematical domains could be described as non-analytic (i.e. synthetic), necessary (non-contingent), and *a priori* (non-empirical).[7]

Frege argued that Kant was right about geometry but wrong in claiming that arithmetical truths are synthetic – i.e. not reducible to Logic. He attempted to show that all the concepts of arithmetic could be defined using only logical concepts and constructs, and all the truths proved using only logic. Other philosophers and mathematicians (e.g. Bertrand Russell some of the time) thought that geometry could also be reduced to logic, e.g. if every geometrical theorem GT is implicitly of the form "If *Axioms* then *GT*" (Russell, 1917). David Hilbert (Hilbert, 2005), and others produced such axiomatisations of Euclidean geometry. But if the axioms and theorems make use of non-logical, but undefined, symbols, such as "point", "line", "lies on", "intersection", then it can be argued, and was argued in (Frege, 1950), that those are not theorems about *geometry*, since they are not about points, lines, planes, circles, and so on, but about some hypothetical logically specified, type of subject matter and it remains to be proved that geometrical entities and relationships, such as points, lines, and incidence, satisfy the axioms.

Frege (*op. cit*) defended against a similar objection, by avoiding use of undefined symbols, and attempting to show that all the concepts of arithmetic (e.g. "0", "1", "+", "−" and "="), could be defined using only logical concepts and constructions, which he thought was not possible for the concepts of geometry, concluding that Kant was wrong about arithmetic, but right about geometry.

He seems not to have considered an objection that could be made his programme, namely that even if he was able to demonstrate that there is a subset of logic that is isomorphic to arithmetic, this is analogous to Descartes' demonstration that there is a subset of arithmetic, including ordered pairs or triples of numbers, and equations relating different pairs or triples of numbers, that is isomorphic to Euclidean geometry. Descartes did not show that geometry is arithmetic, merely that arithmetic can model geometry. Likewise geometry can model arithmetic, a topic not discussed here. In fact, as noted in (Sloman, 1962) even logic is in some sense modelled in geometry insofar as logicians use geometric patterns and operations on geometric patterns when stating and proving logical theorems, as was done long before a domain of operations on bit patterns was used to model a great deal of logical reasoning, in computers.

Frege went to enormous lengths to demonstrate that the concepts of arithmetic were all definable in terms of pure logic and all the truths could be proved using only logical truths and logical inferences. In order to achieve this he produced new analyses of familiar mathematical concepts (including concepts of particular numbers, 0, 1, 2, etc., of number in general, of arithmetical operations, e.g. addition and multiplication), and new purely logical proofs of previously known arithmetical truths. He also extended the content of what had previously been thought of as the scope of logic, for example by generalising the mathematical concept of "function". This concept had previously been used only to refer to mathematical functions

[7] I first attempted to show that Kant was right in my DPhil thesis (Sloman, 1962), though at that time I had never heard of Artificial Intelligence, which I now regard as central to the explication of what Kant meant, as suggested in (Sloman, 1978a).

applied to mathematical objects such as numbers. Frege showed how concepts could be thought of as functions from objects (or collections of objects) of any kind, or from functions, to truth-values. The last claim introduced higher-order functions: functions whose arguments and results could be functions.

This showed how certain logical forms of expression that had been studied at least since Aristotle could be construed as using "higher order" logical functions (i.e. the universal and existential quantifiers) – though some of his logical translations of familiar natural language expressions remain controversial.[8]

Frege, like Bertrand Russell, built on ideas previously developed by Georg Cantor (http://en.wikipedia.org/wiki/Georg_Cantor), usually thought of as the inventor of set theory. One of Cantor's key ideas was that the notion of a one-to-one correspondence between sets was the basis of our notion of cardinality of a set (the answer to "How many things are in X?"), a notion that he extended to infinite sets. Frege added the claim that the notion of a one-to-one mapping and operations related to such mappings (e.g. the combination of two disjoint sets A and B to form a new set C whose cardinality can be derived from the cardinalities of A and B) could all be defined using purely logical apparatus and proved to have the required properties using purely logical methods of proof. For that reason he claimed to have shown that truths of arithmetic were analytic, not synthetic as Kant had claimed.

Like Kant, Frege thought that geometry was not reducible to logic, and therefore that Kant (1781) had been right in claiming that geometrical truths were synthetic and necessary though wrong in making that claim about arithmetic. When David Hilbert demonstrated (Hilbert, 2005) that axioms of Euclidean geometry could be presented in a purely logical formalism and theorems derived using only logical inference Frege did not regard that as proving that geometrical theorems were analytic. I shall later raise similar objections to his claim about arithmetic. There is a very useful discussion of his disagreement with Hilbert in Blanchette (2014).

Frege's claim that arithmetic was reducible to logic received a serious blow when Russell discovered that Frege's logical system could be used to "prove" a contradiction, now known as "Russell's paradox".[9] The discovery of this paradox and others had deep effects on the Frege's and Russell's logicist project but that's not a topic for this paper.

Frege appears not to have reflected on the following parallel between arithmetic and geometry: both *can* be presented logically using logical concepts, definitions, axioms and rules of inference (though with special manoeuvres to avoid paradoxes), but before that was demonstrated both arithmetic and geometry had had a long history in human thought based on modes of thinking and reasoning that were different from (and arguably more biologically primitive than) the sophisticated

---

[8] I don't understand the recent proof of Fermat's last theorem well enough to know whether it can be said to use only pure logical inferences. I suspect not. It is not uncommon for proofs in a mathematical domain to require consideration of a larger domain, e.g. extending the rationals to the reals, extending the reals to complex numbers.

[9] See http://en.wikipedia.org/wiki/Russell%27s_paradox and http://plato.stanford.edu/entries/russell-paradox/ Irvine and Deutsch (2013)

and very abstract forms of logic that logicians, mathematicians and computer scientists are now familiar with. But those earlier modes of thought were clearly mathematical, not empirical.

He thought Immanuel Kant's claim that mathematical truths were both necessary (i.e. incapable of being proved false by facts about the world) and synthetic (i.e. not provable on the basis of logic and definitions alone) was true for geometry but not for arithmetic. But I suggest there is a deep parallel between the two cases, that for some reason Frege rejected, although he came close to considering it in his condemnation of other theories of the nature of arithmetic. For example, he writes in his Introduction "It may well be that in many cases the history of earlier discoveries is a useful study, as a preparation for further researches; but it should not set up to usurp their place", and earlier sarcastically comments (page xix):

> What, then, are we to say of those who, instead of advancing this work where it is not yet completed, despise it, and betake themselves to the nursery, or bury themselves in the remotest conceivable periods of human evolution, there to discover, like JOHN STUART MILL, some gingerbread or pebble arithmetic! It remains only to ascribe to the flavour of the bread some special meaning for the concept of number.

Frege is famous for his anti-psychologism, especially in connection with the nature of mathematics, but perhaps the reason he thinks geometry escapes that charge can be generalised to allow for an arithmetical domain that humans discovered and explored prior to the logicisation of arithmetic, just as they discovered and explored a geometric domain prior to the development of purely logical formalisations of geometry (though with undefined primitive predicates and relations).

A defense of Mill against Frege along those lines is proposed by Kessler (1980), but I want to offer a different defence for an analogy between a non-logical, yet mathematical, domain of arithmetic and a non-logical, yet mathematical, domain of geometry, both to be understood as important for the evolution of animal and human minds, but not because arithmetic and geometry are within the subject matter of psychology. Rather evolution produces some of its advances because it repeatedly develops new forms of psychology to cope with increasingly complex mathematical domains. How it copes, at least in the early stages, uses information structures and modes of reasoning that are different from, but just as capable of supporting mathematical discovery and proof as, logical modes. Frege seems to have accepted something like that as true of geometry, and it seems that his criticism of Hilbert's axiomatisation of geometry is that it ignores the pre-logical form of mathematics used in Euclid's discoveries. I am making the same criticism of Frege's (and Russell's) formalisation of arithmetic using only logic. This is connected with the great difficulty in giving current computers the ability to reason geometrically, discussed below.[10]

This is still sketchy and programmatic, and requires much additional work, so here I'll merely hint at the theory.

---

[10] Also in: `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html`

If there is a non-logical way of thinking about and proving truths of arithmetic, analogous to but different from proofs in geometry, then it would follow for both pre-logical Geometry and pre-logical Arithmetic that they include synthetic necessary truth that can be demonstrated without using empirical evidence, although both also have more abstract *analogues* in the space of logical systems definable using purely logical concepts with theorems provable using purely logical reasoning from axioms expressed using only logical concepts and additional variables.

My DPhil thesis and a derived publication (Sloman, 1962, 1965a), argued that Frege's notion of function was not general enough to accommodate concepts whose applications to potential instances produce truth values that depend on how the world is. I called those "rogators". A major theme of this paper is that tracing biological origins of mathematical concepts requires understanding how they start life as rogators. This is as true of arithmetical and logical concepts as geometrical concepts.

This project should not be confused with the research in developmental psychology and ethology that aims to test whether very young children or non-human animals have some mathematical concepts and understand some mathematical truths. Most of the empirical research I have encountered shows at best that some children and other animals can use concepts partly analogous to mathematical concepts (e.g. "numerosity", which is a perceptible pattern rather than "cardinality", which depends on use of one-one correlations) but lacking some of their core features, and also that the experimental subjects can acquire and use *generalisations* that appear to have mathematical content, but not that they are able to reason mathematically about what is mathematically necessary or impossible. There's a difference between learning *empirically*, for example, that removing some chairs from a room whose human occupants were all sitting on separate chairs earlier, can produce a result that prevents them all sitting at the same time, and understanding mathematically *why* that happens, which depends on a mathematical fact about one-one correlations. Compare Rips et al (2008). I think Piaget understood the problems but lacked the conceptual tools needed for explanatory theories. Piaget (1981, 1983).

6.1 Different conceptions of mathematics

Many non-mathematicians I have encountered have an extremely narrow conception of mathematics, perhaps related to what they were taught in an impoverished mathematical education, which may have included arithmetic, a little algebra, and some memorised geometrical theorems and trigonometric formulae, with a little statistics added if they are empirical researchers. Often they have memorised theorems that they cannot prove. Some of those who have studied mathematics at University level seem to think of mathematics in terms of formal languages, axioms, rules of inference and formal proofs. But that is a relatively recent rather narrow view of mathematics, not shared by all mathematicians. Formalised mathematics can be viewed as merely a subset of the subject matter of mathematics, that was unknown to those who made the earliest mathematical discoveries presented in Euclid's Elements. My (Sloman, 1962) agrees with (Mueller, 1969) on this. Compare what Benoit Mandelbrot wrote in the Foreword to Dauben (1995)

> "What is mathematics? Opinions range the spectrum from a wide, liberal Open Mathematics to a small Fortress Mathematics. For proponents of the former school, which I favour and which I am sure Abraham Robinson also favoured, mathematics is a big rambling building permanently under construction, with many doors and many windows revealing beautiful and varied landscapes. For proponents of the latter, the highest ambition is to wall off the windows and preserve only one door."[11]

Although Euclid's *Elements* included axioms and proofs, that is because the ancient mathematicians had begun to notice that mathematical discoveries are often related to other mathematical discoveries. This is part of what constitutes a mathematical domain. However they were not using (and could not have used) the notion developed at least two millennia later of *formal* proof involving primitive symbols, formation rules, derivation rules – axioms that need not be justified except as specifying the system being investigated – and explicit, formal, definitions of validity and soundness.

Chains of reasoning linking discoveries in a mathematical domain can go in different directions, as shown by the discovery of different, mathematically equivalent, axiomatisations of the same domain. E.g. Boolean propositional calculus can be presented with different initial symbols in terms of which others can be defined, and with different axioms, and different proofs of the same results. Neither the original discoveries in Euclidean geometry nor the original ideas about propositional calculus were concerned with logical derivability in a formal system. For example, propositional calculus can be, and often is, taught in terms of the domain of binary truth functions defined by truth-tables not axioms or rules of inference. Later, truth tables can be used as basis for validating axioms and inference rules. I taught beginners that way for many years, as others have done.

The variety of mathematical investigations and discoveries covers a huge and diverse collection of topics, going far beyond arithmetic and geometry (including, for example, shapes in motion, games of many types, grammars, language types, computational systems, logic, transfinite ordinals, probability, utility maximisation, finite and infinite dimensional vector spaces, and many others). A fairly broad partial overview of the scope of mathematics is provided by Wikipedia: `http://en.wikipedia.org/wiki/Mathematics` though it does not mention the role of (primitive) mathematical competences in organisms discussed below. There is also a brief introduction to the variety of forms of proof, including proofs using spatial reasoning in, `http://en.wikipedia.org/wiki/Mathematical_proof`.

There are systematic ways of discovering new mathematical domains simply by "thinking mathematically" about familiar non-mathematical topics, which turn out

[11] `http://www-history.mcs.st-and.ac.uk/Extras/Mandelbrot_Robinson.html`

to instantiate mathematical domains. I shall present several examples below. Even a child can discover a mathematical domain through play – e.g. discovering the domain of shapes that can be constructed by means of a single rubber band stretched over a flat surface and held in place by pins. (The shapes include an outline capital “T”. Do they include an outline capital “A”? Can you justify your answer?)

The wonderful little book (Sauvy and Sauvy, 1974) shows how young children can be introduced to mathematical thinking by exploring properties of everyday objects such as string, rubber bands, pins, buttons, etc., rather than axiomatic systems. There are also many examples in the work of Piaget and Holt (Piaget, 1981, 1983; Holt, 1967). I believe this is also how Kant thought about mathematics (Kant, 1781; Giaquinto, 2007).

7 The blind theorem-prover

There are different ways evolution makes use of mathematical relationships, some more sophisticated than others, including at one extreme simple organisms whose evolved design implicitly incorporates mathematical competences, and at another extreme human mathematicians who discover and explicitly investigate previously unnoticed mathematical domains, some of which turn out to have important practical applications, as Turing did in his work on computable numbers and later in his work on chemical morphogenesis (Turing, 1936, 1952) – two streams of thought awaiting unification, as suggested in (Sloman, 2013b).

So far those who have thought about the nature of mathematics have mostly thought of mathematical domains investigated by humans. But thinking of natural selection as (blindly) exploring and using mathematical domains can help us achieve a better understanding of both evolution and mathematics. From this point of view, the huge variety of more or less successful products of evolution is in part explained by the huge variety of mathematical domains.

There is an infinity of mathematical domains: After exploring a collection of mathematical domains it is always possible to find new domains by adding or modifying components of old ones, or combining domains, or abstracting from details. For example, the game of Go can be played on boards of different sizes. Each board size determines a mathematical domain. There is no limit in principle to the size of a Go board, though human brain limits may limit the feasibility of developing expertise beyond certain sizes. So even the game of GO generates infinitely many mathematical domains. (`http://en.wikipedia.org/wiki/Go_and_mathematics`)

That mathematics itself is unbounded (as implied in the quotation from Mandelbrot, above), should be no more surprising than that many mathematical structures are unbounded, including the set of integers, the set of ratios, the set of logical proofs and the set of geometrical structures.

This view of mathematics as infinitely extendable may be unfamiliar to most non-mathematicians. I’ll try to support it by presenting various examples of previously unnoticed mathematical domains both in order to demonstrate both how easy it is to discover previously unnoticed domains, and also how varied they are.

Many of the domains in which humans and other animals achieve mastery in their everyday life have mathematical structures: for example tying shoelaces, putting on or taking off clothing, domains that involve separating food into parts before eating, such as cracking a walnut shell or peeling a banana in order to access the edible part, or extracting meat from a carcass; or domains that involve constructing large objects (e.g. nests) by assembling many smaller objects (e.g. twigs, or lumps of mud) possibly requiring lengthy journeys with navigational sub-tasks. In each domain there are collections of possibilities that are generated by a subset of possibilities and modes of combination and variation, and it is possible to discover invariants and impossibilities as well as learning about ranges of possibilities.

One of the consequences of playing with toys or other physical objects is often meeting new domains, some continuous, some discrete, some continuous but with discrete partitions (e.g. because players alternate), some bounded (e.g. Tic-Tac-Toe) others potentially infinite, e.g. making sand-castles). Careful observation can sometimes reveal transitions in understanding: e.g. the child who at a certain stage does not understand why two trucks each with a hook and a ring at opposite ends can be joined by bringing a hook and a ring together, but not by bringing two rings together. In contrast, some birds can learn to make hooks to lift an inaccessible food container by its handle (Weir et al, 2002), and will do this in sufficiently varied ways to indicate a grasp of the mathematical structure of the problem.

Using familiar objects Vi Hart demonstrates a rich variety of mathematical domains in her wonderful videos presenting mathematical "doodles", e.g.: `https://www.khanacademy.org/math/recreational-math/vi-hart` and `http://vihart.com/`.

8 Types of domain mastery

Mastery, or partial mastery, of a domain can take many forms, including acquiring abilities to *perform actions* – possibly to achieve some goal, e.g. crawling, walking, running, catching, throwing things into a bucket, drawing a circle, tying laces, putting on a shirt, or getting food out of a banana or walnut – and abilities to *predict consequences* of events, or processes, including some of one's own actions. Performing and predicting depend on the ability to identify distinct sets of possibilities, e.g. possible actions to perform, and possible future consequences of some type of change. Expertise that comes from accumulating evidence from previous repeated actions or recording results of predictions, constitutes *empirical mastery* in a domain.

Empirical mastery in an individual organism or machine is produced by training in varied situations so that response types are associated with problem types in a memory system shaped by training (for which different mechanisms are available). A novel problem covered by one or more learnt types allows goals to be achieved by adjusting one or more parameters (angle, force, speed, direction), or by combining sub-competences (e.g. a familiar type of perceptual skill used to control a familiar type of movement), or by recognising states that have a winning move, or states that are followed by states that have a winning move, for example. Such learning may be based either on randomly attempted combinations, or by application of meta-competences that recognise relevant aspects of a novel problem type and direct the search for a new solution.

Likewise, natural selection is able to use various adjustments to designs of organisms, including parameter modification or combining previous solution types, for example. But that uses no explicit recognition of a problem type and no explicit meta-knowledge about previously acquired competences encoded in a genome. There is no directed search for a solution to a new problem, though a new solution can be adopted if an unintended design change produces instances with new powers. This need not be at the expense of other species or other individuals, if the resources used by the new individuals are not used by previous individuals, e.g. when organisms move into new previously uninhabited terrain. Natural selection does not have to be competitive all the time!

There are also changes that allow conspecifics to collaborate (e.g. by "division of labour") and changes that allow cross-species symbiotic co-operation. Such evolutionary changes may involve new mathematical domains composed of different sub-domains corresponding to different roles in collaborative relationships. A similar point can be made about the mathematics of competitive relationships. (Simple cases where all the relevant changes involve numerical utilities have long been studied mathematically in the framework of theory of games and decisions.)

For novel *types* of situation, where old competences cannot easily be made useful by modifying parameters, more radical changes are needed, for example the enormous changes between a light sensitive area of skin to a mechanism for acquiring information about enduring structures in the environment using vision. One of the goals of the Meta-Morphogenesis project (Sloman, 2013b) is to investigate types of transition in information-processing that might have occurred at various stages.

More generally, we should aim for a theory of *types* of transition. This will involve new mathematical domains linking old and new designs. One likely result is that there are some transitions that can happen directly from one generation to another, whereas others require intermediate stages. For example, forms of life in which competition for mates occurs cannot emerge in a population that does not use sexual reproduction. So only after evolution of sexual reproduction can there be evolution of cognitive abilities required for perceiving potential mates and their properties, and abilities for recognising and competing with potential rivals for the same mate. Many more mathematically complex routes through evolutionary transitions await future research.

Many different types of competence are required for success in naturally occurring environments. Over-emphasis on the role of embodiment in intelligence has led to excessive focus on robots with practical empirical mastery using sensory-motor loops, as a result of which other kinds of learning and mastery are often ignored (Sloman, 2009b). Both *online* intelligence and *offline* intelligence, e.g. the ability to develop an improved ontology, or to notice a new type of constraint in a space of possibilities, or to develop a new form of representation, are all part of what needs to be investigated and explained.

A type of offline mastery that seems to exist only in a small subset of species is what could be called *mathematical mastery*, using information-processing capabilities that probably evolved relatively recently. This includes the ability to *invent theories* about structures and processes in a domain and to derive consequences about what is or is not *possible* in the domain. This can make it possible to choose a new solution in a new situation by reasoning about a range of possible consequences, without first having to try alternatives to find out what will and what will not work.

9 Mathematical mastery

That sort of reasoning competence is particularly important for satisfying goals that are not specified in full metrical detail: e.g. the goal of getting to the other side of a wall, where the precise trajectory, the precise speed of motion, the precise sequence of moves, etc. is immaterial to the goal. For example, an intermediate goal could be finding or building a rigid object whose top surface is within reach by stepping or clambering, and which is close enough to the top of the wall to allow someone resting on that surface to lift a leg over the wall. Satisfying the goal of getting to the other side of a river may include achieving the intermediate goal of finding a set of suitably arranged stepping stones. In both cases the precise size, shape, location, etc. of the objects used is immaterial, provided certain constraints are satisfied. Representing this range of possibilities using a probability distribution is using a sledgehammer to crack a nut. For more examples, see (Sloman, 1996).

An animal that lacks the mathematical mastery required to design plans may nevertheless be trained by someone else to build an intermediate platform to get over the wall, or to search for and use stepping stones to cross the river. Such training can produce competence without mathematical understanding (like much bad education).

It is also possible to learn as a result of some random sequence of movements that happens, fortunately, to produce success. But in many cases the search space will be too large, and better methods are needed.

Moreover, even when learning based on training by another individual or based on trial and error, produces successful behaviour, the learner may not acquire the ability to use suitable variants of the solutions, because there is no deep understanding of which features of the working solutions are required and which can be varied.

When that additional deep (proto-mathematical) understanding is present it can lead a child who has learned to use a screwdriver as a lever to open a flanged lid, to cope when no screwdriver is available, by *working out* that the required functionality also exists in a different rigid object, another flanged lid, or rigid disc. This sort of discovery, which can occur in pre-verbal children, would be an example of a "toddler theorem"[12].

When present, such theoretical/mathematical mastery does not provide *infallible* powers, partly because things that look similar may behave differently, e.g. two levers with similar appearance, one of which is made of weaker material than the task requires. In sufficiently sophisticated organisms such failures can drive ontology extension, e.g. leading to richer concepts of kinds of stuff, including kinds of stuff with similar appearances but different physical properties (e.g. different strengths). This requires an ontology that extends beyond perceptual categories.

An even deeper understanding can come from insight into structural relationships between features of a complex object: for example understanding why a loosely linked tetrahedron formed from 6 rods will be rigid whereas a cube made in the same way from 12 rods will not. For individuals who understand the differences between triangles and other polygons that would be a "theorem", for others merely an empirical discovery.

These abilities to perceive, understand and use various abstract kinds of affordance in the environment may have been essential precursors to the evolutionary changes that allowed the discoveries later assembled in Euclid's *Elements*. The capabilities required are not yet available in AI or robot systems, for reasons that will be discussed below.

10 What makes a domain mathematical?

In a mathematical domain, relationships can initially be discovered empirically, e.g. by observation or repeated experiment, then later *proved* to hold without exception. The second step is impossible in non-mathematical domains. For example, in the domain of state changes of electric light wall switches (e.g. going up or down) and the domain of state changes in electric lights (going on or off) someone may live in a house where there is a 100% correlation between switches going down and lights going on. But no amount of experimental repetition amounts to a proof. Moreover, the learner cannot examine more closely the process of moving a switch downwards or the process of a lamp beginning to glow, so as to discover deep connections between them.

Many adults know that a power supply or a lamp can fail, fuses can blow, or an electrician can swap the switch connections. There are also lights, e.g. on stairways and corridors, that are controlled by two switches, and how one switch affects the light depends on the state of the other switch. Furthermore the switch wiring conventions in different countries are not the same. So we understand why the mapping between switch states and light states is highly variable.

---

[12] `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html`

In some cases close investigation of variable relationships discovered empirically can reveal special cases of the relationship that have some invariable properties. For example, the light-switching domain can be transformed into a mathematical domain by extending the portion of the world the child can see so as to include hidden cables and connections that change state when a switch is operated, along with power supplies and mechanisms for converting electrical energy into light energy. With those extra details, a user still cannot reason *mathematically* about the domain. That can be done after *idealising* the domain in various ways, e.g. treating the power supplies, electrical conductors, switches, etc. as *infallible* and then reasoning mathematically about what *must* happen and what *cannot* happen in various situations. (Compare the levels of domain mastery in Section 8.)

A simpler example of this transformation of an empirical regularity to something provable might start with the empirical discovery that if a large wheel is used to make a small co-planar wheel in contact with its rim turn, then if the ratio of circumferences is 2 to 1, then, if both wheels have fixed axles, the ratio of rotational speeds is 1 to 2. But further experiments may reveal cases where the ratio is different, and the small wheel turns less than expected. Then further analytical thought may reveal the possibility of distinguishing two cases, one where there is slippage, so that portions of the wheels that are in contact may move at slightly different speeds, and one where there is no slippage – in which case the ratio of circumferences implies that one rotation of the big wheel will be accompanied by two rotations of the small wheel, as can be seen by picturing the rotation, and thinking about how much of each circumference has passed the contact point when the big wheel has completed one rotation.

At that stage a mathematical thinker could generalise the relation to other cases than 2 to 1 ratios, and after learning some more mathematics may realise that the ratio of diameters has the same implication as the ratio of circumferences. (I am skipping complications regarding the ability to compare amount of motion in a straight line and amount of motion in a curved path, and other important developmental complications.)

Transforming the empirically observed rotation ratio into a proved mathematical theorem, and transforming the empirically observed light switching correlation into a proved relationship, both involve re-representing the structures and processes involved, in some cases adding extra components (the originally hidden circuits, connections and light producing mechanisms) and in the other case excluding a type of process, namely slippage or bending, which could also be done by postulating a mechanism producing perfect friction and postulating perfect rigidity.

I leave it to the reader to try inventing a mathematical domain in which the wheel mechanism is altered so that the turn ratio is larger than 2 to 1. It might be done by adding new microstructure to the domain.

In presenting these examples I am using a great deal of common knowledge, probably shared by the vast majority of humans likely to read this! A more detailed theory would be required to explain how that knowledge arises, how it is represented, how its relevance is recognized and exactly how it is used. Such processes might go on in future robots, but we shall first need a better understanding of the innate mechanisms required to support the staged processes of concept formation, theory extension, growth in explanatory reasoning, and development of abilities to create and evaluate explanations and proofs.

The crucial feature of a proof of a mathematical property in a mathematical domain is that all potentially relevant influences have been noted. This provides a *guarantee* that nothing can disrupt what has been proved (if you have made no mistakes), whereas there's no such proof for the initial connection between light switches and lighting states.

In more complex cases, how the proofs used by mathematicians work is not obvious, and I'll return to that later. The most important fact for now is that an individual with the competences that allow certain proofs to be understood may not need to be taught the proof by a teacher or community of others. The learner may not even notice that the proof has been understood and an exceptionless generalisation discovered. It is possible to be an unselfconscious mathematician (as most people are about their linguistic competences).

It is also not true in general that understanding a proof is a matter of checking that various proof rules have been followed without exception, since, during normal development of mathematics, rules are usually discovered only after instances of the rules have been understood Sloman (1968/9). At a much later stage, after the development of metamathematics, explorations of alternative sets of axioms and rules becomes part of mathematics (in the late 19th Century?).

11 Mathematical development

Many animals, including human infants and toddlers, acquire practical non-mathematical mastery of many domains, as do highly trained robots. Acquiring mathematical mastery is less common, and in humans comes later than the empirical practical mastery, sometimes very much later. Meta-mathematical competence is very rare, even in adults. I don't know whether a different educational system could change that.

For example, a child who can put on a shirt expertly may be unable to think about the number of significantly different ways of doing that, or about why starting from a different configuration would lead to the shirt being on back to front. Being able to start from a configuration and perform actions to achieve a desired result is possible without the ability to notice that there are many slightly different initial configurations and trajectories with common constraints that produce functionally equivalent results, whereas other initial configurations, or other trajectories that diverge from the first set that produce different results, like a shirt being on back to front, or an arm going through a neck-hole rather than an arm-home. That sort of thinking about how to divide sets of possibilities into subsets with non-overlapping continuations where the different subsets satisfy different constraints requires cognitive abilities that are not required for performing actions using one set of possibilities successfully.

As far as I know no current robots (in 2013) acquire theoretical/mathematical mastery, in the sense discussed here, partly because researchers are mainly focused on providing empirical mastery through training processes – providing robots with "online" intelligence. This tendency to ignore "offline" intelligence is one of the bad consequences of the over-emphasis on embodied cognition in the last two decades (Sloman, 2009b).

That work ignores the important roles of "offline" intelligence, or "deliberative" competence: namely being able to think and reason about sets of possibilities without acting, including inferring consequences of *types* of situations that do not exist but might, as opposed to reasoning about consequences of a very specific situation, which game engines can already do.

Research in the first two decades of AI attempted to give machines such abilities to reason hypothetically, but most of the work had to be done by human programmers designing suitable formalisms, rules, inference mechanisms, etc., as in the STRIPS problem-solver developed in the 1960s `http://en.wikipedia.org/wiki/STRIPS`. We have yet to produce machines that can produce the theoretical discoveries (and inventions) that the human AI/Robotic researchers achieved and used in programming their machines.

What's missing is the ability to understand *invariants* in a structure, e.g. invariant properties of triangles, that are independent of particular shape, size, location, etc. An example is the difference between a person who can predict that the intersection between a *particular* plane and a *particular sphere* will have a circular boundary, and someone who understands that the prediction holds for *any* plane intersecting *any* sphere anywhere. (A tangent plane can be regarded as a limiting case, where a point is a limiting case of a circle.) What sorts of shapes can result from a plane intersecting a cube?

These abilities to chunk future possibilities into clusters with different constraints and further possibilities can be used for construction of multi-step plans that involve types of situation an individual has never previously experienced. (For further discussion of different levels of deliberative competence, see Sloman (2006).)

Although such hypothetical, counter-factual thinking and reasoning is not normally classified as mathematical it is possible only because there are mathematical domains that thinkers have understood, including domains of possible physical configurations, domains of possible actions in those configurations, domains of possible state-action-state combinations forming a space of possible futures, and many more. Exactly what changes in brain design evolution had to produce in order to make this possible is not clear.

There is also *meta-mathematical mastery*, which involves being able to think about one's mathematical thinking or reasoning and identify steps as valid or invalid, redundant or not, specialised or widely applicable, elegant or clumsy, etc.

Sometimes meta-mathematical mastery of two domains leads mathematicians to discover a common structure leading to a new more abstract domain. An example is noticing the mathematical *group* structure common to addition, multiplication, and various transformations of geometrical operations (e.g. rotations) on 3-D objects.

In other cases, domain combination can *enrich* previously understood domains, e.g. adding forces, masses and accelerations to a domain of topological and geometric constraints on motion, or adding folding operations to Euclidean geometry, as in origami geometry, which starts from a subset of Euclidean geometry (without circles) and add folding of planes subject to constraints (Geretschlager, 1995). An interesting question is whether humans who had never encountered materials like paper or cardboard could have thought of origami geometry. I suspect the answer is yes in principle, though it may be unlikely. Not all humans have the same potential to make mathematical discoveries.

Another kind of mastery of a domain, closely related to meta-mathematical mastery is *Pedagogical mastery* which includes being able to detect whether *other* individuals have empirical or theoretical or meta-mathematical mastery of a domain, and helping them improve their mastery. This requires meta-semantic competences: abilities not only to use information (semantic competences) but also to think and reason about use of information, including use by others. Polya's pedagogical mastery is demonstrated in his (1945).

An aspect of pedagogical mastery identified by Lev Vygotsky is understanding a student's "Zone of proximal development"[13] and using it to set challenges that are neither too easy nor too difficult for the individual[14] – which some learners can do to themselves (self-scaffolding). Explaining what domains are and describing various kinds of domain mastery that individuals can acquire, requires unusual pedagogical mastery – which I am struggling to achieve!

These different forms of domain mastery (empirical, mathematical, pedagogical, self scaffolding, mastery, etc.) can be found in relation to many different domains, including different subsets of mathematical knowledge, e.g. arithmetic, algebra, calculus, topology, geometry, logic, meta-mathematics, theory of computation, and many more. It is unlikely that there will turn out to be one specific information processing mechanism that performs all these tasks, despite claims of those who, like Turing in his (1950), rashly propose totally general learning mechanisms, for example, mechanisms based on neural learning, or information compression, or some form of logical reasoning. Part of the evidence for the diversity of mechanisms (and the information structures on which they operate) is the length of time between the earliest (prehistoric) discoveries leading to ancient recorded mathematics, and the most recent forms of mathematical research, including computer-aided mathematical research. Different domains often require different mechanisms for their mastery.

These ideas about common types of transition with different contents and different mechanisms are closely related the ideas about "representational redescription" in Karmiloff-Smith (1992).[15].

To illustrate these ideas about mathematical domains, the next few sections introduce domains that arise out of familiar structures and processes and have

[13] `http://en.wikipedia.org/wiki/Zone_of_proximal_development`

[14] `http://en.wikipedia.org/wiki/Instructional_scaffolding`

[15] I have begun to summarise and discuss aspects of that work in
`http://tinyurl.com/CogMisc/beyond-modularity.html`

been well studied by mathematicians, though I'll present only simple cases, some continuous and some discrete. I'll then move on to geometry, my main target, and attempt to characterise its role in human and animal competences (i.e. its biological importance) and then present some simple examples of kinds of reasoning that could have led to the development of Euclid's elements in our ancestors thousands of years ago, but seem to be very difficult to give computers, or robots. These tell us far more about human consciousness and its underpinnings than theorems about incomputability or undecidability emphasised by Penrose 1989 and others. Later I'll raise the question what all this tells us about the nature of mathematics, and ask why it has proved so difficult to mechanise human geometrical reasoning abilities.

12 A domain of curves on a doughnut

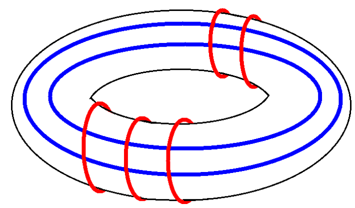

Fig. 1 *Five red and two blue closed non-self-crossing paths on a torus. Are other types possible?*

A doughnut-shaped object (a torus, or ring) is the basis for a domain of *closed non-self-crossing paths* on its surface. In how many different ways can you draw a closed curve on such a surface that nowhere crosses itself (as "**8**" does)? What should count as different curves? Examples are depicted in Figure 1.

Ask whether a loop of stretchable string lying on the surface can be transformed from one closed curve to another by sliding the string around on the surface without cutting the string and without the string ever losing contact with the surface, or crossing itself. If such a transformation from one curve to another is possible the curves are equivalent, otherwise not. A set of curves that can be transformed into one another without any cutting and joining (or passing through solid material), form an "equivalence class". It should be obvious from the figure that the blue curves are in the same equivalence class, since each blue curve can be smoothly transformed into any other in the surface of the solid, without any cutting or joining. Likewise the red curves (if each is a separate ring, not joined to the others out of sight) are in the same equivalence class, though the transformations between them have a different character. Can you tell whether the red curves and the blue curves are in the same equivalence class (ignoring the colour difference)? Answering this requires not only failing to find a smooth transformation between a red and a blue curve, but detecting that it is *impossible* for such a transformation to exist, just as you can detect that it is impossible for line in a plane, straight or curved, to connect a point inside a circle in that plane to a point outside the circle without somewhere touching the circle. Are you sure that is impossible? Why?

You can try a range of transformations that fail and eventually be sure there's nowhere else to go to find another transformation. Some people will find that less obvious than others. This is a subset of topology[16], a branch of mathematics not normally taught to beginners, though there's no reason why topology should not be taught to children, informally, for example using ideas in Sauvy and Sauvy (1974). Many board games introduce children to different sorts of routes in a space. (E.g. "Go directly to jail" in *Monopoly*, and snakes vs ladders in *Snakes and ladders*.)

A different question: Are there other possible closed curves on the surface of a torus that are not in the same equivalence class as either the red or the blue curves in the figure? How many equivalence classes of non-crossing closed curves are there on a plane surface, or on a sphere? How many classes are there on a torus? (After you have reached a your own conclusion, look at Figure 2, below.)

A set of equivalent smooth curves on a smooth surface, like the set of curves of which a small sample is shown in red in Figure 1, is a *continuous* set: between any two curves there are infinitely many others and each can be smoothly deformed into any of the others. Likewise the curves equivalent to the blue curves form a continuous set. But no member of the red set can be continuously deformed into a member of the blue set, or *vice versa*, so in the case of the torus there are discontinuities between classes of curves. Why can't a red curve be smoothly transformed into a blue curve or vice versa, without damaging the torus? You can try to find a way of doing it and be convinced that you have tried all the possibilities. It is always possible in principle that you've missed something: humans are not infallible at mathematical reasoning – (Lakatos, 1976) famously showed that even outstanding mathematicians can make mistakes. However mistakes can be detected and corrected. But in simple cases it is possible to be sure that you have considered all the possible moves. There's nowhere else for the transition to go. Discovering that requires use of "offline intelligence", defined in Section 29.

How do you do this reasoning about what is and is not possible? Do you use logic? Most people would be incapable of saying what they know about a torus using only logic, without any diagrams, 3-D examples, or gestures indicating shapes.

Can you give a logical proof that there's no smooth transition between a red curve and a blue curve in Figure 1? Is using a logical formalism a requirement for reasoning about transformations of curves on a torus? If not, what alternatives are possible?

Humans thinking about this must use some internal structures and processes encoding information about the domain and the constraints on processes in the domain, aided in complex cases by external diagrams or other objects. Some problems about curves on a torus can be made easier to think about by noticing that two appropriate cuts on the surface of the torus allow the torus to be flattened (with minor distortion) to form a rectangle. What brain mechanisms allow that discovery? How could a machine make that discovery? What did evolution have to do in order to make produce these human capabilities? I don't think anyone has the faintest idea, though the apparent abilities of other intelligent animals to reason about what is or

[16] http://en.wikipedia.org/wiki/Topology

is not possible in a situation suggest that some of the relevant competences evolved in other species.[17] I'll offer some conjectures later.

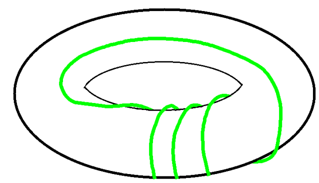

Fig. 2 *Another non-crossing closed path on a torus. There is no upper bound to the number of different such paths. Why not? How does this differ from the paths in Figure 1?*

13 Mathematisation of a domain

What starts as an empirical discovery can sometimes be transformed into a mathematical discovery after thinking about what has been learnt and understanding why that must *always* be the case. Why are no exceptions possible to the discovery that if you count 1, 2, 3, ... in synchrony with turning coin over, then even numbers coincide with a repeat of the initial state when coin flipping is combined with counting? What would have to happen for an odd number to coincide with a repeat of the initial state? Can you be sure that's impossible?

In this case, even though the domain is infinite, composed of an unending stream of states with two components, a coin state and a number counted, it is possible to notice a pattern that the process cannot break out of. To see that, you have to think about what happens during a state transition – the coin is reversed, and the number goes from odd to even or even to odd, another reversal. Why can't there be two consecutive odd numbers, or two consecutive even numbers? To answer that you have to specify what the odd/even distinction is. How is the problem changed if instead of counting you simply continue alternating between two different words, e.g.: cat, dog, cat, dog,... ?

Another way to combine coin flipping with counting, or with a potentially infinite odd/even alternation, is to start with two coins, and generate a sequence of flips where either coin can be flipped at any time, but never both coins at the same time. By counting at the same time as flipping, a new domain is created. By playing with examples and thinking about the patterns found, you may be able to find new unbreakable patterns that are generated. Can you explain why they are unbreakable?

You may discover that if the initial state has both coins with heads up then after an odd numbered flip the two coins will be in different states and after an even numbered flip they will be in the same state, no matter which coins have been flipped, and in what order as long as it is only coin at each counting step. What difference will it make if instead of starting in the same state the coins start in

[17] For further discussion of requirements for "internal languages" see Sloman (1971, 1978b, 1979, 2008, 2011); Sloman and Chappell (2007)

opposite states: one heads up and one tails up? This is an example of discovering a structural relationship between two domains when they are merged to form a new domain.

With three coins there are more complex mappings between numbers of flips and properties of the coin state sequences – left as an exercise in mathematical exploration, for the reader.

Discovering such regularities empirically, by recording frequencies of associations, requires quite different information-processing mechanisms from those needed for understanding *why* the regularity *must* exist. Much current research into learning in children and robots fails to distinguish the two sorts of learning and two related concepts of causation – Kantian and Humean causation[18]. In part the failure comes from a common belief that the kind of *necessity* discovered in mathematical reasoning is just a *very high probability*. This is not the case when what is discovered is a structural limitation on what is possible, not a probability.

Being able to represent something as impossible requires conceptual and representational competences whose precise nature is not easy to specify. Making such discoveries requires an information processing architecture whose capabilities are not restricted to discovering correlations and their frequencies. Observed correlations tell you nothing about what *must* or what *cannot* happen. Mathematical insights of the kinds described here go beyond that. However an observed instance can refute a suspected mathematical generalisation, as the history of mathematics shows.

The ability to discover mathematical necessities probably evolved long after the ability to detect and use correlations. The ability to use discovered correlations is often implemented as a mechanism that adjusts preferences between selections – with varying degrees of sophistication in the mechanisms by which past evidence and preferred outcomes influence the adjustments, some discussed in Russell and Wefald (1991) – though like many AI researchers they assume resource allocation problems arise out of use of a single CPU, ignoring the biological solution of dedicating different processors with different capabilities to different tasks.

A robot that learns *from experience* that three coins added to two coins always gives five coins, or that counting from left to right gives the same result as counting from right to left, is not learning mathematics, even if it ends up always giving the right answers to test questions. The non-mathematical robot does not treat this generalisation any differently from observation-based generalisations about which way to move a light switch. It uses only evidence, not proof. As far as I can tell, this is true also of the very impressive robots developed in Ben Kuipers' "Bootstrap learning" project[19] The robots learn empirically, using correlational evidence, but are not able to discover mathematical proofs, or use pre-human proto-mathematical modes of reasoning about affordances. Likewise the impressive toddler-like types

---

[18] A distinction presented with Jackie Chappell, in these talks: `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac`

[19] `http://web.eecs.umich.edu/~kuipers/research/ssh/bootstrap-learning.html`

of robot learning based on reward-free motivation reported in Ugur (2010) (though I need to study this work more closely).

14 Domains in the physical environment

Many domains are related to the properties of different kinds of matter: e.g. water, sand, syrup, mud, porridge, kitchen salt, caster sugar, normal sugar, skin, hair, various kinds of cloth, various kinds of paper (tissue paper, writing paper, sheets of cooking foil, sheets of plastic film). Some kinds of matter are fluid or flexible, and some of them can interpenetrate e.g. sand and water forming mud. Kinds of matter that are rigid and impenetrable are the basis of many processes of construction of shelters, clothing, tools, and machines. They are also the basis of certain kinds of mathematical reasoning.

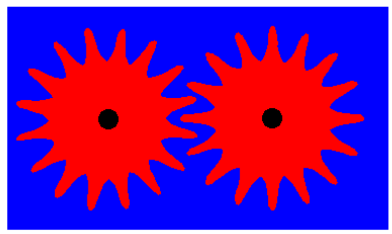

Fig. 3 If the gears are made of rigid, impenetrable, material and the two axles are fixed, what can you conclude about how the gears rotate? Rigidity and impenetrability are mathematical constraints, though their existence has physical causes.

Children playing with examples of such domains may discover properties of various structures and processes empirically. They sometimes get things wrong, like the child of a colleague who was surprised when she failed to grasp a vertical column of water flowing from a bath tap. But in some cases they understand that the domain specification has implications that can be derived simply by reasoning, without using evidence.

For example, if two co-planar gear wheels are made of rigid, impenetrable material, and their teeth are meshed, as in Figure 3, do you have to find out empirically how rotating one wheel affects the other? If one wheel turns clockwise the other *must* turn counter-clockwise since otherwise at least one of the wheels is not made of rigid and impenetrable material. Understanding this requires thinking about what happens when one of the meshed teeth moves upwards or downwards. What is implied by rigidity and impenetrability of the material? What does that allow you to infer about motion of a tooth in contact with the other wheel?

A child who grasps the generalisation needs to have an understanding of impenetrability and rigidity (impossibility of shape change) that is independent of the size and shape of a rigid object. The impenetrability implies that if a tooth moves upwards and comes into contact with the lower surface of part of the other wheel, and continues moving, then the part of the other wheel being touched must also move upwards. If that wheel is rigid, and has a fixed axle, then as one part moves up the whole wheel must rotate about the axle. Rigidity implies that no portion of the surface of the object can change shape, such as acquiring a new crack, fold, dent or bump. Similar considerations apply to motion of levers, and to pulleys with unstretchable strings going round them. In all these cases, as with electric light switches, and the routes by which light travels, we need to separate out the empirical discovery of how things actually work from the mathematical study of implications of particular ways of working.

For the latter, we do not need to start from logical specifications of (e.g.) rigidity, contact, or rotation, Some learners who are incapable of logically deriving theorems from axioms expressed using boolean operators and quantifiers can, after a certain process of development, grasp the mathematical impossibility of a counter example to the gear wheel generalisation, e.g. the impossibility of a pair of rigid impenetrable meshed gearwheels turning in the same direction about fixed axles. (Contrast empirical Bayesian learning.)

Biological evolution somehow produced the ability to detect and reason about necessary connections between structures and processes, but as yet we do not know what the information-processing architecture is that makes this possible – though it seems to depend on a meta-cognitive ability to reflect on what has been done. This is connected with, but different from, Gibson's ideas about affordance perception.

15 Domains of view-changes and epistemic affordances

Many animal actions alter the information available to the animal. Different sorts of information need different actions, provided by tasting, smelling, listening, prodding, pulling, twisting, threatening, and many more. Some of the actions assume that the entity being investigated is itself an information-processor, e.g. a conspecific with useful information to impart, or a competitor or predator from whom information should be kept.

A subset of animals use vision as one of the major sources of information, if not the dominant sort. For such animals there is a great deal to learn *from* vision, but also a great deal to learn about *vision*, e.g. about how to change what visual information is available, or what new visual information would be provided by various actions, that alter what is visible, and how it is visible, e.g. with more or less detail.

In particular, many animals have to learn that moving around makes different things become visible or invisible. As you walk round a house you can see different parts of it, so different information becomes available in a sequence that depends on the structure of the house (as Kant noticed). We can say that motion of a perceiver alters the "epistemic affordances" available to the perceiver. A special case of this is indicated in Figure 4. What can be seen through a door from a particular viewpoint changes as the viewpoint changes, because the viewer moves left or right or forwards or backwards, what can be seen in the room changes.

These relationships may be used by many animals. For some of them, information about the role of vision as a source of information may be built into their

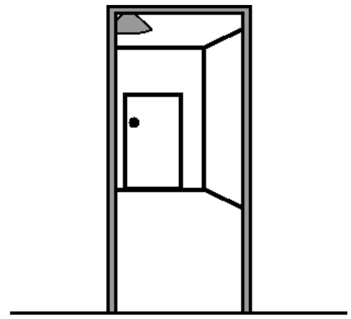

Fig. 4 Example "toddler theorems" about visibility: If you move towards or away from the open door, or sideways to the left or to the right how does what is visible through the door change?

information processing mechanisms as a result of their evolutionary history. They simply use vision to obtain practical information and don't consider any alternatives, or ask how it works. For some species, how vision is used, and which actions are deployed to alter visual information could be results of individual empirical learning. Correlations between how they move and what they can see are noticed and used to influence future movements when information is required. They may learn (implicitly) *that* rotating an object gives new information about its surface without knowing *why* it does – namely because it produces new unobstructed linear pathways between parts of the surface and the individual's eyes.

For animals or robots with additional cognitive competences it may be possible to *work out* rather than learn empirically what follows from the assumption that visual information travels in straight lines. An individual with an understanding of straight lines could work out that as a viewpoint approaches a fixed doorway, the lines defining the limits of what is visible from that viewpoint diverge increasingly, or equivalently, the walls block a reducing subset of lines of sight. What can you infer about moving backward, or moving left or right? How? This seems to require some of the key ideas about straightness in Euclidean geometry. It may depend on an ability to imagine a downward facing view (plan-view) of the whole scene, showing how horizontal view angles change as the viewer moves. The opportunities to learn and use such relationships vary according to where individuals live. A built environment with vertical walls and door frames provides epistemic opportunities missing for bush-dwellers or desert-dwellers – they have only much more complex configurations where boundaries of visible portions of scenes are rarely vertical lines. (I have omitted counter examples to the claim about gaining more information on moving towards the doorway. These are left for the reader to explore.)

Other actions that can alter epistemic affordances include rotating an object, to make new parts of it visible, removing intervening objects that block some or all of the desired view, or removing something covering the object, such as a lid, or a towel resting on it. In this case we need to separate the fact learnt empirically that visual information (light) travels in straight lines (for which there are well known exceptions produced by variations in transmission media, and gravitational fields) from the mathematically derivable implications, which I am suggesting even young children and some non-human animals learn and use. It may be possible to show that in some cases the learning is empirical initially, then replaced with mathematical reasoning. (A non-empirical truth can be derived from empirical facts Sloman (1965b), even though there are quicker, more general, deeper, non-empirical derivations of the same generalisation, requiring a more complex cognitive architecture.)

In the last few centuries, use of microscopes, telescopes, and cameras, including video cameras, have contributed new information-altering options. The use of these additional devices alters the set of possible routes and patterns of flow of visual information. Humans can learn to think about this and reason about what will happen in new situations, e.g. when designing a surveillance mechanism. An interesting question is whether playing with virtual spaces can damage formation of some mathematical understanding of spatial relationships in young children, e.g. if the theories about acquisition of geometric concepts in Nicod (1930) are correct.

16 What's the source of arithmetical truths?

It is sometimes suggested that social conventions or stipulations are the source of mathematical truths: as if humans had room to agree that 3 plus 5 should equal 8. In order to discuss this we have to get clear what is meant by 3, 5, 8 and other numbers are and what "plus" means. These concepts have been used for centuries in many practical applications, in connection with payments, numbers of soldiers or other participants in organised activities, measures (e.g. of length, area or weight) and with relations between discrete sets: e.g. the set of chairs at a table and the set of people sitting on them. Ensuring that there are enough chairs for each person to have one, and no chairs are unused can be done by using results of counting, where counting allocates a numerical label to a set of objects.

Some readers may need to be reminded that integer arithmetic as originally discovered, and still widely used, was not based on Peano's axioms: the axioms merely expressed an important relatively recent (late 19th Century) discovery *about* the positive integers, namely that Peano's axioms along with recursive definitions of concepts of addition and multiplication, provide a logical specification for which positive integer arithmetic, as understood centuries earlier, is a model.[20] Negative integers, ratios, subtraction and division require additional constructions.

What are the original number concepts? The key notion required to answer this is the idea of a one to one correspondence between two sets of entities, as illustrated in Figure 5 It is a curious fact that humans had been using one to one correspondences for centuries, for example whenever they counted objects, but also when assigning people to places at a table, when performing dances composed of male-female couples, and in many practical tasks such as making gloves designed for normal human hands. (The reader is invited to think of additional examples.)

[20] The axioms are presented briefly in `http://mathworld.wolfram.com/PeanosAxioms.html`, with a more extended presentation and discussion in `http://en.wikipedia.org/wiki/Peano_axioms`

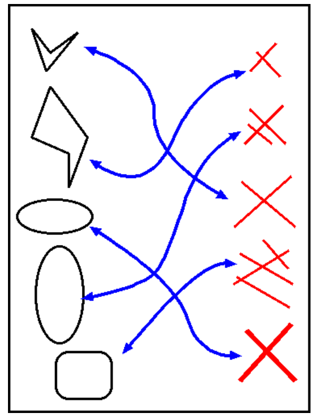

Fig. 5 A one to one mapping between the set of objects on the left made of closed curves, some smooth some not, and the set of objects on the right made of groups of overlapping lines. The blue double-headed arrows indicate the mapping, but there are several more possible mappings between the two sets. If the two sets remain unchanged, it is clear that the one-one mapping can be changed, e.g. by moving the arrow-heads on one side of the diagram. How many distinct one-one mappings are possible between the two sets?

Frege (1950) Once the structures and operations have been specified there is no room for arbitrary decisions: the structures of the mappings and the concatenation relationship, make only one outcome of the extended counting operation possible. A philosopher might argue: “But what if the result comes out different, with no mistakes being made?”

A possible reply is that an examination of the process of combining two such one to one mappings so as to produce a new one to one mapping shows that only one outcome is possible. We can see that that must be so when the sets are small. It is not possible to have the same clarity of vision for addition of very large sets. But we can work out that the case of large sets is approached step by step from the case of small sets being added and it is possible to see that if one case of adding M and N entities produces a total of K entities and another case produces a different total of L, then we can map the two together and must find an error in at least one of them.

This needs more detailed argument and demonstrations of how errors can be detected though for most readers that will be obvious.

Some mathematicians and philosophers have doubted that forms of reasoning that work for small numbers, or simple structures can be extended to arbitrary situations without becoming empirical.

Notice that I am not claiming that the principle of mathematical induction is used as a premiss in a logical argument. The principle is itself a summary of a mode of reasoning that in each case is justified by the contents of the case. The cases are not justified by the principle. Likewise inferences of type *modus ponens* are not justified by a logical rule: the rule summarises what has been discovered about that form of inference without using the rule.[21] However it is possible for a mathematical reasoner to abstract the general principle from particular cases and understand why the principle works in all cases. There are also important differences between cases, for example differences between induction applied to numbers generated by a successor function, and “structural” induction, applied to infinite

[21] Assuming that a rule is required leads to absurdity as shown in Carroll (1895). See also Sloman (1968/9).

classes of structures generated by more complex construction processes – widely used in computer science. Compare the comment on the disjunctive syllogism in Section 17.

"3+5=8" expresses an unavoidable property of the domain of mappings from sets of entities to numerals expanded with operations for merging sets and concatenating mappings. For many people, that example may be too abstract to be obvious, which is why I started with much simpler examples of domains which don't come with a large collection of "baggage"(good and bad) produced by past philosophers discussing them.

A major missing part of the argument is identification of the particular biological competences that led up to, and provide mechanisms for, these human mathematical competences. Insofar as those competences are products of biological evolution we need explanations of how the relevant evolutionary mechanisms work. Insofar as they are products of learning and development in individuals, we need detailed descriptions of what those mechanisms are and how they work.

Some of the problems were discussed, but without mechanisms specified, by Piaget in his last two books Piaget (1981, 1983). The ideas about "Representational Redescription" described in Karmiloff-Smith (1992) also seem to be very relevant, as discussed earlier.

Although there are impressive computer based theorem provers, not all human mathematical competences have been modelled or replicated, some because nobody has tried, and others because they have proved difficult to specify or model computationally. In some cases that is because of the modelling resources chosen.

For example, McCarthy and Hayes (1969) proposed that a variant of first order logic should suffice for all AI purposes, in particular because it was metaphysically adequate (everything that needed to be expressed as true in AI theories could be expressed using logic) it was epistemologically adequate (all the knowledge that intelligent agents required could be expressed in logic) and it was heuristically adequate (logical forms of representation could suffice for efficient reasoning). Sloman (1971) suggested that they might be mistaken at least as regards heuristic adequacy. It is also possible that logical forms of representation are epistemologically inadequate to express the type of information used and the means of manipulating it when our ancestors first discovered Euclidean geometry.

More generally, the possibility of biological mechanisms and evolution of those mechanisms depends on the existence and reliability of physical mechanisms that achieve reproduction and maintenance of complex structures, discussed in more detail in Ganti (2003). Compare Kauffman (1995) (NB. Reliability does not need to be perfect.) Brian Goodwin's idea that there are "Laws of Form" constraining evolution may also turn out relevant here.[22]

The processes involved in physical reproduction, growth, repair and the processes involved in provision and use of energy and various kinds of information in organisms all have complex mathematical structures. The mechanisms that reliably achieve those functions also have mathematical structures. If the physical world

[22] See `http://en.wikipedia.org/wiki/Brian_Goodwin`.

could not support mechanisms that reliably operate within mathematical constraints required by complex enduring molecules and more complex multi-functional chemical structures require, then complex forms of life would not be able to endure, replicate and diversify.

17 Mathematical theory formation

When proofs have been created for a collection of different but related cases, it is sometimes possible to group them into an elegant theory covering all the cases in a systematic way: and that's what Euclid's Elements did for a large chunk of spatial reasoning about 2-D surfaces and 3-D structures, though I don't believe much is known abut the earliest stages of discovery leading to the Elements. Mathematicians and AI researchers are still extending that work, e.g. see Ida and Fleuriot (2012).

Aristotle attempted to do something similar for logic, but from the 19th century it was realised that he had barely scratched the surface; and mathematical studies of logical reasoning were extended by Boole, Peano, Frege, Russell, and others. The details are not important here except insofar as they reveal that far from logical reasoning providing the *basis* of mathematics, as some have suggested (e.g. Russell in Russell (1917)), logic is itself a topic that can be studied mathematically, with different systems of logic explored in much the same way as different systems of geometry and different number systems have been.

For example, one form of logical reasoning is the disjunctive syllogism, e.g. from premisses (P1) "Fred is in the kitchen **or** in the bathroom" (P2) "Fred is **not** in the bathroom" it follows logically that (C) "Fred is in the kitchen". It is sometimes suggested that in order to understand the validity of such reasoning one must (a) have learnt a rule of inference (e.g. the rule *From "p or q" and "not q" infer "p"*) and (b) must have noticed that the rule applies to this example – an idea famously challenged in Carroll (1895). But that's back to front: a rule like the disjunctive syllogism is simply a generalisation of what one can learn by inspecting special cases, whose validity is evident because if P1 asserts that there are only two possibilities, and P2 denies one of them, then if P1 and P2 are true the second possibility of P1 must be true. Instead of using such rules of logic in order to think mathematically, one needs to think mathematically in order to understand why such a logical rule is valid, which can start from seeing its validity in a particular case, and then generalising, just as reasoning about joining points on a cube can start from a particular case then generalising to all cubes, because they have the same structure as the particular case, at a certain level of abstraction. You may be able to illustrate this with a theorem you have discovered about rubber bands, or tying shoelaces, or uses of cupboards. For more on mathematical and logical necessity see Sloman (1965b, 1968/9).

18 Can we explain and replicate human geometric reasoning?

Among many competences that all humans seem to have, though they differ in the details, is the ability to learn euclidean geometry to some level, even though, sadly, this is not taught to all children. The kind of competence I am interested in is not the ability to memorise and recite definitions or theorems of geometry, e.g. the theorem that internal angles of a triangle sum to half a rotation, or Pythagoras' theorem. Neither is it the ability to use those theorems in practical applications.

Instead it is the ability to understand *why* a theorem is true: why there *cannot* be any counter examples. Some philosophers have argued that understanding that there can be no counter examples is possible only for truths that are either simple matters of definition, e.g. "All bachelors are unmarried" or can be derived from definitions using nothing but pure logic. For example, if "bachelor" is defined to mean "unmarried adult male" and "x is an uncle of y" is defined to mean "x is the brother of a parent of y or is married to the sister of a parent of y" then by logic we can derive the impossibility of an only child being a bachelor uncle.

What evolution has produced has many facets that are already the subject of intense scientific research, including microbes that altered the climate on earth, ecosystems in which millions of different types of organism can coexist, sub-microscopic organisms that can kill animals as large as whales or elephants, immune systems that resist such invaders, eggs that can turn themselves into caterpillars that turn themselves into soup-filled cocoons that turn themselves into butterflies, and many other amazing (and *beautiful* – a topic for another day) achievements.

The vast majority of research on evolution so far seems to have been concerned with (a) which physical structures, are products of biological evolution, including organisms of all sizes, shapes, and habitats, (b) how the physical structures grow and reproduce, (c) what sorts of behaviours organisms can produce in which environments, (d) what sorts of social or cultural behaviours, including forms of communication, are found in different species or in symbiotic interactions, (e) which physical/chemical mechanisms and processes are involved in the processes of reproduction, growth and development, and evolution, and (f) which sorts of sensory-motor morphologies and neural mechanisms make all those behaviours possible.

There is a further unifying collection of problems that seem not to have received much attention, namely: what are the varieties of information processing that make all those changes possible: e.g. what information contents, what information sources, what forms of representation of information, and what information-processing mechanisms are used, what are they used for, and how did they evolve on an initially lifeless planet?

One of the things that makes this research hard is that most of the processes and mechanisms involved in acquisition, analysis, interpretation, storage, derivation, combination, and use of information items are invisible. There are fossil records of products but not processes. Another source of difficulty is that many researchers fail to notice the need for some forms of information-processing because their assumptions about what animals can do, or need to do in order to produce observed behaviours are mistaken. A few examples follow:

- *Vision as pattern recognition:* There is much research on animal or machine vision that studies only ways in which portions of images can be segmented and labelled as belonging to some category, e.g "ball", "cow", "desk", or "elephant". This ignores many facts about vision, such as that complex objects can be seen and acted on (e.g. pushed, broken, eaten) without being recognized or categorized. It also ignores the fact that an important aspect of visual perception is identification of three-dimensional structure, e.g. the existence of surface fragments that have different curvatures or orientations, and the varying spatial relationships between parts, e.g. these thorns grow out of the same stem, but point in different directions.
- *Vision as reversing the projection process:* Many researchers, especially those influenced by David Marr Marr (1982) believe that the function of animal vision is to take in retinal image data and infer the 3-D structures from which light bounced into the viewer's eye.
  There are many problems with this proposal, including the impossibility in general of reversing the process, though in special cases, e.g. using two eyes and stereo decoding mechanisms it may be possible to compute depths, orientations, curvatures, etc. of visible surface. Gibson (1979) in contrast argued that the function of animal vision, and more generally animal perception, was not to reverse the projection process but to acquire information about actions available to the organism to enable it to meet its biological needs.

Gibson's general point is right, but the types of information organisms may need are more subtle and complex than he proposed. Some of the additional functions are closely related to perception of mathematical relationships, e.g. characterising not just what is actually happening and how it relates to possible actions of the perceiver, but all sorts of possibilities inherent in a situation and constraints on those possibilities. For more largely unnoticed functions and limitations of animal vision and requirements for possible mechanisms see Clowes (1971); Sloman (1978a, 1982, 1989, 1996, 2011), Hochberg's work reported in Peterson (2007) and many others focusing on different subsets of the phenomena.

19 The Side Stretch Theorem (SST) and area of a triangle

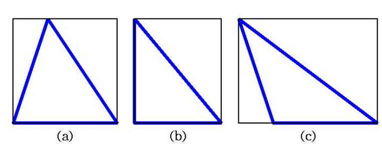

Fig. 6 Some additional lines added to these figures allow reasoning about equivalences between different areas in the diagrams, leading to a formula for area of a triangle. Can you derive it?

Many non-mathematicians have learnt a formula for computing the area of a triangle even if they cannot prove that it is correct. Most are unaware of the problems of defining the notion of "area" for an arbitrary shape, eventually solved by using the notion of two limits as the shape is filled with successively smaller squares so that as the squares get smaller the outer boundary of the enclosed squares increasingly approaches a limit, as does the outer boundary of all the squares contained inside or overlapping the boundary. This notion goes back to ancient Greece.[23] An extended online discussion of ways of thinking about this is still under development, exploring ways of thinking about the area of a triangle and how to compute it.[24]

The "standard" proof of the formula for the area of a triangle uses of ways of fitting triangles into rectangles and employs a prior formula for area of a rectangle, as hinted at in Figure 6. Readers are invited to reconstruct a proof.

For now, let's focus on a much simpler theorem, the Side-Stretch Theorem depicted in Figure 7, which, as far as I know has never previously been noticed, though it may be equivalent to a known axiom or theorem in one of the many formulations of Euclidean geometry.

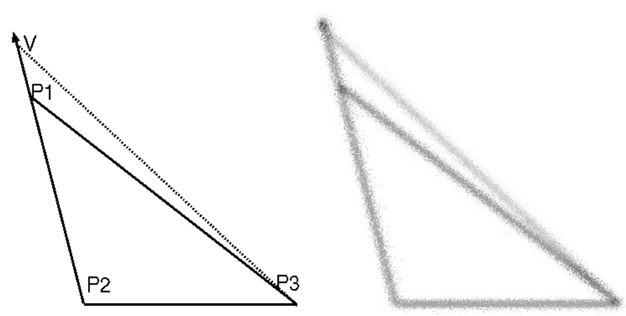

Fig. 7 The Side Stretch Theorem (SST): extending or shrinking the length of side P_1P_2 alters the area of the triangle. If point P_1 moves further from vertex P_2 then the area of triangle $P_1P_2P_3$ increases. Thinking about this can use a messy, blurred diagram, on paper or imagined

SST states that IF a vertex (P_1) is moved on a fixed line (from P_1 in direction V) while the far end of the line (P_2) is unchanged and another side of the triangle (P_2P_3) remains unchanged, then the area of triangle $P_1P_2P_3$ will increase. In order to prove this theorem it is necessary to think about what can change in the diagram in Fig 7, and what the consequences of such a change would be.

The relationship between direction of motion of the vertex V and whether the area increases or decreases, or the changes in containment corresponding to motion of V, can be seen to be *invariant* features of the processes. But it is not clear what information-processing mechanisms make it possible to discover that invariant

[23] `http://en.wikipedia.org/wiki/Integral#Pre-calculus_integration`

[24] `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html`

relationship before the connections between geometry and coordinate systems have been noticed.

We are not discussing probabilities here, only what's possible or impossible. A robot with statistical learning capabilities might discover a regular relationship between moving the vertex up or down whether a new shape contains or is contained by an old shape, but lack the ability to discover that the connection is inviolable.

All the mathematical discoveries we have been discussing are very different from the discoveries currently made by AI programs that collect large numbers of observations and then seek statistical relationships in the data generated, which is how much robot learning is now done. The kind of learning described here, when done by a human, does not require large amounts of data, nor use of statistics. There are no probabilities involved, only invariant relationships: if a vertex P_1 moves along one side, away from the opposite end of that side P_2, and the other two vertices, P_2 and P_3 do not move, then the new triangle must contain the old one. Therefore the new area must be larger than the old one. This is not a matter of a high probability, not even 100% probability. It is about what combinations of states and processes are *impossible*.

A "virtual discontinuity" can be based on continuous change. As a vertex V moves along one side *away* from the other end of that side, P_2, the area will increase continuously. However, if the vertex moves in the opposite direction, *towards* the other end, i.e. V moves toward P_2, then the change in direction of motion necessarily induces a change in what happens to the area: instead of *increasing*, the area must *decrease*.

As the vertex moves with continuously changing location, velocity, and/or acceleration, there are some unavoidable *discontinuities*: the direction of motion can change, and so can whether the area is increasing or decreasing. But there are more subtle discontinuities, which can be crucial for intelligent agents.

A "virtual discontinuity" can occur during continuous motion with fixed velocity and direction. If a vertex V of a triangle starts beyond a position P_1 on one of the sides of the triangle and moves back towards the other end of the line, P_2, then the location of V, the distance from the other end, and the area of triangle VP_2P_3 will all change *continuously*.

But, for every position P on the line through which V moves, there will be a *discontinuous* change from V being further than P from the opposite end (P_2), to V being nearer than P to the opposite end. Likewise, there will be a discontinuous change from *containing* the triangle with vertex P to *being contained* in that triangle, and the area of the triangle with vertex V will change discontinuously from being greater than the area of the triangle with vertex P to being smaller. Between being greater, or containing, to being smaller, or being contained there is a state of "instantaneous equality" separating the two phases of motion.

These discontinuities are not *intrinsic* to the motion of V, but involve a *relationship* to a particular point P on the line. The same continuous motion can be interpreted as having different virtual discontinuities in relation to different reference points or structures on the route of the change.

If an observer has identified the location P, the discontinuity may be noticed by the observer. But there need not be any observer: the discontinuity exists in the space of possible shapes of the triangle as V moves along one side.

There are many cases where understanding mathematical relationships or understanding affordances involves being able to detect such virtual discontinuities based on relational discontinuities (phase changes of a sort). For example, a robot that intends to grasp a cylinder may move its open gripper until the ‘virtual’ cylinder projected from its grasping surfaces down to the table contains the physical cylinder. Then it needs to move downwards until the gripping surfaces are below the plane of the top surface of the cylinder, passing through another virtual discontinuity. Then the gripping surfaces can be moved together until they come into contact with the surfaces of the cylinder: in that case a physical, *non-virtual*, discontinuity. An expert robot, or animal, instead of making the three discrete linear motions could work out (or learn) how to combine them into a smooth curved trajectory that subsumes the three types of discontinuity. But without understanding the requirements to include the virtual discontinuities a learning robot could waste huge amounts of time trying many smooth trajectories that have no hope of achieving a grasp.

After discovering this strategy in relation to use of one hand a robot or animal may be able to use it for the other hand. Moreover, since the structure of the trajectory and the conditions for changing direction are independent of whose hand it is, the same conceptual/perceptual apparatus can be used in perceiving or reasoning about the grasping action of another individual capable of grasping with a hand. It may even generalise to other modes of grasping, e.g. using teeth, if the head can rotate. (The concept of a “mirror neuron” which fires both when the perceiver performs an action and when another individual performs the same action, may actually be part of a detector for changing mathematical relationships: the same changes in relationships between surfaces can occur whether some of the surfaces are parts of the perceiver or parts of something else. This could be an important source of generalisation of competences, and have nothing to do with perception of actions of conspecifics.)

The examples above show that a perceiver looking at a triangle may not only able to see and think about the particular triangle displayed, but can also use what is learnt from the perceived triangle, e.g. The Side-Stretch theorem or other theorems, to support thinking and reasoning about large, indeed infinitely large, sets of *possible* triangles, related in different ways to the original triangle.

The concept of an infinitely large set being used here is subtle and complex and (as Immanuel Kant noted) raises deep questions about how it is possible to grasp such a concept. For this discussion we need only note that if we are considering a range of cases and have a principled means of producing a new case different from previously considered cases, then that supports an unbounded collection of cases.

20 Implications for biological meta-cognition

In some cases the new configurations thought about in mathematical discoveries include additional geometrical features, specifying constraints on a new figure, For example, after considering a vertex moving along a side of a triangle we can consider it moving along a median, or along a line parallel to or perpendicular to one of the sides.[25] Such constraints, involving lines or circles or other shapes, can be used to limit the possible variants of a starting shape, while still covering infinitely many cases. However, the infinity of possibilities is simplified by making use of common features, or invariants, among the infinity of cases. Some mathematical thought experiments are more complex than others. The Side-Stretch Theorem considered only two cases: a vertex moving further from or towards another vertex. However thinking about a vertex moving on a line parallel to the opposite side of the triangle, or a vertex moving on a line perpendicular to the opposite side produces more qualitatively distinct cases. For example, There are infinitely many perpendiculars to the base of a triangle, adding complexity missing in the previous examples. This is why mathematical reasoning often requires case-analysis, and if a case is missed a proof will be erroneous.

This ability to think about infinitely many cases in a finite way seems to depend on the biological *meta-cognitive ability* to notice that members of a set of perceived structures or processes share a common feature that can be described in a meta-language for describing spatial (or more generally perceptual) information structures and processes. An example would be noticing that between *any* two stages in a process there are intermediate stages, and that between any two locations on a line, two thicknesses of a line, two angles between lines, two degrees of curvature, there are *always* intermediate cases, with the implication that there are intermediate cases between the intermediate cases and the intermediate cases never run out.

(We can ignore the difference between a set being dense and being continuous – a difference that mathematicians did not fully understand until the 19th Century. I have been using “continuous” loosely to cover both cases.)

This ability to notice that some perceived structure or process is continuous, and therefore infinite, is meta-cognitive insofar as it requires the process of perceiving, or imagining, a structure or process to be monitored by another process which inspects the changing information content of what is being perceived, or imagined, and detects some feature of this process such as indefinite divisibility. A more complex meta-cognitive process may notice an invariant of the perceived structure or process, for instance detecting that a particular change necessarily produces another change, such as increasing area, or that it preserves some feature, e.g. preserving area.

This sort of ability contrasts with the ability to reason about discrete cases, for example logical structures where a new disjunct is added, or an expression is repeatedly negated. Researchers have been developing logical theorem provers (and planners) with meta-cognitive capabilities for several decades, since that is easy

[25] See Note 24

to do in list processing or logical programming languages (like Lisp and Prolog). There has been work on giving computers meta-cognitive reasoning abilities about spatial reasoning in cases where the spatial structures are discrete, e.g. arrays of dots or squares Jamnik (1999). Some progress on reasoning about continuity, in an automated tutor for calculus is presented in Winterstein (2005). But neither of those has the features of a child exploring spatial structures and discovering euclidean geometry.

For example, learners can find it obvious that the arguments used in connection with the Side-Stretch theorem, and earlier in connection with curves on a toroid or paths on a cube, do not depend on the locations, orientations, sizes, or exact shapes of the examples considered. So the proofs cover infinitely many different cases. This seems to be connected with the fact that diagrams used in mathematical reasoning do not have to exhibit mathematical precision, which would normally be impossible anyway, when producing proofs on a blackboard or sheet of paper. In a logical or algebraic formalism, generality is typically achieved by using variables for which many different constants can be substituted. In geometrical and topological reasoning generality may be expressed in the fact that a particular structure is capable of being continuously deformed into many different cases without important structural relationships changing.

Noticing an invariant topological or geometrical relationship by abstracting away from details of one particular case is very different from searching for correlations in a large number of particular cases represented in precise detail. For example computation of averages, co-variance and various other statistics requires availability of many particular, precise, measurements, whereas the discovery processes demonstrated above do not require even one precisely measured case.

An observation made in Lenat and Brown (1984) concerning Lenat's AM and Eurisko programs, which learnt through unsupervised exploration of structured domains, is relevant here: *"If the language or representation employed is not well matched to the domain objects and operators, the heuristics that do exist will be long and awkwardly stated, and the discovery of new ones in that representation may be nearly impossible."* We may one day discover that the sort of design effort that AI researchers have to put into some of their programs for independent exploration and discovery were preceded by the "design effort" of evolution in producing animals with powerful learning abilities tailored to the environments in which they evolved, as opposed to the totally general learning mechanisms sought by some AI researchers. Moreover, learning about spatial structures may require very different forms of representation from learning about arithmetic or programming.

The transitions in biological information processing required for organisms to have this sort of meta-cognitive competence are still largely unknown. But I suspect they form a very important feature of animal intelligence in organisms as different as dolphins, squirrels, elephants, crows and apes. The genetic mechanisms underlying such forms of learning and reasoning may share commonalities across different animal forms, despite differences in sensors, effectors and modes of locomotion, because the mechanisms address common environmental challenges, such as finding routes, disassembling food, bringing food to helpless young in a shelter or nest, moving obstacles out of the way or avoiding them, building shelters, using matter to manipulate matter in various ways. The commonalities come from the spatial structures of environments rather than morphological commonalities of organisms.

Such evolved, possibly convergently evolved, features could then have provided the basis for further evolutionary transitions, in some species, including development of meta-meta-meta... competences required for some aspects of human intelligence, as speculated in Chappell and Sloman (2007). Perhaps evolution first produced organisms that are able to detect and make use of invariant relationships, e.g. in controlling movements, selecting materials, using tools, and only later produced the meta-cognitive competences required for noticing the reasoning processes, although the meta-cognitive competences may then have supported more advanced versions of the cognitive competences.

Allowing results of such learning and reasoning to be communicated to other individuals required further evolutionary steps. I have speculated that this started with modifications of actions allowing and then enhancing the side-effect of communicating intentions and indicating where help might be useful. These action-modifications might later have developed into sign languages, before there was vocal communication Sloman (1979, 2008). Further research, including experimental AI modelling, is required.

Some of these points about the need for intelligent systems to use diverse forms of representation were made, though perhaps less clearly in Sloman (1971) which also emphasised the fact that for mathematical reasoning the use of external diagrams is sometimes essential because the complexities of some reasoning are too great for a mental diagram. Every mathematician who reasons with the help of a blackboard or sheet of paper knows this, and understands the difference between using something in the environment to *reason* with and using physical apparatus to do *empirical* research, though it took some time for many philosophers of mind to notice that minds are extended.

21 Toward robot mathematicians discovering geometry

It is likely to be some time before we have robot mathematicians that understand, or independently discover the Side-Stretch theorem and similar theorems, or discover the elegant proof of the Triangle Sum theorem by Mary Pardoe shown in Figure 8, and discussed in `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-sum.html`, or which can think about how to compute the area of a triangle, or can discover the existence of prime numbers by playing with blocks (in the manner described in Sloman (2013a)), or can perceive and make use of the many different sorts of affordance that humans and other animals can cope with (including, in the case of humans: proto-affordances, action affordances, vicarious affordances, epistemic affordances, deliberative affordances, and communicative affordances), illustrated in Sloman (2011).

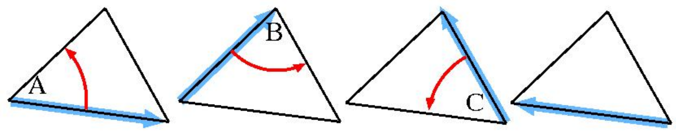

Fig. 8 Mary Pardoe's proof of the triangle sum theorem (see text)

A deeper question is whether there are features of the information-processing engines developed and used by evolution that are not modelled in Turing machines or modern digital computing systems, or which have totally intractable complexity on Turing machines but not on special purpose spatial-reasoning machinery.

If there are important differences between Turing machines and chemical computers they may depend on some of the following:

- Chemical processes involve both continuous changes in spatial and structural relations and also the ability to cross a phase boundary and snap into (or out of) a discrete stable state that resists change by thermal buffeting and other processes. This stability could rely on quantum mechanisms.
- They allow multiple constraints to be exercised by complex wholes on parts, that permit certain forms of motion or rotation or chemical behaviours but not others.
- Some switches between discrete states, or between fixed and continuously variable states can be controlled at low cost in energy by catalytic mechanisms. Compare Kauffman (1995).

It is clear that organisms used chemical computation long before neural or other forms were available. Even in organisms with brains, chemical information processing persists and plays a more fundamental role (e.g. building brains and supporting their functionality). Does chemical computation have a more important role than anyone has realised in supporting evolution, or individual learning or reasoning? This is just a question: I have no answers at present. The CogAff architecture schema and ideas about evolved virtual machinery may provide a framework for gap-filling: `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html`

22 Conclusion

Despite presenting more questions than answers I have tried to explain in outline connections between the possibility of mathematical thinking, exploration, discovery and proof and the existence of a huge variety of domains that include structures, relationships, and processes about which discoveries can be made by thinking in a special way, using *offline* intelligence, that is very different from the processes of collecting and summarising evidence for use in *online* intelligence, that now dominate much of AI including robotics.

I have cast doubt on “deflationary” theories that suggest that mathematics is a product of human thought or culture, or of biological evolution, while allowing that some of the processes of evolution and learning may instantiate areas of mathematics whose possibility is part of the fabric of reality, whether noticed by anyone or not.

A hypothesis that some may find very surprising is that the possibility of biological evolution of kinds hypothesised here implicitly depended on natural selection “discovering” forms of mathematical structure and derivation that were at first “blindly” used in production of new, increasingly complex types of individuals, then later implemented in learning and discovery processes in individuals who thought about and found new uses for those mathematical principles.

At a still later stage evolution produced mechanisms allowing some of those mathematical thinkers to become aware of some of their thinking and in some cases to find ways of criticising and improving it. On other planets where life emerges this might not happen – as it did not happen here for several billion years (as far as we know).

This form of meta-cognition, supported communication between individuals about mathematical reasoning and eventually produced massive collaborative activities leading to Euclid’s elements and other collections of communicable mathematical results. In the process the individuals became increasingly aware of what they were doing and able to communicate thoughts the nature of the activity. This generated disagreements about the nature of mathematics that remain unresolved, but perhaps the theories presented here will one day lead to agreement.

These processes can be construed as natural selection (or the biosphere) undergoing a process analogous to what Karmiloff-Smith calls “Representational Redescription” when it happens in individuals (Karmiloff-Smith, 1992).

These processes were recently enormously enhanced by the availability of digital computers and the development of many computational tools and applications. But some apparently simple forms of human and animal reasoning, e.g. about affordances and properties of geometrical shapes, have proved hard to implement on computers and it is possible that importantly different forms of computing machinery, will be required before robots can match a wider range of animal capabilities. Perhaps the Turing-inspired meta-morphogenesis project will generate new ideas about how this could happen.

Acknowledgements I would like to acknowledge the stimulation of philosophical discussions about mathematics with Ian Hacking many years ago when we were both students and more recently by email. Mary Leng dragged my attention back to philosophy of mathematics after it had drifted, by inviting me to give a talk in Liverpool in 2008, which forced me to pull many lines of thinking together. Alison Sloman has helped me by teaching me biology and reacting to many of my examples and forcing clarification or correction. Our grandson and many other babies and toddlers have provided examples of development of human proto-mathematical competences. Annette Karmiloff-Smith’s notions of “Representational Redescription” helped to focus and clarify my thinking about domains and how knowledge about a domain can be transformed. Jackie Chappell co-authored some of the work developing these ideas and improved my understanding of the diversity of evolutionary influences. Barry Cooper prompted new lines of thought by inviting me to contribute to Cooper and van Leeuwen (2013), especially part IV, and again by inviting this contribution. I have discussed these issues intermittently with Alan Bundy since 1973. Jacques Fleuriot drew my attention to examples of recent work on geometry. Colleagues and students in the School of Computer Science, University of Birmingham have provided critical comments, examples and challenges over many years. Cathy Legg introduced me to relevant aspects of the work of C.S.Peirce. Justin Clarke-Doane usefully challenged these ideas in recent conversations. Squirrels and magpies in our garden have humbled me. If only Alan Turing had lived on and developed his ideas about chemical morphogenesis on an evolutionary scale....

23 Preamble: evolution of information processing

24 Pervasiveness of mathematical domains

25 Evolution and finite brains

26 Evolving mathematicians

27 Mathematical domains

28 What sorts of machines can be mathematicians?

29 A conjecture

30 Steps towards mathematical intelligence

31 Evolved and engineered solutions

32 Non-human mathematical competences

33 Domain of lines on a cube

34 Some discrete domains

References

Blanchette P (2014) The frege-hilbert controversy. In: Zalta EN (ed) The Stanford Encyclopedia of Philosophy, spring 2014 edn

Carroll L (1895) What the Tortoise said to Achilles. Mind 4(14):278–280, URL http://fair-use.org/mind/1895/04/what-the-tortoise-said-to-achilles

Chappell J, Sloman A (2007) Natural and artificial meta-configured altricial information-processing systems. International Journal of Unconventional Computing 3(3):211–239, URL http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#717

Clowes M (1971) On seeing things. Artificial Intelligence 2(1):79–116, URL http://dx.doi.org/10.1016/0004-3702(71)90005-1

Cooper SB, van Leeuwen J (eds) (2013) Alan Turing: His Work and Impact. Elsevier, Amsterdam

Dauben JW (1995) Abraham Robinson : The creation of nonstandard analysis. A personal and mathematical odyssey. Princeton University Press, Princeton, NJ,

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

Ganti T (2003) The Principles of Life. OUP, New York, Eds. Eörs Szathmáry & James Griesemer, Translation of the 1971 Hungarian edition

Geretschlager R (1995) Euclidean constructions and the geometry of origami. Mathematics Magazine 68(5):357–371, URL http://www.jstor.org/stable/2690924

Giaquinto M (2007) Synthetic a priori knowledge in geometry: recovery of a Kantian insight. URL http://www.ucl.ac.uk/~uctymdg/RECOVERY-OF-A-KANTIAN-INSIGHT-web.pdf, presented at The Royal Institute of Philosophy Annual Conference 2007: Kant and Philosophy of Science Today

Gibson JJ (1979) The Ecological Approach to Visual Perception. Houghton Mifflin, Boston, MA

Hilbert D (2005) The Foundations of Geometry. Project Gutenbert, URL http://www.gutenberg.org/ebooks/17384, translated 1902 by E.J. Townsend, from 1899 German edition

Holt J (1967) How Children Learn. Pitman Publishing

Ida T, Fleuriot J (eds) (2012) Proc. 9th Int. Workshop on Automated Deduction in Geometry (ADG 2012), Informatics Research Report, University of Edinburgh, Edinburgh, URL http://dream.inf.ed.ac.uk/events/adg2012/uploads/proceedings/ADG2012-proceedings.pdf

Irvine AD, Deutsch H (2013) Russell's paradox. In: Zalta EN (ed) The Stanford Encyclopedia of Philosophy, winter 2013 edn

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

Jamnik M (1999) Automating Diagrammatic Proofs of Arithmetic Arguments. PhD thesis, Edinburgh University, URL http://www.dai.ed.ac.uk/papers/documents/pt9915.html

Kant I (1781) Critique of Pure Reason. Macmillan, London, translated (1929) by Norman Kemp Smith

Karmiloff-Smith A (1992) Beyond Modularity: A Developmental Perspective on Cognitive Science. MIT Press, Cambridge, MA

Kauffman S (1995) At home in the universe: The search for laws of complexity. Penguin Books, London

Kessler G (1980) Frege, Mill and the Foundations of Arithmetic. Journal of Philosophy 77 Issue 2:65–79, URL http://people.virginia.edu/~gpk2n/kessler-FregeMillAndTheFoundationsOfArithmetic.pdf

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

Lakatos I (1980) Falsification and the methodology of scientific research programmes. In: Worrall J, Currie G (eds) Philosophical papers, Vol I, Cambridge University Press, Cambridge, pp 8–101

Lenat DB, Brown JS (1984) Why AM and EURISKO appear to work. Artificial Intelligence 23(3):269–294, URL http://www.aaai.org/Papers/AAAI/1983/AAAI83-059.pdf

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

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

Mueller I (1969) Euclid's Elements and the Axiomatic Method. British J for Phil of Science 20(4):289–309, URL http://www.jstor.org/stable/686258

Nicod J (1930) Foundations of Geometry and Induction. Routledge and Kegan Paul, London, UK, translated from earlier French edition

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

Peterson MA (2007) The Piecemeal, Constructive, and Schematic Nature of Perception. In: Peterson MA, Gillam B, Sedgwick HA (eds) Mental Structure in Visual Perception: Julian Hochberg's Contributions to Our Understanding of the Perception of Pictures, Film, and the World, OUP, New York, pp 419–428

Piaget (1983) Possibility and Necessity Vol 2. The role of necessity in cognitive development. U. of Minnesota Press, Minneapolis, tr. by Helga Feider from French in 1987

Piaget J (1981) Possibility and Necessity Vol 1. The role of possibility in cognitive development. U. of Minnesota Press, Minneapolis, tr. by Helga Feider from French in 1987

Polya G (1945) How To Solve It. Princeton University Press, Princeton

Popper K (1984) Critical Remarks on the Knowledge of Lower and Higher Organisms the So-Called Sensory Motor Systems. In: Creutzfeld O, Schmidt RF, Willis WD (eds) Sensory-Motor Integration in the Nervous System., Springer,, Berlin, pp 19–31

Rips LJ, 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 (1917) Mysticism and Logic and Other Essays. Allen & Unwin, London

Russell SJ, Wefald EH (1991) Do the Right Thing: Studies in Limited Rationality. MIT Press, Cambridge, MA

Sauvy J, Sauvy S (1974) The Child's Discovery of Space: From hopscotch to mazes – an introduction to intuitive topology. Penguin Education, Harmondsworth, 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. PhD thesis, Oxford University, http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#706

Sloman A (1965a) Functions and Rogators. In: Crossley JN, Dummett MAE (eds) Formal Systems and Recursive Functions: Proceedings of the Eighth Logic Colloquium Oxford, July 1963, North-Holland Publishing Co, Amsterdam, pp 156–175, URL `http://tinyurl.com/BhamCog/07.html#714`

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

Sloman A (1968/9) Explaining Logical Necessity. Proceedings of the Aristotelian Society 69:33–50, URL `http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#712`

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

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

Sloman A (1978b) What About Their Internal Languages? Commentary on three articles by Premack, D., Woodruff, G., by Griffin, D.R., and by Savage-Rumbaugh, E.S., Rumbaugh, D.R., Boysen, S. in *Behavioral and Brain Sciences* Journal 1978, 1 (4). Behavioral and Brain Sciences 1(4):515, http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#713

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

Sloman A (1982) Image interpretation: The way ahead? In: Braddick O, Sleigh A (eds) Physical and Biological Processing of Images (Proceedings of an international symposium organised by The Rank Prize Funds, London, 1982.), Springer-Verlag, Berlin, pp 380–401, URL `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, URL http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#7

Sloman A (1996) Actual possibilities. In: Aiello L, Shapiro S (eds) Principles of Knowledge Representation and Reasoning: Proc. 5th Int. Conf. (KR '96), Morgan Kaufmann Publishers, Boston, MA, pp 627–638, URL http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15

Sloman A (2006) Requirements for a Fully Deliberative Architecture (Or component of an architecture). Research Note COSY-DP-0604, School of Computer Science, University of Birmingham, Birmingham, UK, URL http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604

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

Sloman A (2009a) A demonstration of unconscious seeing. URL http://www.cs.bham.ac.uk/research/projects/cogaff/misc/unconscious-seeing.html, online demonstration

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

Sloman A (2010) How Virtual Machinery Can Bridge the "Explanatory Gap", In Natural and Artificial Systems. In: Doncieux S, et al (eds) Proceedings SAB 2010, LNAI 6226, Springer, Heidelberg, pp 13–24, URL http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#sab

Sloman A (2011) What's vision for, and how does it work? From Marr (and earlier) to Gibson and Beyond. URL http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk93, online tutorial presentation, also at http://www.slideshare.net/asloman/

Sloman A (2013a) Meta-Morphogenesis and Toddler Theorems: Case Studies. URL http://tinyurl.com/CogMisc/toddler-theorems.html, online discussion note

Sloman A (2013b) Virtual machinery and evolution of mind (part 3) meta-morphogenesis: Evolution of information-processing machinery. In: Cooper SB, van Leeuwen J (eds) Alan Turing - His Work and Impact, Elsevier, Amsterdam, pp 849–856, URL http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1106d

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

Sloman A, Chrisley R (2003) Virtual machines and consciousness. Journal of Consciousness Studies 10(4-5):113–172, URL http://www.cs.bham.ac.uk/research/projects/cogaff/03.html#200302

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

Turing AM (1936) On computable numbers, with an application to the Entscheidungsproblem. Proc London Math Soc 42(2):230–265, URL http://www.thocp.net/biographies/papers/turing_oncomputablenumbers_1936.pdf

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

Turing AM (1952) The Chemical Basis Of Morphogenesis. Phil Trans R Soc London B 237 237:37–72

Ugur E (2010) A Developmental Framework for Learning Affordances. PhD thesis, The Graduate School of Natural and Applied Sciences, Middle East Technical University, Ankara, Turkey, URL http://www.cns.atr.jp/~emre/papers/PhDThesis.pdf

Weir AAS, Chappell J, Kacelnik A (2002) Shaping of hooks in New Caledonian crows. Science 297(5583), URL DOI:10.1126/science.1073433

Winterstein D (2005) Using Diagrammatic Reasoning for Theorem Proving in a Continuous Domain. PhD thesis, University of Edinburgh, School of Informatics., URL http://www.era.lib.ed.ac.uk/handle/1842/642

Wittgenstein L (1978) Remarks on the Foundations of Mathematics, 3rd Edition. Blackwell, Oxford, Eds. G. H. von Wright, R. Rhees & G. E. M. Anscombe