

Video presentation with online notes
**Why can't (current) machines reason like Euclid
or even human toddlers?**
(And many other intelligent animals)[*]

**Invited talk for IJCAI 2017 Melbourne:
Workshop on Architectures for Generality and Autonomy**
<http://cadia.ru.is/workshops/aga2017/>

42 Minute Video

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/ai-cogsci-bio-sloman.webm>

Shorter extracts from video available [below](#).

Extended discussion [below](#).

Aaron Sloman

<http://www.cs.bham.ac.uk/~axs/>

[School of Computer Science, University of Birmingham](#)

[*] So much for fears about an AI singularity this century...

Notes and video lecture for remote presentation at IJCAI Workshop, Melbourne, 19 Aug 2017

Workshop: Architectures for Generality and Autonomy

<http://cadia.ru.is/workshops/aga2017/>

Detailed contents of this document [listed below](#).

This page is still under development (being reorganised)

27 Sep 2017: Sections that were in the wrong order are now fixed

29 Sep 2017: NEW SECTION [First draft specification of the non-Turing, but Turing-inspired, membrane machine.](#)

RELATED TALKS WITHOUT VIDEO

A 20 min. talk given in Geneva on 15th June 2017 at

[3rd Japan-EU Workshop on Neurorobotics](#), co-located with

[Building Bodies for Brains & Brains for Bodies on 16th June.](#)

The notes and the video also expand on a 20 minute presentation at

[ICCM 2017 July 2017 Warwick University](#)

ICCM Paper: http://www.cs.bham.ac.uk/research/projects/cogaff/sloman_iccm_17.pdf

Invited tutorial presentation (using <https://appear.in>) 3rd August 2017, at BICA 2017 in Moscow:

<http://bica2017.bicasociety.org/>

Expanded abstract for PTAl17, Leeds 3-5 Nov 2017

Huge but unnoticed gaps between current AI and natural intelligence ([HTML](#)) ([PDF](#))

THIS DOCUMENT (much expanded during Sept & Oct 2017) is

HTML <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/ijcai-2017-cog.html>

Short URL: goo.gl/dB9WYz

PDF <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/ijcai-2017-cog.pdf>

PDF is derived from the html version.

Video presentation (42 min) for the IJCAI AGA Workshop on 19th Aug 2017,
available here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/ai-cogsci-bio-sloman.webm>

More options <http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17>

(Using extracts from an earlier version of this web site.)

Four extracts from that video are linked separately below. Search for **[VIDEO LINK]**.

NB

The (unscripted) video lecture skipped through an earlier version of this web page.

I apologise for any confusion caused. The video covers only a subset of this document.

I also apologise for the lack of a document more closely related to the video, and for disorganisation of this document (work in progress!).

CONTENTS (Order partially arbitrary -- to be improved later)

(Some of the contents have explicit links to parts of the video presentation.)

[PARTIAL OVERVIEW](#)

[Related work in Biosemiotics](#)

[VIDEO PRESENTATION FOR IJCAI AGA WORKSHOP](#)

[Some Themes](#)

[Reasoning that is easy and reasoning that is difficult for computers](#)

[Pre-history of ancient mathematics](#)

[\(Propositional\) Logical reasoning: easy for computers](#)

[Figure Logic](#)

[Geometrical/topological reasoning: much harder for computers](#)

[Figure Stretch-internal](#)

[Figure Stretch-external](#)

[Comments beyond the video presentation](#)

[A NON-TURING, BUT TURING-INSPIRED MEMBRANE MACHINE?](#)

(Added during Sept 2017)

[What is mathematical discovery? \(Euclid, Kant and Einstein\)](#)

[Relation to Kant's philosophy of mathematics and Turing's latest ideas](#)

[Can AI reasoning systems replicate the ancient mathematical discoveries?](#)

[Non-local geometrical impossibilities](#)

[Figure Reutersvärd](#)

[Geometrical vs logical reasoning](#)

[Figure Multiblocks](#)

[Some challenges for the reader](#)

[Replicating human understanding of geometry](#)

[Mary Pardoe's proof of the triangle sum theorem](#)

[Kant-inspired notes on mathematical discovery](#)

[Hume's types of reasoning](#)

[Note: Max Clowes](#)

[Possibility and depictability](#)

[Relevance to \(cardinal and ordinal\) number competences](#)

(Expanded 12 Sep 2017)

[What are foundations for mathematics?](#)

[The Meta-Morphogenesis \(Self-informing universe\) project](#)

[CONJECTURE What would Turing have done after 1954?](#)

[Graham Bell's view of evolution](#)

[CONJECTURE: Evolution's uses of evolved construction kits](#)

[Figure FCK: Fundamental Construction Kit \(FCK\)](#)

[Figure DCK: Derived Construction Kits \(DCKs\)](#)

[Deep design discoveries](#)

[Why is this important?](#)

[Construction kits for information processing](#)

[Metacognition evolves](#)

[Biological uses of abstraction](#)

[Mathematical consciousness](#)

[Designing designs](#)

[Figure Evolution](#)

[A highly intelligent crow](#)

[Toddler with pencil video](#)

[Evolution of human language capabilities](#)

[Changes in developmental trajectories](#)

[Epigenesis in organisms and in species](#)

[Figure WAD:](#)

[Figure EPI:](#)

[Evolution as epigenesis](#)

[Other species](#)

[Disembodiment of cognition evolves](#)

[Are non-Turing forms of computation needed?](#)

[The importance of virtual machinery](#)

[The future of AI](#)

[Links to Examples](#)

[Angle trisection](#)

[REFERENCES](#)

[About this document](#)

PARTIAL OVERVIEW

Some deep, largely unnoticed, gaps in current AI, (e.g. gaps between human and artificial mathematics) and what Alan Turing might have done about them if he had lived several more decades after publishing his 1952 paper on The Chemical Basis of Morphogenesis ([Turing 1952](#))

There are deep explanatory gaps in Current AI, Psychology, Neuroscience, Biology and Philosophy, concerned with things we don't understand about the following:

- use of structured **internal** languages in many species, and in pre-verbal humans, required for expressing contents of percepts, motives, intentions, beliefs, plans, etc. (As explained in [Sloman\(2015\)](#))
- grasp of mathematical (e.g. geometrical, topological) features of spatial structures and processes, in humans and many other species (however, on this planet only humans, so far, can reflect on, talk about, and systematise what they know);
- gaps in the explanatory power of current neuroscience, psychology, philosophy and AI;
- gaps in philosophical thinking about types and functions of consciousness; (E.g. how can we get robots to have mathematical qualia like Euclid?)
- gaps in evolutionary theory regarding how major mechanisms and competences evolved;
- gaps in developmental theory regarding how mechanisms and competences develop from a fertilised egg;
- gaps in our understanding of the nature of mathematical discoveries, especially ancient discoveries, by Archimedes, Euclid, Pythagoras, Zeno and many others; ([Euclid's Elements available online](#))
- gaps in our understanding of unreflective mathematical competences in other species (e.g. how adult elephants represent information about a baby elephant in trouble in a fast flowing river and work out how to rescue it, or how weaver birds make nests).

Some thinkers are excessively optimistic about what we already understand about the current state and future prospects of AI, and currently used forms of computation, while others are excessively pessimistic, about all forms of computation.

Both extremes need to be rejected in favour of new deep investigations into what organisms actually can do, how they do it, and what it is possible to replicate in human-designed machines including current computers, and possible future machines.

The accompanying 42min [video lecture](#) and this document are intended to give a taste for a middle way: a relatively new kind of exploration of abilities of biological evolution to produce new forms of information processing -- i.e. new forms of computation in a more general sense of computation than Church-Turing computation, including chemical computations, involving a mixture of continuous and discrete changes, instead of only discrete changes: a topic Turing was investigating before he died [Turing\(1952\)](#). His paper on chemistry based morphogenesis triggered the idea of the "Meta-Morphogenesis" (M-M) project, or "Self-informing universe" project, described below, first proposed in [Sloman \(2013b\)](#). The project extends earlier research on Kant-inspired philosophy of mathematics, now presented in a new context.

ADDED 4 Oct 2017:

Towards a "Super Turing Membrane Machine" [below](#) and in [related document](#).

NOTE: I suggest that any theory of consciousness that cannot accommodate mathematical competences of the sorts discussed here is worthless, even though most philosophers who write about consciousness ignore mathematical consciousness. Immanuel Kant was a notable exception in [1781](#).

Related work in Biosemiotics

Added 14 Aug 2017:

Alexei Sharov pointed me at deep, closely related work by the [Biosemiotics](#) research community, e.g.:

<http://www.biosemiotics.org/biosemiotics-introduction/>

<https://en.wikipedia.org/wiki/Biosemiotics> <https://en.wikipedia.org/wiki/Zoosemiotics>

I shall later try to write something about the connections, and will add links to documents on this site.

[VIDEO LINK]

VIDEO PRESENTATION FOR IJCAI AGA WORKSHOP 19 AUG 2017 (Approx 42 min)

This is available at

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17>

The video has the following main parts

- Brief introduction to the problem of accounting for ancient mathematical competences.
- Demonstration and brief discussion of related competences of weaver birds and other animals.
- Demonstration of 17.5 month toddler apparently forming and testing conjectures about 3D topology.
- Links to ancient mathematical discoveries assembled and documented by Euclid.
- Key features of such mathematical discoveries identified by Kant but not yet replicated in AI systems, or understood by neuroscientists.
- Examples of spatial reasoning about consequences of stretching a triangle in one dimension. (This text adds discussion of reasoning about possible and impossible 3D configurations of blocks, [below](#).)
- Introduction to key ideas of the M-M project, including evolution of new derived construction kits, especially construction kits for information processing systems.
- Implications for theories of epigenesis: Waddington's Epigenetic Landscape contrasted with the "Meta-configured Genome" theory: layered/staggered stages of gene expression where abstract later layers derive parameters for instantiation from products of lower layers interacting with the environment.
- Conjecture that this may hold clues about what's missing in (current) AI, including perhaps new chemistry-based forms of computation (suggested by Turing's work).

After a brief high level overview, the video comments on extracts from a BBC video showing weaver birds at various levels of competence building nests. The builders have to develop rich topological competences, e.g. concerning knots and the processes by which knots can be made, including formation of loops, grasping of free ends, pushing a free ends through a loop, and pulling a loop tight. (The video shows only short fragments of nest construction. The full weaver bird video is <https://www.youtube.com/watch?v=6svAlgEnFvw>)

The presentation then comments on a video recording of a pre-verbal child who is apparently able to think about 3-D topology, and do topological experiments, before she is able to talk about such matters. The relevant portion of the full video is available separately here:

[VIDEO LINK]

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/#infant-pencil>

The original toddler+pencil video (without commentary) was previously linked from a web page discussing "Toddler theorems" that illustrate mathematical competences used unwittingly by young humans [Sloman, \(2013c\)](#).

Some examples of differences between logical and geometrical reasoning are used to illustrate features that make logical reasoning (in propositional calculus) much easier to replicate (or simulate?) in a computer than ancient geometric reasoning. It includes some possibly new, partially analysed, examples of geometric reasoning.

[VIDEO LINK]

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/#geom-discovery>

The next section of the video introduces some of the ideas about evolution that have emerged since 2012, in the Turing-inspired Meta-Morphogenesis project, ([the M-M project](#)) including [the theory](#) of fundamental and derived construction kits of various sorts, and aspects of biological evolution that may eventually lead us to understand how biological evolution, later followed by cultural evolution, could have produced the competences of ancient mathematicians, building on competences shared with other intelligent species.

[VIDEO LINK]

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/#m-m-intro>

The video ends with a very brief introduction to a new theory of epigenesis, also summarised [below](#).

[VIDEO LINK]

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/#epigen>

Note: The original toddler+pencil video (without commentary) was previously linked from a web page discussing "Toddler theorems" that illustrate mathematical competences used unwittingly by young humans [Sloman, \(2013c\)](#).

Some Themes

Many animal competences, including competences of toddlers and weaver birds, that are not results of statistical learning or inferred probabilities, are related both to mathematical features of the environment, and to mathematical aspects of motor control, path selection, and perception of structures, and functions. Biological sensors are not good at precise measurement. Instead, visual systems seem to make use of partial orderings of spatial distances and sizes of perceived objects and object-parts. In many contexts these structural (semi-metrical) relationships are both more accessible and more useful than the inaccessible metrical details, e.g. distances, sizes, angles, times, etc. (Stereo vision is accurate only for nearby objects.)

Many robot designers attempt to give robots abilities to perceive and reason about precise distances, sizes, etc. resulting in systems that are not very robust. An alternative, which I think evolution "discovered" long ago, is to use comparisons rather than absolute values, i.e. partial orderings not total orderings of size, distance, area, angle, etc. [Sloman \(2007-2014\)](#). I'll try to show [below](#) how such partial orderings can have important roles in mathematical discovery.

Conjecture: abilities in pre-verbal humans and other animal species, to detect and reason about semi-metrical spatial proto-affordances, were evolutionary precursors of mathematical competences of the great ancient mathematicians, who identified and thought about some of the mathematical structures and processes in the environments in which humans and other animals evolve and behave.

The required meta-cognitive competences (discussed by Kant?), illustrated in examples below, include abilities to choose which possible changes to think about and how to think about them: selecting "affordances for mental action" in the terminology of [McClelland, \(2017\)](#). Many such affordances include attending to proto-affordances in the environment -- e.g. possible processes not related to the perceiver's needs or potential benefits.

Examples are presented in [Sloman\(2007-2014\)](#), [Sloman\(2008\)](#) and also below, e.g. in connection with [Fig Stretch-internal](#) and pictures of impossible scenes.

[BACK TO CONTENTS.](#)

Reasoning that is easy and reasoning that is difficult for computers

Some familiar forms of reasoning about geometrical structures and processes, seem to be impossible or very difficult to model using current computers, as illustrated [below](#). I'll present further examples involving non-local interactions between partial orderings, inspired by Reutersvärd's triangle [later](#).

There have also been philosophical and/or mathematical publications analysing, and in some cases defending, Euclid's use of diagrams, but without any attempt to automate reasoning with diagrams as a contribution to the cognitive science of mathematics. Examples are [Miller \(2007\)](#) and works by Ken Manders [\(1998\)](#), [\(2008\)](#). (I apologise if I have missed important aspects of their work.)

[[VIDEO LINK](#)] A video presentation contrasting logical and spatial reasoning is [here](#).

After presenting an example of purely logical reasoning easily implemented in computers, I'll present [below](#) an example of a mathematical discovery about effects of moving one vertex of a triangle that is close to Euclid's axioms for congruence, but different. I made the discovery while trying to find a simple unfamiliar example for a workshop presentation. I have no idea whether it has ever previously been noticed, but I am sure it will not be surprising to readers of this document. Likewise, the axioms in Euclid's *Elements* were important [discoveries](#), not arbitrarily adopted "starting points". I don't know of any theory of learning/perception that can explain or model those ancient discovery processes. I suspect that some of them are regularly repeated by young children, without anyone noticing.

Those discovery and reasoning processes, and the epistemological and semantic differences between mathematical and empirical discoveries pointed out by Kant, are usually ignored by developmental psychologists, neuroscientists, and (most) AI theorists.

(Piaget was an exception, though it seems to me that he lacked the required conceptual tools, i.e. computational tools, needed to say anything deep about explanatory mechanisms.)

My aim in presenting these examples, is to give my audience detailed first hand experience of what I claim is missing from AI (and not at all easy to add). These examples of reasoning about triangles, as far as I know, are not included in Euclid's *Elements* or in standard textbooks on Euclidean geometry, but should be obvious to most people, once they have encountered them.

I'll contrast those geometric discovery processes with reasoning in propositional logic, presented first, which can easily be modelled on computers.

I'll also contrast

(a) use of computers to simulate [spatial processes](#), as is commonly done in video games, weather forecasts, engineering design tools, etc.,

with

(b) use of computers to simulate [mathematical reasoning about spatial processes leading to new mathematical discoveries](#).

The latter is much harder to model computationally than the former.

[[VIDEO LINK](#)] The video presentation on this problem is available [here](#).

Pre-history of ancient mathematics

I think the deep, ancient, geometric discoveries build on deep "proto-mathematical" (or "pre-mathematical"?) competences shared with other intelligent species (e.g. nest building birds, squirrels, elephants, orangutans, hunting mammals), some of which are already present in pre-verbal humans as illustrated in the toddler with pencil video [below](#).

Some parts of these spatial reasoning competences are already present in pre-verbal human toddlers, though they are very hard to investigate because failure of a child to perform as an experimenter hopes says [nothing](#) about what that child can or cannot do. The problem may be a failure of communication, or a failure on the part of the experimenter to motivate the child or to trigger the right competences in the child, which could be triggered in a totally different social or non-social practical interaction. (This is one of the reasons studying and explaining [possibilities](#) is a

more fundamental aim of science than studying and explaining laws, as explained in [Chapter 2 of Sloman 1978](#),

I'll start with logical reasoning that is easy to replicate using current computers, and then go back to the more difficult challenge presented by the type of reasoning used in ancient geometric discoveries.

[BACK TO CONTENTS.](#)

(Propositional) Logical reasoning: easy for computers

A lot of work has been done on automated theorem proving using logic, arithmetic and algebra. This reasoning uses manipulation of notations allowing only discrete transitions, e.g. modifying a symbol, adding or removing a premise. The treatment of elementary propositional logic using truth-tables is similarly discrete. Elementary examples are concerned only with sets of possibilities in a finite space of possibilities. For example, in propositional logic, many students learn how operators related to the words, "not", "and", "or", "if ... then..." can be defined using truth tables. Then the validity of inferences can be checked using truth tables, though the sizes of the truth tables grow exponentially with the number of propositions (or propositional variables) involved.

A toy example is reasoning in propositional calculus, illustrated here:

Figure Logic	
Premises: P or Q not Q	By starting from the truth tables for "or" and "not" you can reason about which possible combinations of truth values for P and Q (and R in the second example) will make the premises true and ask whether the conclusion can be false in any of those cases.
Conclusion: P	
Premises: P or Q or R not Q	If it is impossible for all the premises to be true while the conclusion is false, then the inference is valid.
Conclusion: P	

Here the symbol ":-" can be read as "therefore", to indicate the conclusion of an inference to be evaluated.

In the first (upper) example, where only two propositions P and Q are involved there are only four possible combinations of truth values to be considered. And that makes it easy to discover that no combination makes both premises true and the conclusion false.

In the second case, for each of those four combinations R may be true or false, so the total number of possibilities is doubled. But it is still a finite discrete set and can be examined exhaustively to see whether it is possible for both premises to be true and the conclusion false. I assume the answer is obvious for anyone looking at this who understands "or" and "not". Checking for validity in propositional calculus involves exhaustive search through finite sets of discrete possibilities, so it is easy to program computers to do this. More generally, it is the discreteness of the structures and

transitions and the fact that there are only finitely many elementary components that allows such reasoning to be modelled on Turing machines, and their descendents digital computers.

Things get more complex if the propositional variables, e.g. P, Q, etc. instead of being restricted to two possible truth values (T and F) can take other intermediate values, or even vary continuously. The simple discrete reasoning, based on truth-tables, using **or** and **not** etc., will have to be replaced by something mathematically much more complex.

(a topic investigated by Tarski, Zadeh and others -- "Fuzzy Logic"

https://en.wikipedia.org/wiki/Fuzzy_logic)

That option will not be discussed further in this document, as it is not relevant to my main concern -- to understand very ancient geometrical and topological mathematical reasoning, concerning spaces of possibilities that are not composed of discrete combinations of discrete units, but continuously variable angles, lengths areas, curvatures, volumes, etc. Moreover, in a Turing machine or digital computer, elementary components are cleanly separated and it is not possible for one component to gradually occupy more of the space occupied by another, whereas in geometrical reasoning, illustrated below, entities can be superimposed moved continuously, and their shapes, parts and relationships can change non-discretely.

NOTE: The video associated with this document starts with examples of spatial and topological perception and reasoning by birds and a pre-verbal toddler. Many more examples of spatial intelligence, involving squirrels, elephants, nest-building birds, and other animals can be found by searching online videos. They are all relevant because the ancient mathematical abilities that are the core topic of this document build on, and are tightly connected with, a wide range of evolved perceptual, control, and decision-making capabilities in pre-verbal humans and other animals, about which I suspect very little is currently understood. When those abilities to perceive, reason, and plan are properly described it turns out that they are all missing from current AI programs and robots, including many that display impressive actions after much training or very careful programming (e.g. Boston Dynamics robots). For more on the discrepancies, see [Sloman \(2007-2014\)](#). I am not claiming that it is impossible to give future robots the required reasoning abilities, merely that there are important features of those abilities, closely connected with Kant's philosophy of mathematics, that are rarely noticed.

The gap to be bridged is between acquiring statistical generalisations from samples of a space of possibilities, and grasping necessary connections and impossibilities in the space.

[Necessity and impossibility are not points on a scale of probabilities.](#)

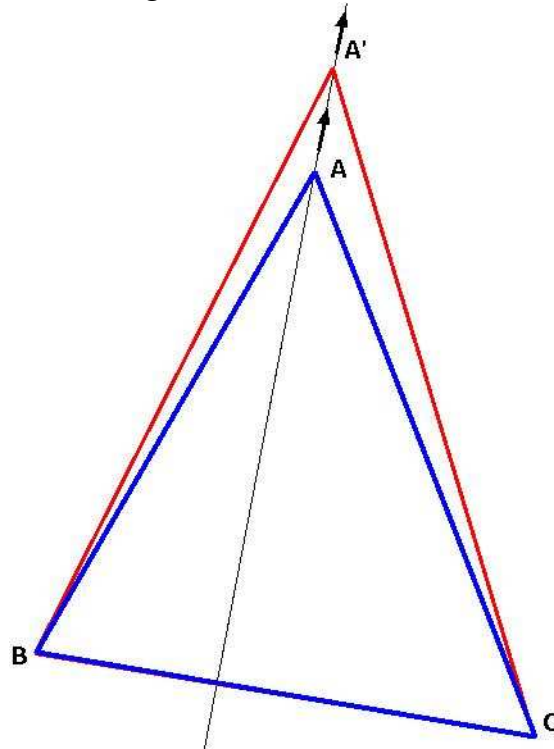
Geometrical/topological reasoning: much harder for computers

It is not yet clear whether computers can replicate the discoveries made by ancient mathematicians, over 2000 years ago, long before the discoveries in modern logic, formal reasoning, set theory, and Descartes' "translation" of geometry into arithmetic using coordinates.

Ancient geometrical discoveries, were not concerned only with collections of discrete possibilities. Euclidean geometry (like much of real life perception and action) is concerned with smoothly varying sets of possibilities, including continuously changing shapes, sizes, orientations, curvature and relationships between structures. In some cases a smooth change in one relationship can produce discrete changes, for instance a point moving smoothly in a plane, from outside a circle to inside the circle. I'll start with a deceptively simple example, and later reveal hidden complexity,

discussed in [a separate document](#).

Figure Stretch-internal



This example is presented in the accompanying video. [\[VIDEO LINK\]\(Vid 2\)](#)

Contrast the simple truth-table analysis above, involving only discrete options, with reasoning about deformation of a triangle. What will happen if you start with a triangle, like the blue triangle in [Figure Stretch-internal](#), and move the top vertex away from the opposite side along a line going through the opposite side? What happens to the size of the angle at the vertex, as the vertex A moves further from the base, BC? Here the red triangle illustrates one of the possible new triangles that could result from moving the top vertex along the thin line (which goes between B and C).

One way to reason about the above figure is to consider what happens if you move the **new** red vertex downwards while keeping the two red lines with the same orientations, so that the angle between them remains fixed, and the two sides retain their lengths. In that case as the new vertex moves down towards the old position with the angle between the red lines fixed, both the new red sides at their bottom ends will pass through the base of the old blue triangle between B and C, the two bottom corners of the blue triangle. Getting them to pass through the two bottom corners will require widening of the angle.

So, since the angle made at the top of the new triangle must be widened to match the original top angle, the new angle, produced by moving the vertex upwards while the base of the triangle is unchanged, must be smaller than the original angle at the top of the blue triangle.

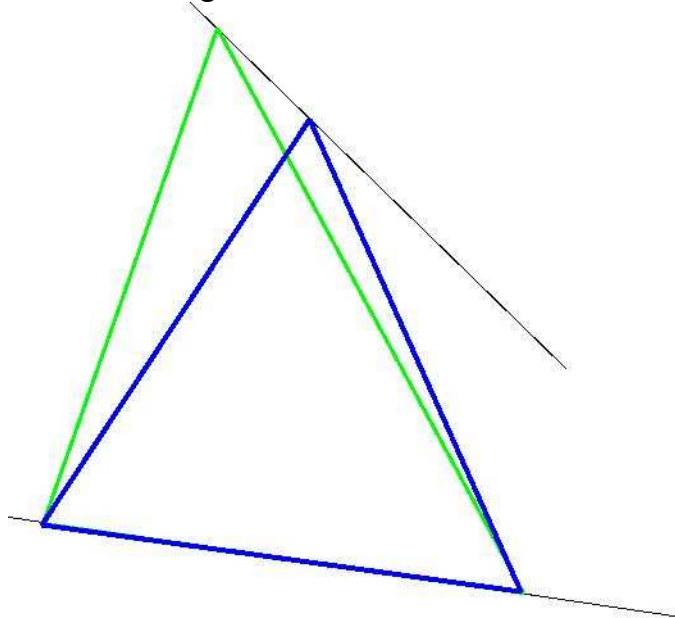
This shows that as the top vertex is moved further and further up the thin line, while continuing to pass through the two bottom corners, the angle between the two red sides will continually get smaller.

Another strategy is to notice that the line through the vertex divides the triangle into two smaller triangles sharing two vertices and one side. You can then consider what happens to each smaller triangle as the top vertex moves upward, or downward, along the line through the original vertex and the opposite side.

As far as I know this is not one of the standard theorems in Euclid's *Elements*. I don't even know whether anyone else has ever formulated this fact about Euclidean triangles, but I hope all readers will find this inference as obviously valid as the first logical example using "or" and "not" above, despite the fact that unlike the logical example, the triangle example involves an infinite (smoothly varying, "as if infinite") set of possible shapes rather than a discrete set of cases that can be examined exhaustively.

I am **not** claiming that no existing geometry theorem prover can derive answers to the questions about what happens to the angle at the moving vertex as it moves further from the opposite side of the triangle. For example, by using Cartesian coordinates for points and lines and it may be possible, using arithmetic, trigonometry, and a logically precise formulation of Euclid's axioms and postulates to derive the required inequality between before and after angle sizes. But that is almost certainly not what readers of this paper do, and it could not have been done by ancient mathematicians, since Cartesian coordinates were not invented until about two millennia later (https://en.wikipedia.org/wiki/Cartesian_coordinate_system).

Figure Stretch-external



Vertex moves upward along a line through base extended.

The problem of how the angle changes gets more complex if it is moved along a line that intersects the base outside the triangle, i.e. not between B and C, as in [Figure Stretch-external](#), above. Readers are invited to explore that case. A separate document offers a partial analysis (stimulated by Auke Booij when this talk was presented in Birmingham in Sept 2017), with some unanswered questions, here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/deform-triangle.html>

A further branch was later added in this document

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/apollonius.html>

after Diana Sofronieva heard me talk about this in Leeds on 3rd Nov 2017. She produced a

surprisingly complex and deep answer to one of the questions I raised, using the Apollonius construction in Euclidean geometry. Her answer is presented in a new separate document: <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/apollonius.html>

NOTE:

This problem about how [angles](#) change as one corner of a triangle moves is loosely related to ways of reasoning about how [the area](#) of a triangle changes as the triangle is deformed (the "area stretch" theorems) explored in [another document](#).

**Note: Comments beyond the video presentation
(Added during Sept -Oct 2017)**

I am not claiming that there are some facts of geometry or topology that are "self-evident", a type of claim with a long history discussed in [Shapiro \(2009\)](#). Instead, my questions are: What mechanisms make such mathematical discoveries possible? How do they work? When did they evolve? Which species have them? Can they be replicated on digital computers? How do they benefit organisms that have them? How could similar mechanisms benefit future robots or other AI systems?

Shapiro summarises the self-evidence claim (without endorsing it):

"We know the axioms, individually, to be truths about their subject matter, and this knowledge does not rely on anything else."

I discuss this notion in a separate document -- this one is not concerned with self-evidence, despite the historical importance of that notion in discussions of mathematical and logical knowledge. I regard it as a "red herring". See <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/self-evidence.html>

**A NON-TURING, BUT TURING-INSPIRED MEMBRANE MACHINE?
Added 27-28 Sep 2017 (Work in progress)**

I have some half-baked ideas for a specification for such a machine, which requires replacing important features of the specification of a Turing machine, with different (more general) features. E.g. instead of a tape there is a surface (something like a paintable, wipeable, stretchable, membrane) and instead of a finite set of symbols there is a finite set of mechanisms for producing or removing marks on the surface.

Perhaps it will need several concurrently active surfaces performing different storage functions on different spatial and temporal scales, with information constantly updated (after suitable transformations to compensate for saccades, etc.), as suggested in [Trehub\(1991\)](#).

Instead of the ability to identify discrete symbols in locations of a TM tape the new machine will need to be able to detect and compare structures, processes, and relationships between them, that can be produced or changed by operations on the surface, including continuous changes of part of the surface contents, and continuous changes of location and size of region currently under inspection. (Contrast the discrete linear array of readable/writable locations in a Turing machine.)

There is NO requirement for the structures, processes and relationships to be specified or achieved with high precision: I suspect the biological mechanisms used cannot support that (unlike digital computers). Instead it can produce, maintain, and detect [partial orderings](#) of location, size, curvature, thickness, direction, etc., such as would be useful for an organism controlling motion towards a target (maintain it approximately centrally in the field of view), avoiding contact with obstacles (maintain gaps between heading and line of sight to near edge of obstacle), or preparing to grasp something (move hands and hand-parts so as to decrease distances from contact points, and to keep gaps between grasping contacts larger than the diameter of the thing to be grasped, etc.

Continued in a separate document (Nov 2017):

The Super-Turing membrane machine, for perception and spatial reasoning.

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/super-turing-geom.html>

NOTE on 3D requirements(13 Oct 2017):

Most of the discussion here is restricted to 1-D and 2-D structures and processes. However we need something more general, e.g. for thinking about a screw rotating in a nut and why the circular rotation parallel to a plane forces a translation to occur perpendicular to the plane. Contrast the relationship between translation and rotation when a ball or a cylinder rolls along a surface without slipping.

Other examples include watching an animal climb a tree or assemble a nest, or dismember a carcass, all of which involve 3D structures and processes, about which it is possible to reason mathematically (e.g. what happens if the direction of rotation of a nut on a screw reverses?). This document ignores the 3D requirements (temporarily). Perhaps the word "membrane" needs to be replaced by something less constrained to 2D structures.

One of the requirements for this membrane would be that some of the items added to it will move under control of the external environment, e.g. if you watch a vehicle moving past along a road. Others will be added by internal information processing mechanisms for various purposes, and may be movable under internal control: e.g. to see if some perceived item can fit through a gap. One strategy is to insert a (rough) copy of the item and slide the copy to the perceived gap to see whether it can clearly fit within the gap. If it's a 3D scene, there may have to be shrinkage or expansion of the item moved depending on whether the gap is closer or further away than the item. All of these judgements will use low precision mechanisms and as a result not all such questions will have definite yes/no answers, as illustrated in [Sloman \(2007-2014\)](#). (Most robot vision systems are designed on totally different principles aiming for accuracy of perception of location, distance, size, etc. I suspect this results in mechanisms that are incompatible with human-like perceptual mechanisms, but may perform better on some narrowly constrained tasks.)

The mechanism proposed here is partly like "The FINST Visual Indexing Theory" of Zenon Pylyshyn and colleagues around 1990 summarised in

<http://ruccs.rutgers.edu/val-current-research/23-labs/val/231-the-finst-visual-indexing-theory>

However, I don't think Pylyshyn ever proposed using his mechanism to explain geometrical discoveries made by ancient mathematicians (and young children, intelligent non-human animals, etc.). The use was mainly to allow internal processes to refer to objects of attention.

It seems likely that evolution first produced versions of this control membrane for use in online control of actions, e.g. steering motion towards gaps, or around obstacles, or towards something to be grasped or eaten. I conjecture that at a later date it somehow modified the mechanisms to support [deliberative](#) operations, considering and reasoning about possible alternative futures: e.g. considering translating something in the surface to check the possibility of some motion (e.g. can something fit through a gap, or even can the individual itself fit through a gap, which requires a kind of "detachable self-model").

Evolution might later have produced mechanisms for reasoning about consequences of a possible motions of other objects unrelated to the perceivers actions, intentions, needs, etc. (I called these possible motions "proto-affordances" in [Sloman\(2008a\)](#). The label seems to have been invented independently more than once.)

I suspect that various evolutionary developments of these capabilities gradually produced richer and richer mechanisms, including [meta-cognitive](#) mechanisms that enabled the sort of reasoning used to answer the question posed in [Figure Stretch-internal](#), and many more.

It is possible that simplified, much less formal, versions of the processes of development of geometry without points presented by [Dana Scott](#) were used.

These are very much quarter-baked ideas that require much more precise and detailed specification (possibly answering some of the deep questions about vision raised by [Craik \(1943\)](#), e.g. how do messy neural mechanisms represent straightness?)

For reasons spelled out by Trehub, the contents of the surface will not correspond directly to retinal stimulation (or area V1 in brains), but will normally be constantly updated according to changes in retinal stimulation, and/or various predictive or "replay" mechanisms, perhaps using temporary copies of parts of the visual information structure.

If what humans are visually aware of corresponds closely to the contents of such a surface (or surfaces) then that would explain why we are not aware of the blind spot: its existence, i.e. the absence of visual information at that (changing) location, would never need to be recorded. Explaining this was not one of Trehub's intentions. It was a consequence of his theory that I pointed out to him.

In addition to information about contents of different locations in the visual field, the visual information store will need different sorts of information about recent history, and inferred processes (e.g. directions of motion of perceived items).

We should not jump to the conclusion that this relatively slowly changing record of rapidly changing retinal stimulation should consist of regular rectangular arrays that might be the first choice of a software engineer -- when available image-grabbing hardware has that structure, unlike anything in animal brains. Some AI vision researchers in the late 1960s (e.g. [Max Clowes](#)) proposed that computer vision systems could use logical descriptions of visual structures, including image structures and scene structures. But I suspect the required forms of logic were not available to biological evolution -- rather they were inventions of human logicians in the last few centuries. I suspect that some unknown form of "analogical" representation [Sloman \(1971\)](#) was used. [Trehub](#) made a similar suggestion.

Some of the mechanisms used for detecting and classifying spatial changes that occur will be related to mechanisms that allow the contents sometimes to be changed under system control rather than control by the environment. That could include system created "shadow copies" of parts of the available information that can be manipulated and compared with other parts, including historical records (e.g. needed for detecting rotations of enduring perceived objects).

These copying, manipulating, comparing, mechanisms should allow information to be recorded at various levels of abstraction, and for varying time-scales, that allow both useful online control decisions to be taken (e.g. changing direction while chasing a moving object, changing grasp while the hand approaches an object to be grasped) and also allows useful generalisations to be discovered.

There is a lot more to be said about the requirements for a non-Turing membrane machine, but the key idea is that we need to generalise Plato's example of Meno drawing in the sand at the same time as generalising Turing's ideas!

(I don't yet know exactly how these mechanisms relate to the ideas regarding affordances for mental action discussed by McClelland (2017), but there seem to be close connections.)

Humans, in addition, have **meta-cognitive** processes able to reflect on processes of perception and reasoning, and the relationships between the discoveries they lead to, whereas I suspect other species can merely use the abilities in considering, choosing and executing spatial actions, often in novel situations: perhaps a few species can also recognise impossibilities and necessary consequences, though explaining **how** their brains, and human brains, achieve this is likely to be non-trivial. Merely finding which portions of brains are involved, using standard procedures in neuroscience, does not explain what they are doing or how it works, in particular; how **necessary connections** between structures and processes are identified.

Without that there can be empirical discoveries, but not mathematical discoveries. A non-mathematician may stumble across the fact that attempts to arrange 11 buttons in a regular MxN array, with M and N both greater than 1 always fail, without understanding why. (Pat Hayes once informed me that a conference receptionist had complained that she could not always arrange the remaining delegate badges in a regular array. She did not realise that she had almost (re-)discovered prime numbers.) See <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html#primes>

I don't think anyone knows how brains could implement the sorts of mechanisms proposed here. It may be that Turing's interest in chemistry-based morphogenesis was just the first step in a potentially very complex investigation of uses of chemistry to implement information processing mechanisms like the proposed non-Turing membrane machine discussed above. (Still too ill-defined.)

NOTE:

The mechanism required here is related to "The FINST Visual Indexing Theory" of Zenon Pylyshyn and colleagues around 1990, summarised in <http://rucss.rutgers.edu/val-current-research/23-labs/val/231-the-finst-visual-indexing-theory> I don't know whether Pylyshyn attempted to use his mechanism to explain geometrical discoveries made by ancient mathematicians (and young children, intelligent non-human animals, etc.). I think the mechanism was mainly intended to allow internal processes to refer

to objects of attention, including moving objects. The ideas are relevant to my topic, though at present I only dimly recall the details. (I think Pylyshyn and I have discussed some of these issues in the very distant past.)

[BACK TO CONTENTS.](#)

What is mathematical discovery? (Euclid, Kant and Einstein)

Most AI researchers, like many psychologists, neuroscientists and even some philosophers ignore the problem of explaining features of human mathematical intelligence noted by Immanuel Kant in 1781, namely features of the kinds of discovery made by Euclid, Zeno, Archimedes, Pythagoras, and others over 2000 years ago, and even earlier in the case of Babylonian mathematicians

<http://www-history.mcs.st-andrews.ac.uk/Indexes/Babylonians.html>

As Kant pointed out, those ancient mathematical discoveries:

(a) are not empirical -- though they may be triggered/awakened by experience e.g. you don't need to have your eyes open when reasoning about geometry!

(b) are not derivable from definitions using logic (i.e. they are not analytic, in Kant's sense) -- and the discoveries are not necessarily made using modern logical/algebraic mechanisms, or starting from modern "foundations", discussed [below](#) (most of them listed in [Sakharov \(2003ff\)](#));

(c) are discoveries of non-contingent facts, i.e. [necessary](#) truths.

Analysing 'necessarily' (one of the main themes of [my 1962 DPhil thesis](#)) is non-trivial. Necessity has nothing to do with probability. Necessary truth and impossibility arise from structural constraints of various kinds, including topological and geometric constraints. So necessity and impossibility are very different from concepts of 100% and 0% probability, which refer to ratios of measures of some sort. Statistical evidence, e.g. repeated success or repeated failure in some attempted construction, may [suggest](#) a necessary connection or an impossibility, but such suggestions do not amount to proof, and they are often false.

This is why current "deep learning" mechanisms based on statistical evidence and probabilistic reasoning [cannot](#) replicate the ancient mathematical discoveries (and variants illustrated in this document, e.g. in the [Stretch-internal](#), [Stretch-external](#), [Reutersvärd](#), [Figure Multiblocks](#) examples): all require consideration of continuously varying collections of spatial possibilities without investigating instances of all the possibilities. For related reasons, the kinds of discovery about what is possible with a pencil and a hole, tested by the toddler presented in [the video](#) accompanying this document cannot have been based on statistical learning.

[\[VIDEO LINK\]](#)

(On another occasion I'll try to explain why I think most intelligent animals use powerful learning mechanisms not yet available in current AI systems. Some clues are in this document.)

The modal concepts, e.g. 'necessary', 'possible', 'impossible', used to describe such discoveries also have nothing to do with "Possible world semantics" developed in the 20th Century
https://en.wikipedia.org/wiki/Possible_world
<https://plato.stanford.edu/entries/possible-worlds/>

It would be more accurate to say (in agreement with Kant I think) that the ancient mathematical discoveries are concerned with possible and impossible variations of **particular fragments** of **this** universe than alternative **complete** universes, which certainly never entered my mind when I learnt geometry at school.

[Barbara Vetter\(2011\)](#) seems to have developed related criticisms of "possible world semantics", though apparently without any reference to Kant, or mathematics.

(d) Many of those ancient mathematical discoveries are still in constant use by engineers, scientists and mathematicians all over this planet.

(Though in the UK very few students now learn to make such discoveries by solving geometric problems and finding proofs, which I think is disgraceful. As a result, for many learners, mathematical education mainly consists of memorising mathematical facts, with little understanding.)

To summarise: in some species whose members faced problems of interacting with structures and processes in a 3D spatial environment, evolution somehow produced mechanisms implemented in brains that made it possible to reason about possible structures and processes without actually constructing them in the environment. This reasoning used "analogical" as well as "Fregean" (e.g. logical) forms of representation [a distinction explained in [Sloman \(1971\)](#) and Chapter 7 of [Sloman 1978](#)]. In some cases such reasoning revealed interesting constraints on possibilities, i.e. impossibilities and necessary connections, discussed below.

It is not clear to what extent other species have similar forms of mathematical meta-cognition, e.g. allowing them to discover that certain actions that can be thought about cannot ever succeed, e.g. unlinking linked rings without breaking them.

Such recognition of impossibilities and related necessities may have been useful terminators for useless attempts at action, without explicit use of modal concepts, like "can" or "impossible". A more positive example of proto-mathematical reasoning might be a crow's understanding (a) how to make a hook from a straight piece of stiff wire and (b) how to use the hook to lift a bucket of food out of a vertical tube, mentioned [below](#).

I suspect that evolution of abilities to think with and reason about logical operators (e.g. "not", "and", "or", "all", "none") came much later, building on and extending the spatial reasoning mechanisms in only a subset of species with the spatial reasoning abilities.

Although these forms of reasoning could be used in taking real decisions to act on or make predictions about the environment, their use is not dependent on how things are in the environment: they can also be used to reason about hypothetical situations, or possible future situations. But the mechanisms are not infallible (as shown by Imre Lakatos in his [Proofs and Refutations \(1976\)](#)).

Relation to Kant's philosophy of mathematics and Turing's latest ideas

My own work on these (and related problems) began over half a century ago: In [Sloman\(1962\)](#), I tried to expound and defend [Kant's](#) philosophy of mathematics. After learning about AI, starting to program (around 1971), and criticising logicism in AI [Sloman \(1971\)](#), I hoped to develop working explanatory models e.g. explaining mathematical developments in young children based on non-logical mathematical reasoning. But progress has been very slow, and deep gaps in our understanding still go largely unnoticed by most researchers.

Progress was accelerated around 2012 when discussions of Alan Turing's [\(1952\)](#) paper on morphogenesis during his centenary year suggested the idea of the Meta-Morphogenesis (M-M) project (<https://goo.gl/9eN8Ks>), studying evolution of varieties of information processing, since the earliest life forms -- including ways in which products of evolution can alter mechanisms and processes of evolution, partly through use of Derived Construction Kits (DCKs), discussed further [below](#).

Trying to understand intelligence by studying only human intelligence is as misguided as trying to understand life by studying only human life.

So the project investigates evolution of human and non-human information processing, from the simplest forms onwards -- a rich and very varied subset of the space of possible minds [Sloman\(1984\)](#).

"Information" is here used in the sense of [Jane Austen \(1813\)](#) (in her novel *Pride and Prejudice*), not the sense of [Claude Shannon\(1948\)](#). Austen's sense involves semantic content, which can often be useful in controlling decisions or actions. Shannon's measure merely referred to mathematical properties of the syntactic forms used for storing or transmitting information, disregarding what is referred to or expressed, or how it can be of use.

As a novelist and shrewd observer of humanity in her surroundings, Austen was interested in what could be done *with* information content, whereas Shannon's work was mainly concerned with what could be done *to* information vehicles, e.g. storing them, transmitting them, compressing them, correcting distortions, etc. Shannon understood this contrast very well, but his unfortunate choice of terminology -- possibly under pressure from his employer Bell Laboratories(?), who sold information transmission and storage services -- has confused generations of scientists, philosophers and artists, as explained in [Sloman \(2011\)](#).

The most important feature of semantic (Austen) information is that it can be used, e.g. for selecting or controlling action, whether immediately on acquisition of the information (as in thermostats and microbes) or at some later time, possibly in combination with information from different sources. Meta-cognitive mechanisms allow some kinds of information to be used in controlling what is done with other kinds of information. In the last 60 years or so, since the development of computers, and virtual machinery running on computers) our understanding of uses of information has increased enormously, though the majority of scientists and philosophers ignore most of what has been learnt.

Mathematical competences enrich what can be done with information, including deriving new information from old. A theory about the nature of human consciousness, or consciousness in general, must include ancient and recent forms of mathematical consciousness -- partly analysed by Kant ([1781](#)), as summarised [below](#).

Any theory of consciousness that says nothing about mathematical discovery must be incomplete or incorrect, as explained, with examples, [below](#), especially examples of ancient geometric reasoning.

I'll also say a little below about requirements for fundamental physics to support all the known forms of life and the products of biological evolution that made them possible, including fundamental and derived, concrete and abstract, construction kits mentioned [below](#).

An adequate account of fundamental physics must explain how the physical universe can support all the fine details of biological evolution, including evolution of ancient mathematical minds, and their precursors.

The M-M project page provides background information <https://goo.gl/9eN8Ks>.

Can AI reasoning systems replicate the ancient mathematical discoveries?

Since the 1950s there has been work on automated geometrical theorem proving, apparently followed only by a small subset of the AI research community. It seems to have started with a "pencil-and-paper" simulation by Marvin Minsky, which inspired the geometry theorem prover reported in [H. Gelernter, 1964](#). For a brief historical discussion See Margaret Boden's summary in her *Magnum Opus* [Boden \(2006\)](#), section 10.i.b

Impressive more recent work on geometrical theorem proving, which I have not yet studied in detail can be found in [Chou, et al., 1994](#), and other publications by the same authors. However, my understanding is that this work uses reasoning in a logical framework, starting from a variant of Euclid's axioms, supplemented with heuristic use of arithmetical models to block some searches and increase efficiency. As far as I can tell, there is no attempt to model or replicate the processes by which the axioms were originally discovered, or human uses of vision or visual imagination in reasoning that demonstrates [impossibility](#) or [necessity](#).

It is sometimes easy to create a set of axioms describing some portion of the world, e.g. the layout of a building, and to use those axioms to prove that the only route from room A to room B goes past room C. But the fact that that is a "theorem" in that system does not make it true, let alone *necessarily* true. For example the map may be incorrect, missing out a door. Moreover there may be a wall that blocks an alternative route, but walls can be modified by creating new openings through them. So the fact that the map includes the wall does not make it a [necessary truth](#) that there is no route through the location of the wall. What can easily be made false is not a necessary truth. What is a necessary truth is that if a certain wall is in place and is impenetrable then there is no route for a person to move from one side of the wall to the other through the wall, though there may be other longer routes, or routes involving tunnelling, jumping, or flying.

The fact that the impossibility persists while certain constraints are retained can be a (mathematical) necessary truth, unlike a contingent impossibility that is due to somebody not trying hard enough, or not knowing about readily available drilling tools or the lack of a power supply for a drilling tool.

[BACK TO CONTENTS.](#)

Non-local geometrical impossibilities

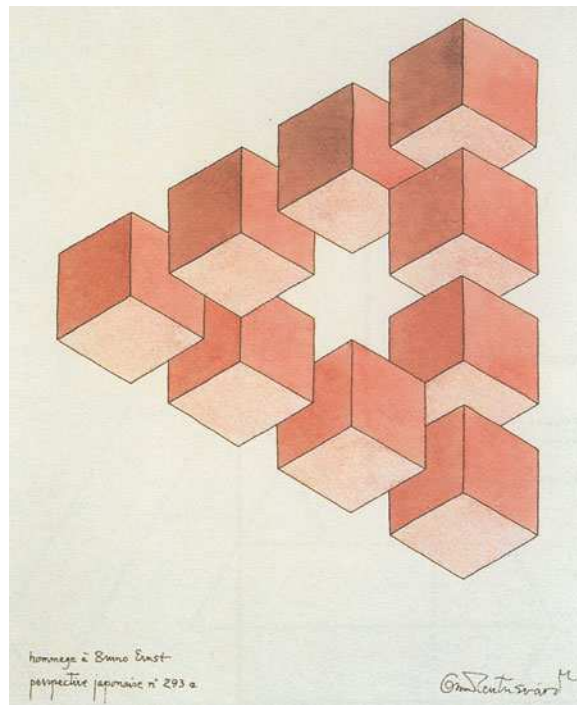
One of the features of spatial cognition that I have been emphasising is ability to detect possibilities and impossibilities (or what Kant referred to as "necessary" spatial relationships -- relationships that cannot be violated).

The first example above ([Figure Stretch-internal](#)) involved perception of directions of continuous change in one feature as another feature changes continuously where all the changes involve containment of one shape by another. In the second example above that is no longer the case: the deformation of the triangle produces overlapping shapes (with a challenge for the reader).

Another kind of mathematical spatial insight involves detection of global impossibility in a static structure with multiple parts, where the parts taken pairwise, or in larger groups are perfectly possible. The impossibility depends on transitivity of relations like [further from the viewer](#) or [above](#) or [further to the right](#), and the impossibility of both "X is further than Y" and "Y is further than X". I.e. the [further than](#) relation is anti-symmetric as well as transitive. How is that discovered?

In the cases shown, possibility/consistency can be restored by removing a subset of the configuration. This is superbly illustrated in [Figure Reutersvärd](#) below. The collection of depictions of configurations of blocks in [Figure Multiblocks](#) below was inspired by Reutersvärd's picture. (The richness of structure discussed here is what makes me prefer the Reutersvärd triangle to the much better known Penrose triangle as an example revealing required features of (adult) human-like cognition.)

Figure Reutersvärd (1934)

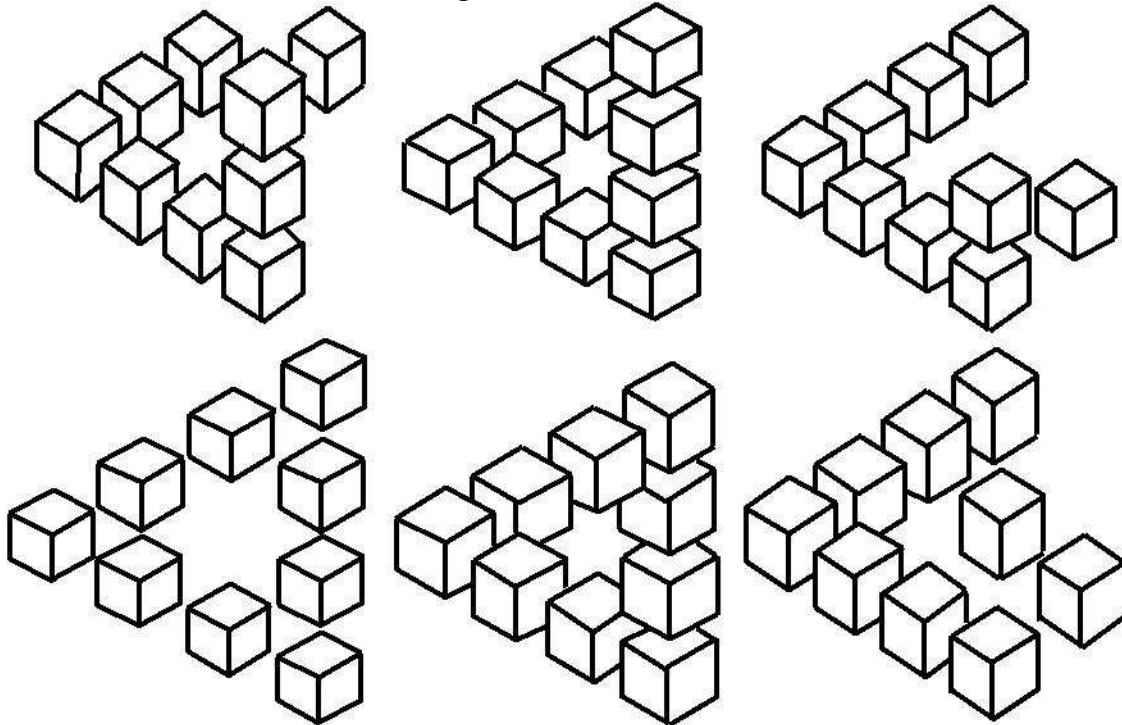


Available at:

http://www.toutfait.com/issues/volume2/issue_4/multimedia/shearer/images/04_big.jpg

The Reutersvärd picture is much richer than the standard impossible triangle because it includes a large collection of affordances (anticipating Escher): For example, if all the blocks are about the size of your fist, consider all the locations in the 3D structure depicted where you could insert your flattened hand between the blocks. Moreover consider all the possible ways in which blocks could exchange their positions in the scene depicted, by being moved continuously through the space. So the picture depicts not only a rich 3-D structure with many relationships between parts, it also implicitly depicts *possibilities for change*, either by moving blocks or inserting a hand or some other object at various locations in the scene. And despite all that, what it depicts is impossible, because it depicts cycles of relationships that are necessarily (why?) transitive and irreflexive. (Details left as an exercise for readers.)

Figure Multiblocks



Some challenges for the reader:

- Which of the 3D configurations of blocks depicted above are spatially possible if interpreted in a "natural" way and which not?
- Why not? What modes of reasoning suffice to demonstrate the impossibility? (Consider transitivity and other features of spatial relations. Of course, the 2D configurations are all possible, as the pictures demonstrate directly.)
- If a group of blocks depicts a geometrically impossible configuration, can possibility be restored by removing any **one** of the blocks in the group?
- What sort of ontology for perceived spatial structures does the impossibility detector require?
- How do the mechanisms required for the earlier examples (reasoning about continuous deformations of a triangle) relate to:
(a) mechanisms required to detect the impossibilities depicted in [Figure Multiblocks](#)?
(b) mechanisms required to detect possible pictorial changes that restore geometrical/physical possibility to the scene depicted?
- Why are distance estimates for individual blocks, or their corners, edges or faces, not needed in order to detect the impossibilities?

The main reason: The examples involve partial orders that are transitive (if A is more than B and B more than C then A is more than C) and antisymmetric (if A is more than B then B cannot be more than A). https://en.wikipedia.org/wiki/Partially_ordered_set. Partial orders without absolute numerical measures suffice for many aspects of visual perception, as illustrated in detail in [Sloman \(2007-2014\)](#). This contradicts many theories about visual perception, and challenges many AI vision projects. Additional examples of non-local

impossibilities are presented in a separate document, including [grid impossibilities](#), [numerical mismatches](#), and impossibilities involving [numerical divisibility](#)

My informal observations suggest that very young children do not see anything wrong with some of the pictures that older humans (in our culture) fairly quickly see as impossible. This may be because the younger children have not yet developed a grasp of non-local angle size comparisons or distance comparisons or because they have not yet grasped the transitivity and antisymmetry of "further", or because they cannot do the required reasoning in some particular cases. All of these require fairly sophisticated information processing mechanism, which I suspect are not yet included in robot visual systems. Neither are they innate in humans, though simpler related competences may be innate in some *precocial* species (e.g. chicks that peck for food soon after hatching).

That suggests that the required "impossibility-detection" mechanisms develop later than other perceptual mechanisms (just as learning to read text develops later than ability to understand spoken or signed language), though I don't know how different such geometrical development would be in a different community, e.g. a forest dwelling, or desert dwelling, or arctic dwelling community that encounters no rectangular objects.

Could the required spatial inferences be easier for members of a deaf community who use only signed rather than spoken language? That would require development of different spatial competences in early childhood. (Compare [Deaf Studies Trust \(1995\)](#)).

The discussion of epigenetic mechanisms [below](#) explains why connections between genetic potential and actual realisation can be very indirect, and partly environment-dependent (also illustrated dramatically by toddlers who are now competent users of internet-connected tablets, which did not exist when their genes were produced by evolution).

Geometrical vs logical reasoning

Unlike the geometrical reasoning discussed here, the example of symbolic reasoning in [Figure Logic](#) above involves consideration only of *discrete* alternatives (the possible truth-values for P, Q, and R, the propositions, and the resulting truth values of premises and conclusion). So it is easy to program a digital computer to examine all the possible cases and discover that the first inference, involving only P and Q is valid and the second is not, because the first premiss would be true if R is true, even if P is false.

On the other hand the kinds of reasoning presented in [Figure Stretch-internal](#), [Figure Stretch-external](#), [Figure Reutersvärd](#) and [Figure Multiblocks](#) all require consideration of non-discrete spaces, in which continuous deformation of shape, size and distance can occur (through infinitely many locations if the lines are infinitely thin, as assumed by Euclid).

Related examples involving changing mostly non-metrical visual affordances are discussed and demonstrated in [Sloman \(2007-2014\)](#). That paper illustrates the need for visual systems to perceive continuous spatial processes resulting either from movements of objects in the environment or changes of viewpoint, and to reason about possible and impossible consequences of such motions. This goes some way beyond James Gibson's ideas about affordances in his [\(1979\)](#), although he indirectly inspired those extensions. As far as I know Gibson did not relate perception of affordances to human mathematical abilities to reason

about topological and geometrical structures, relationships and processes.

The above examples suggest that in order to solve the geometric/topological problems human mathematicians (and in some cases toddlers and other non-mathematicians) can solve, a digital computer would have to be able exhaustively to explore all possible configurations of some spatial structure to ensure that none of them refutes the theorem, unlike the computing systems that can reason about finite discrete spaces in solving problems like those in [Figure Logic](#). Yet human brains cannot go through infinitely many operations while looking at a geometrical proof e.g. using [Figure Stretch-internal](#). So what are brains doing, and can it be replicated on digital computers running suitably designed virtual machinery?

It seems that some ancient animal brains evolved abilities to do the kinds of reasoning I have just summarised, on which humans later constructed abilities to discover mathematical facts about Euclidean geometry. Whether the same abilities can be implemented on a discrete computer is not obvious.

A possible response is to search for a way to implement a new kind of virtual machine able to [represent](#) the required kinds of continuity without itself going through infinitely many states, using abilities somehow to examine and partition infinite collections of possibilities. I don't rule out the possibility that such virtual machines can be implemented on digital computers, though I think it will require new software engineering ideas.

And it may turn out to be very computationally expensive on digital computers -- but not on intra-neuron chemical computers [Grant \(2010\)](#).

It is sometimes thought that computer graphical systems already do such geometrical reasoning, e.g. when they work out where two trajectories will intersect, or what the effects of a complicated collision could be. However, a computer-based graphical simulation system can generate and display a [simulation](#) of a [particular](#) continuously changing configuration, whereas the ability to detect [invariants](#) across large (potentially infinite) classes of such changes requires something very different -- and more difficult to specify. The examples above implicitly specify a partial set of requirements to be satisfied by such a reasoning system.

Replicating human understanding of geometry

Can AI model or replicate the ancient geometrical discovery processes? This is not a question about deriving consequences from Euclid's axioms. The mathematicians who discovered the geometrical facts in Euclid's elements were not selecting an arbitrary set of axioms and then finding out what could be derived from them logically, although it is true that one can explore the consequences of adding or removing various constructions or axioms. For example, Archimedes (?) discovered that adding the "neusis" construction made it possible to trisect any angle, as discussed [below](#).

The [possibility](#) of adding that construction to Euclidean geometry was not a theorem derivable from Euclid's axioms. It was a substantive mathematical discovery extending Euclidean geometry, although some ancient mathematicians did not like it and thought it should not be taught to young mathematicians! (E.g. it was not included in the Euclidean geometry that I learnt at school.)

I believe David Hilbert did not include the neusis construction in his use of logic and arithmetic to model Euclidean geometry. I don't know whether he knew it could be included, but left it out because it was not included by Euclid.

Mary Pardoe's proof of the triangle sum theorem

Another example of mathematical discovery: a former student at Sussex University, Mary Pardoe discovered (around 1970) a related extension to Euclidean geometry that makes it possible to prove the triangle sum theorem (*Internal angles of a triangle sum to half a rotation*) without using Euclid's parallel axiom as the standard proofs do. Her proof, explained here, is very memorable: <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-sum.html>

That was a real mathematical discovery, which I don't believe she ever published, apart from telling me about it. At the time I did not have the good sense to advise her to get it published, e.g. in a journal concerned with mathematical education. I have demonstrated the proof to many audiences, including mathematical audiences. I have never met anyone who had previously encountered the proof, though it is possible that, like many mathematical results, it has been discovered several times, independently, and treated as an unimportant curiosity.

For someone trying to understand the cognitive/brain mechanisms that are capable of generating or comprehending such a proof, it is not a mere curiosity. In particular, there is no obvious answer to the questions:

What brain mechanisms make it possible for such a proof to be discovered?

What brain mechanisms make it possible for such a proof to be understood?

The Pardoe proof is also discussed briefly in

Andrea Asperti, Proof, Message and Certificate,
in *AISC/MKM/Calculamus*, 2012, pp. 17--31,
<http://www.cs.unibo.it/~asperti/PAPERS/proofs.pdf>

And in this slide presentation with the same title, (including a comment on the proof by Dana Scott, who claimed -- unconvincingly in my opinion -- that the proof depends on Euclid's parallel axiom):

<http://www.cs.unibo.it/~asperti/SLIDES/message.pdf>

As Scott pointed out, it is possible to use Euclid's parallel axiom to demonstrate that the Pardoe proof works: but her construction is simple and clear, and makes no mention of parallel lines. Unless I (and several other people) have missed something her proof obviously works without any further justification required. So any claim that the proof somehow depends on Euclid's parallel postulate needs to be substantiated.

On a non-planar surface, such as the surface of a sphere, both the Pardoe proof and other standard proofs of the triangle sum theorem break down, since on a curved surface (e.g. a sphere, an ellipsoid, a torus) interior angles of different triangles can add up to more than or less than a straight line, though that depends on extending the notion of straightness of lines to non-planar surfaces (e.g. the path of minimal length between two points, assuming different lengths are always comparable).

I don't know of any current AI reasoning system that is capable of discovering this sort of extension to Euclidean geometry, or even inventing a new postulate concerning rotation and translation of line segments and then attempting to investigate the consequences of using it in place of the parallel postulate.

I call the result P-geometry, presented in an unfinished document [Sloman \(2010-2017\)](#) started in 2010, but still unfinished. The aim is in part to find out how much of Euclidean geometry can be re-created if the parallel postulate is replaced by the Padoe postulate: *Any line in a plane can be rotated about any point on the line, and successive angles of rotation in the same direction can be added to produce a total rotation.* (Is this provably equivalent to the parallel postulate?). A slight generalisation allows rotation in opposite directions to have different signs, so that rotations can be both added and subtracted to form a total.

More important early discoveries concerning alternatives to Euclidean geometry were used in Einstein's general theory of relativity, which implied that physical space is non-Euclidean, and led to the prediction confirmed by Eddington and colleagues in 1919. Philosophers I met around 1958 thought that that fact had refuted Kant's philosophy of mathematics -- a serious error, as I tried to show in my DPhil thesis [Sloman\(1962\)](#). But my arguments left gaps that I later hoped could be filled using AI to model the mathematical discovery processes of ancient mathematicians -- a goal not yet achieved, for reasons that are not yet clear.

What cognitive mechanisms enable humans to make such discoveries, or to recognize the truth of one of Euclid's axioms, e.g. one of the "axioms of congruence" specifying conditions for two triangles to be congruent? Despite the fact that sets of mathematical axioms are sometimes thought of as arbitrary starting points for a branch of mathematics, they were usually originally major mathematical discoveries. A clear example is the axiom of induction, which is one of Peano's axioms for arithmetic. It is often presented as an axiom that defines what numbers are. However that was not an arbitrary stipulative definition. The axiom of induction, and what can be done with it, were major mathematical [discoveries](#), not mere stipulations.

Kant-inspired notes on mathematical discovery

[VIDEO LINK]

(Expanding parts of [Part 2](#) of the recorded video presentation.)

This work started long before I heard about Artificial Intelligence or learnt to program. After a degree in mathematics and physics at Cape Town in 1956, I came to Oxford in October 1957, intending to do research in mathematics (after further general study). Because I did not like some of the compulsory mathematics courses (e.g. fluid dynamics) I transferred from mathematics to Logic with Hao Wang as my supervisor, and became friendly with philosophy graduate students, with whom I used to argue. This eventually caused me to transfer to Philosophy. I am still trying to answer the questions about mathematical knowledge that drove me at that time.

The philosophers I met (mostly philosophy research students) were mistaken about the nature of mathematical discovery as I had experienced it while actually doing mathematics. E.g. some of them accepted David Hume's categorisation of claims to knowledge, which seemed to me to ignore important aspects of mathematical discovery.

Hume's types of reasoning

1. Hume's first category was "abstract reasoning concerning quantity or number", also expressed as knowledge "discoverable by the mere operation of thought". This was thought to include only "trivial knowledge" consisting only of relations between our ideas, for example, "All bachelors are unmarried". Kant labelled this category of knowledge "Analytic".

It is sometimes specified as knowledge that can be obtained by starting from definitions of words and then using only pure logical reasoning, e.g.

"No bachelor uncle is an only child".

2. Hume's second category was empirical knowledge, i.e. knowledge gained, and tested, by making observations and measurements i.e. "experimental reasoning concerning matter of fact and existence". This would include much common sense knowledge, scientific knowledge, historical knowledge, etc.
3. His third category was everything that could not fit into either the first or second. He described the residue as "nothing but sophistry and illusion" urging that all documents claiming such knowledge should be "committed to flames". I assume he was thinking mainly of metaphysics and theology.

Warning: I am not a Hume expert. For more accurate and more detailed summaries of his ideas search online. e.g.

https://en.wikipedia.org/wiki/David_Hume

<https://plato.stanford.edu/entries/hume/>

The philosophers I met in Oxford in the late 1950s seemed to believe that all mathematical knowledge was in Hume's first category and was therefore essentially trivial. (My memory is a bit vague about 60 year old details.)

But I knew from my own experience of doing mathematics that mathematical knowledge did not fit into any of these categories: it was closest to the first category, but was not trivial, and did not come only from logical deductions from definitions.

I then discovered that, in his [1781 book](#), "*Critique of Pure Reason*", Immanuel Kant had criticised Hume for not allowing a category of knowledge that more accurately characterised mathematical knowledge, namely knowledge that was

- (a) non-analytic, i.e. synthetic,
- (b) non-empirical (*apriori*), and
- (c) about necessary truths and falsehoods (impossibilities).

Kant's characterisation, though not easy to understand, seemed to be much closer than anything else I read at that time to my own experience (as a mathematics student) of learning and doing mathematics, especially geometrical and topological reasoning.

But the philosophers I met then (around 1958-9 and a bit later) thought Kant's ideas about mathematical knowledge being non-trivial and non-empirical were mistaken because he took knowledge of Euclidean geometry as an example. They thought Kant had been proved wrong when Einstein and Eddington showed that space was not Euclidean, by demonstrating the curvature of light rays passing close to the sun:

https://en.wikipedia.org/wiki/Euclidean_geometry#20th_century_and_general_relativity

This argument against Kant was misguided for several reasons. In particular it merely showed that human mathematicians could make mistakes, e.g. by thinking that 2D and 3D spaces were necessarily Euclidean.

One of several formulations of the parallel axiom:

In a Euclidean plane surface, if **P** is any point, and **L** any straight line that does not pass through **P**, there will exactly one straight line through **P** in the plane, that never intersects **L**. I.e. there is a unique line through **P** parallel to **L**.

However, before Einstein's work, mathematicians had previously discovered that not all spaces are necessarily Euclidean and that there were different kinds of space in which the parallel axiom was false (elliptical and hyperbolic spaces). If Kant had known this, I am sure he would have changed the examples that assumed the parallel axiom.

Removing it leaves enough rich and deep mathematical content in Euclidean geometry to illustrate Kant's claims, including the mathematical discovery that a Euclidean geometry without the parallel axiom is consistent with both Euclidean and non-Euclidean spaces. That is as good an example of a non-analytic necessary truth as any Kant presented.

He could have used the discovery that Euclidean geometry without the parallel axiom can be extended in three different ways with very different consequences as one of his examples of a mathematical discovery that is not derivable from definitions by logic, and is a necessary truth, and can be discovered by mathematical thinking, and does not need empirical tests at different locations, altitudes, or on different planets, etc.

While working on my DPhil I also encountered some of Wittgenstein's work on philosophy of mathematics, [Wittgenstein \(RFM\)](#), including his view expressed as: "[For mathematics is after all an anthropological phenomenon](#)", which I thought was completely mistaken (for reasons that have expanded since then!). What I now think is summarised in an incomplete document on "[Multiple Foundations For Mathematics](#)" (work in progress).

In 1962 I completed my DPhil thesis defending Kant, now online [Sloman\(1962\)](#)

I went on to become a lecturer in philosophy, but I was left feeling that my thesis did not answer all the questions, and something more needed to be done. So when Max Clowes, a pioneering AI vision researcher, came to Sussex university and introduced me to AI and programming I was eventually persuaded to try to show how AI could support Kant, by demonstrating how to build a "baby robot" that "grows up" to make new mathematical discoveries in roughly the manner that Kant had described, including replicating some of the discoveries of ancient mathematicians like Archimedes, Euclid and Pythagoras.

Note: Max Clowes

Max Clowes died in 1981. A tribute to him with annotated bibliography is here.
<http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#61>

As Kant had pointed out in 1781, this would require intelligent robots to have a form of learning totally different from both

- methods based on exploring logical consequences of arbitrary collections of axioms and definitions as done in AI theorem provers
- methods based on collecting statistical evidence and performing probabilistic reasoning, including mechanisms that use "deep learning" and forms of probabilistic reasoning.

The latter methods are logically incapable of demonstrating truths of mathematics, which are concerned with necessities and impossibilities, not mere probabilities.

(Including some that human toddlers and intelligent non-human species seem able to discover, even if unwittingly, as I have tried to demonstrate, e.g. in a partial survey of "toddler theorems": [Sloman \(2013c\)](#).)

The 1962 Kant-inspired thesis implied that intelligent robots, like intelligent humans, need forms of mathematical reasoning that are not restricted to use of logical derivations from definitions, and are also different from empirical reasoning based on experiment and observation.

Examples presented above illustrate such reasoning in different ways (e.g. [Figure Reutersvärd](#), [Figure Stretch-internal](#), [Figure Stretch-external](#), [Figure Multiblocks](#)) but there are many more examples, including those discussed in

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/torus.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/rubber-bands.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/changing-affordances.html>

Encouraged by Max Clowes I published a paper (at IJCAI 1971) that challenged the "logician" approach to AI proposed by John McCarthy, one of the founders of AI, as presented in McCarthy and Hayes (1969). My critique, emphasising heuristic benefits of "analogical" representations is [Sloman \(1971\)](#), re-published as [chapter 7](#) of [Sloman 1978](#),

As a result I was invited to spend a year (1972-3) doing research in AI at Edinburgh University. I hoped it would be possible to use AI to defend Kant's philosophical position by showing how to build a "baby robot" without mathematical knowledge, that could grow up to be a mathematician in the same way as human mathematicians did, including, presumably the great ancient mathematicians who knew nothing about modern logic, formal systems of reasoning based on axioms (like Peano's axioms for arithmetic) and did not assume that geometry could be modelled in arithmetic as Descartes had shown.

I published a sort of "manifesto" about this [in 1978](#) (*The Computer Revolution in Philosophy*, now freely available online, with additional notes and comments.) Chapter 8 included some observations on the nature of number competences, which derive from various ways of understanding and using 1-1 correspondences (bijections):

<http://www.cs.bham.ac.uk/research/projects/cogaff/crp/#chap8>

However, the problems proved more difficult than I had anticipated. I still do not know whether some new kind of virtual machine implemented in very powerful computers can model all the biologically based mathematical discovery processes leading up to Euclidean geometry, or

whether some new kind of computing technology is required, perhaps based on models of sub-synaptic chemical information processing of the kinds discussed in [Grant \(2010\)](#). If brains do make significant use of the millions of molecules and molecular interactions within each synapse then digital (e.g. transistor based) computers will not be able to match brains for many decades, or perhaps centuries, to come. At this stage I merely mention that as a possibility. It may be that a similar possibility had occurred to Turing when he was working on his [1952 paper](#).

Leaving open details of mechanism and timescale for now, I'll provide examples of the kinds of mathematical discovery that do not seem to fit within currently known forms of computation, which use logic and arithmetic as the basis of all non-empirical reasoning.

Possibility and depictability

What makes something geometrically impossible is not (as Wittgenstein once suggested in his *Tractatus*) that it cannot be depicted spatially.

There are several old spatial depictions of spatial impossibilities including Hogarth's 1754 engraving "Satire on false perspective" illustrated in https://en.wikipedia.org/wiki/Satire_on_False_Perspective and many well known pictures of impossible objects by Roger Penrose and M.C. Escher, and a not so well known 1934 picture below by Swedish artist, Oskar Reutersvärd.

[BACK TO CONTENTS.](#)

Relevance to (cardinal and ordinal) number competences

Expanded: 12 Sep 2017 (Still in need of revision)

(Not included in the video recording.)

Our analysis of number concepts is very different from those commonly used in discussions of whether number competences are innate, and in common ways of testing for different sorts of number competence in children or other animals.

[References needed]

It is often assumed that number competences are concerned primarily with allocation of numerical labels to groups of objects or to positions in an ordered sequence. Cardinal number competences are demonstrated in answers to "how many Xs are there" or "are there more Xs than Ys", and are often thought to take two different forms.

One form is imprecise allocation of an estimated number using a mechanism that combines estimate of density of items in some part of the visual field with estimation of the area, and forms the product of the two estimates. So as density increases or the area increases the estimated number will increase. Such mechanisms can in some cases easily decide whether there are more items in one group of objects than another, e.g. if the density is the same and one area is much larger, or if the areas are the same and one density is much higher. However, I would expect any such system to be poor at comparing numbers in regions where both the density and the area are different unless both changes are in the same direction, e.g, smaller area and lower density. Estimation should also fail if density varies wildly within a visual region.

A similar approximate mechanism can also be used for heard sounds, using frequency and time duration instead of density and visual area.

Another type of number competence, which yields a precise number label, not an estimate, is based on a pattern-matching process for small numbers ("subitizing"). It is possible that some organisms have evolved brains with many stored patterns usable for rapid size or density comparisons, allowing accurate parallel numerosity detection.

It may even be possible (though cruel) to spend a large amount of time training young children to develop such numerosity classifiers that work rapidly and reliably for larger collections than most humans can label accurately without counting. But that sort of mechanism need not support any understanding of either cardinal or ordinal arithmetic, and it will always have definite size bounds, unlike the use of serial counting mechanisms.

All the precise processes of number identification depend on checking for one-one correspondences either by direct parallel matching between members of two groups, or by rearranging items so that the correspondence is physically perceivable (e.g. both groups lined up in parallel equally spaced rows), or by counting both groups. Use of counting depends on the fact that one-one correspondence is transitive, enabling a sequence of number names to be used as an intermediary between two sets to be compared.

Research on these topics can include tests for innateness, identification of ages at which competences are achieved, cross-cultural or cross-species comparisons, and abilities to do various kinds of reasoning about one-one correspondences. There are also attempts to identify neural mechanisms involved though as far as I know, nobody has a theory about how neural nets can represent general knowledge about one-one correspondences and their uses, including making discoveries such as symmetry and transitivity of the relations.

The use of such number concepts without the ability to use spoken or signed words as number names can be investigated in young children or other animals via indirect tests, e.g. investigating whether an animal can tell that the number of people leaving the area of its nest is smaller than the number previously seen arriving.

The mechanisms mentioned so far do not support full number competences available to ancient mathematicians. It is not always noticed that the concept of cardinal number used and studied in mathematics, science and engineering is very different from the ability to assign a label to a collection of items. The former depends on understanding the second concept of number, based on one-to-one correspondence (*bijection*) between members of two sets. If there is a bijection then that allows one set to be used for various purposes to answer questions about the other set. That can also be relevant to practical problems, e.g. deciding whether enough chairs have been fetched for a dinner party, or enough places set, without requiring all the guests to stand around the table.

If a short object is used to measure the lengths of larger objects by repeatedly laying the former alongside the latter, in adjoining positions, then that enables lengths of two large objects to be compared without bringing them together. (There are many special cases not discussed here.)

The creation of a collection of memorised number names allows comparisons to be made between numbers or lengths of items that are at different locations, which can be useful for many practical purposes. I suggest these competences derive from more fundamental competences

involved in detecting and making use of information about one-to-one correspondences for many practical tasks that initially do not require reference to numbers, but are later discovered to be achievable more simply by using an abstract ordered set of information items (e.g. words) as intermediaries in one-to-one correspondences. Some of the computational mechanisms required are illustrated schematically in [Chapter 8](#) of [Sloman 1978](#),

A full understanding of the uses of such number names requires understanding that one-to-one correspondence is a transitive and symmetric (and therefore reflexive) relation. All of this requires information processing capabilities concerned with use of one-to-one correspondences that seem not to have been noticed by most psychologists and neuroscientists studying number competences, although philosophers and logicians have understood this since Hume at least, and consequences studied by Frege, Russell and many others.

I doubt that anyone understands how human brains represent information about 1-1 correspondences and how they discover mathematical properties of such correspondences, for instance that the relationship must be transitive and symmetric. Piaget's work suggests that the transitivity is not understood by children until the 5th or 6th year [Piaget\(1952\)](#). What sort of brain development is required for that?

A partial analysis of cognitive mechanisms required for a range of practical uses of 1-1 correspondences was presented in Chapter 8 of [Sloman 1978](#), but without any suggestions about neural implementations. I doubt that anyone knows what brain mechanisms allow the transitivity to be discovered and used.

Mature numerical competence requires the ability to understand, in some intuitive way, that one-to-one correspondence is **necessarily** a transitive and symmetric relationship. That capability can then later be the basis of more sophisticated abstract notions of number divorced from their practical applications. The detailed history of such developments, perhaps centuries before Euclid, may be forever inaccessible to us -- requiring creative but disciplined speculation regarding historical details, and their products.

Those pre-historical abstract notions used by early inventors to solve practical tasks, could later have become the subject of discussion, experimentation and investigation, eventually leading to discovery of the existence of prime numbers, the invention of negative numbers, the discovery of ratios (rational numbers), and so on: much of which had already been achieved by the time of Euclid without making any use of the modern logic-based axiomatic method.

The investigation of properties of numbers by experimenting with various operations on them could eventually have become an activity found to be of interest in its own right independently of the practical uses of the discoveries. In this context, "experimentation" does not mean testing in a wide variety of physical or geographically diverse situations, as required in physics, chemistry, biology, psychology, and other sciences. Rather, as Kant realised, the experiments leading to mathematical discovery were primarily **thought experiments**, even though the thinking may have been partly external, like Plato's slave boy (Meno) learning from diagrams drawn in sand. https://en.wikipedia.org/wiki/Meno#Dialogue_with_Meno.27s_slave
<http://www.cs.bham.ac.uk/research/projects/cogaff/crp/#chap7>

Most AI researchers building systems that need to make use of numerical concepts and techniques make use of the fact that modern computers are supplied with many numerical competences ready for use -- depending the ability of collections of electronic binary switches to represent bit patterns and mechanisms for performing binary arithmetic operations, on which an enormous variety of more complex numerical competences can be built. For scientific and philosophical purposes this can have the bad effect that attention is diverted from the question how products of biological evolution implement required mathematical competences. In particular if video cameras produce rectangular arrays of photon-sensors and support arithmetic-based operations for scanning receptor values along straight lines, or circular arcs, etc. this can hide the importance of research on how related low level structures and relationships are identified in messy collections of biological sensors: a question raised in [Craik \(1943\)](#) before digital computers were available. It is possible that the superficially more messy and inaccurate biological mechanisms have deep unnoticed powers that are required for full replication of animal visual competences.

[BACK TO CONTENTS.](#)

What are foundations for mathematics?

I have tried to give a brief, oversimplified, introduction to Kant's ideas about mathematical knowledge, and I have tried to illustrate some of the ways in which the discoveries of ancient mathematicians (and related discoveries made by very young children and non-human animals), that seem to fit Kant's ideas, don't naturally fit in the space of mathematical forms of reasoning and discoveries so far made by AI systems running on computers.

In the last two centuries there has been a lot of research on foundations for mathematics, most of it focused on mathematical foundations for mathematics, i.e. trying to find some subset of mathematics from which all of the rest can be derived. For examples, see the links in [Sakharov \(2003ff\)](#).

A separate document is an attempt to distinguish additional types of foundation for mathematics:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/maths-multiple-foundations.html>
as follows:

- Neo-Kantian (epistemic/cognitive) foundations,
The features of human/animal cognition that make mathematical discoveries possible.
- Proto-cognitive foundations,
evolutionary precursors (e.g. subsets) of the cognitive mechanisms making human mathematics possible
- Mathematical foundations,
A subset of mathematics, from which the rest can be derived, mathematically.
- Meta-foundational foundations for mathematics
A (mathematical? philosophical?) framework for analysing and comparing different proposed mathematical foundations for mathematics. Discussions of this type are often part of discussions about rival proposals for mathematical foundations.

- Biological/evolutionary foundations
Features of biological evolution (and its physical/chemical basis, and its products) that, over billions of years, have implicitly made and used mathematical discoveries, e.g. in control mechanisms, in genetic abstractions enabling powerful control mechanisms (e.g. homeostatic mechanisms) to develop, ... etc., along with genetic forms of representation to "encode" schematic designs for such mechanisms, instantiated during development of individual organisms.
 - Cosmological/physical/chemical foundations
How the physical/chemical universe constrains the kinds of mathematics required for its description and how it makes possible the production, by evolution or engineering, of types of machines (including organisms) with abilities to discover, make use of, and in some cases reason about the mathematical features. Some examples are in [Schrödinger\(1944\)](#)
 - Metaphysical/Ontological foundations
What, in general, explains/constrains the possibility of mathematical truths, and the possibility of their being discovered, proved, and used.
 - Multi-layered foundations
The various layers of support for mathematical structures, mathematical control systems, mathematical reasoning capabilities, in products of evolution and engineering.
For future AI systems that replicate human mathematical capabilities (and possibly more), all the above topics will have to be addressed.
-

[BACK TO CONTENTS.](#)

The Meta-Morphogenesis (Self-informing universe) project

The Turing-inspired Meta-Morphogenesis project was proposed in the final commentary in *Alan Turing - His Work and Impact*, a collection of papers by and about Turing published on the occasion of his centenary [Cooper \(2013\)](#)

The project defines a way of trying to fill gaps in our knowledge concerning evolution of biological information processing that may give clues regarding forms of computation in animal brains that have not yet been re-invented by AI researchers.

This may account for some of the enormous gaps between current AI and animal intelligence, including gaps between mathematical abilities of current AI systems and the abilities of ancient mathematicians whose discoveries are still being used all over world, e.g. Archimedes, Euclid, Pythagoras and Zeno.

Evolution of information processing capabilities and mechanisms is much harder to study than evolution of physical forms and physical behaviours, e.g. because fossil records can provide only very indirect evidence regarding information processing in ancient organisms. Moreover it is very hard to study all the internal details of information processing in current organisms. Some of the reasons will be familiar to programmers who have struggled to develop debugging aids for very complex multi-component AI virtual machines.

Because fossil records of information processing or the mechanisms used are not available, the work has to be highly speculative. But conjectures should be constrained where possible by things that are known. Ideally these conjectures will provoke new research on evolutionary evidence and evidence in living species. However, as often happens in science, the evidence may not be accessible with current tools. Compare research in fundamental physics (e.g. [Tegmark \(2014\)](#)).

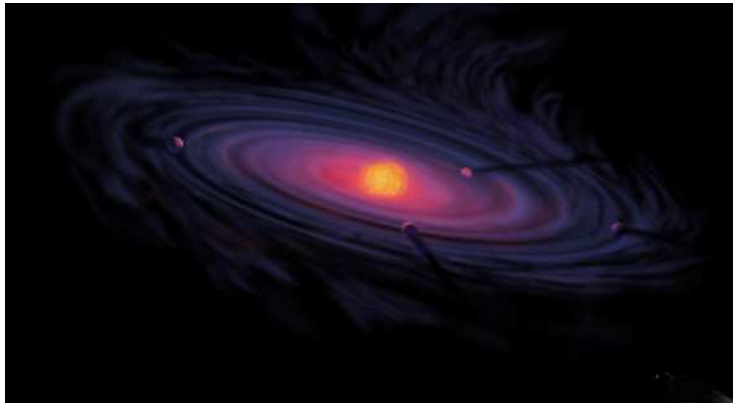
The project presents challenges both for the theory of biological evolution by natural selection, and for AI researchers aiming to replicate natural intelligence, including mathematical intelligence. This is a partial progress report on a long term attempt to meet the challenges. A major portion of the investigation at this stage involves (informed) speculation about evolution of biological information processing, and the mechanisms required for such evolution, including evolved construction-kits. The need for which has not been widely acknowledged by evolutionary theorists. **[VIDEO LINK]** The part of the accompanying 42minute video introducing the project is available here: [here](#).

A lot of work has been done on the project since 2012, some of it summarised below, especially the developing theory of evolved [construction kits](#) of various sorts [Sloman\[2017\]](#), but there are still many unsolved problems, both about the processes of evolution and the products in brains of intelligent animals.

I am not primarily interested in AI as engineering: making useful new machines. Rather I want to understand how animal brains work, especially animals able to make mathematical discoveries like the amazing discoveries reported in Euclid's *Elements* over 2000 years ago.

My interest in AI (which started around 1969) and my work on the The M-M project (since late 2011), arose out of my interest in defending Immanuel Kant's philosophy of mathematics in his [\(1781\)](#), and partly from my conjectured answer to the question: 'What would Alan Turing have worked on if he had not died two years after publication of his 1952 paper on Chemistry and Morphogenesis [\(Turing 1952\)](#). According to Google Scholar, this is now the most cited of his publications. though largely ignored by philosophers, cognitive scientists and AI researchers.

I suspect that if Turing had lived several decades longer, he would have tried to understand forms of information processing needed to control behaviour of increasingly complex organisms produced by evolution, starting from the very simplest forms produced somehow on a lifeless planet produced by condensed gaseous matter and dust particles, later followed, over many millions of years, by increasingly complex organisms, with increasingly complex forms of information processing, including the kinds that led to the ancient mathematical discoveries reported by Euclid, presumably building on earlier discoveries concerning good ways to solve practical problems. [That is the M-M project](#).



[\[NASA artist's impression of a protoplanetary disk, from Wikimedia\]](#)

How could this come about?

I have nothing (at present) to add to conjectures by others about the initial, minimal forms of life, e.g. see, for example, [Ganti \(2003\)](#), and [Froese et al\(2014\)](#).

However, controlled production of complex behaving structures needs increasingly sophisticated information processing:

- in processes of reproduction, growth and development ([Schrödinger \(1944\)](#) had some profound observations regarding mechanisms for storing and using information required for reproduction);
- for control of behaviour of complex organisms reacting to their environment, including other organisms.

In simple organisms, control mainly uses **presence** or **absence** of sensed matter to turn things on or off or sensed **scalar values** to specify and modify other values (e.g. homeostasis and chemotaxis).

As organisms and their internal structures become more complex, the need for **structural** rather than **metrical** specifications increases.

Many artificial control systems are specified using collections of differential equations relating such measures. One of several influential attempts to generalise these ideas is the 'Perceptual Control Theory (PCT)' of [William T Powers](#).

But use of numerical/scalar information is not general enough: It doesn't suffice for linguistic (e.g. grammatical or semantic) structures or for reasoning about topological relationships, or processes of structural change e.g. in building complex nests, in chemical reactions, in programming, or in engineering assembly processes -- or 'toy' engineering, such as playing with meccano sets, tinker toys, Lego, etc. It also cannot describe growth of organisms, such as plants and animals, in which new materials, new substructures, new relationships and new capabilities form -- including new information processing capabilities.

For example, the biologically important changes between an egg and a chicken cannot be described by changes in a state-vector. Why not?

[\(Left as an exercise for the reader: there are several reasons.\)](#)

Turing's Morphogenesis paper ([1952](#)) also focused on mechanisms (e.g. diffusion of chemicals) representable by scalar (numerical) changes, but the results included changes of structure described in words and pictures. As a mathematician, a logician and a pioneer of modern computer science he was well aware that the space of information-using control mechanisms is not restricted to numerical control systems.

For example a Turing machine's operation involves changing linear sequences of distinct structures, not numerical measures, as used in analog computers.

In the last half century human engineers have discovered, designed and built additional increasingly complex and varied forms of control in interacting physical and virtual machines.

That includes control based on

grammars, parsers, planners, reasoners, rule interpreters, problem solvers and many forms of automated discovery and learning.

(Note: It is widely believed, in some academic quarters, that these aspects of symbolic AI have been proved irrelevant to real intelligence. But that belief is an educationally harmful myth.)

Much progress has been made replicating aspects of those competences in computers, especially in the last 50 years, with enormous acceleration during the 21st Century, though the education of AI researchers does not produce much insight into what is still missing from AI: teachers, funding agencies, future employees, and their employers like to focus on successes, i.e. on techniques and theories can be used successfully, however limited they may turn out to be in future decades.

Long before humans began thinking about such matters, and even long before humans existed, biological evolution produced and used increasingly complex and varied forms of information in construction, modification and control of increasingly complex and varied biological processes, including use of information in production of new organisms. Evolution also produced increasingly complex varied mechanisms allowing organisms also to acquire and use information in controlling their behaviours (including internal behaviours such as growth and learning).

[BACK TO CONTENTS.](#)

CONJECTURE What would Turing have done after 1954?

Reflecting on the fact that Turing was still quite young when he died in 1954, two years after publication of his [Morphogenesis paper](#), led me to the conjecture that if had lived several decades longer, he might have produced new theories about many intermediate forms of information in living systems and intermediate mechanisms for information-processing: intermediate between the very simplest forms and the most sophisticated current forms of life.

This would fill gaps in standard versions of the theory of natural selection. E.g., the theory does not explain what makes possible the many forms of life on this planet, and all the mechanisms they use, including the forms that might have evolved in the past or may evolve in the future. Moreover, it does not explain how the physical/chemical mechanisms available when evolution began on this planet it was already possible for evolution to produce mathematical minds, like those of Euclid, Archimedes, etc.

The standard theory of natural selection merely assumes such possibilities are supported by the physical universe and purports to explain how a subset of actually realised possibilities persist, and some of the consequences that follow.

Graham Bell's view of evolution:

For example, the noted biologist Graham Bell wrote "Living complexity cannot be explained except through selection and does not require any other category of explanation whatsoever" [Bell\(2008\)](#). This ignores questions related to the fact that there must have been deep features of the early physical universe that made possible all the many forms of life that have evolved (and many more that have not!), including animals capable of discovering and using Euclidean geometry, long before the tools of modern logic, arithmetic and algebra were available.

Only a few defenders of Darwin, e.g. [Kirschner & Gerhart](#), seem to have noticed the need to explain

- (a) what mechanisms make possible all the options between which choices are made, and
- (b) how what is possible changes, and depends on previously realised possibilities.

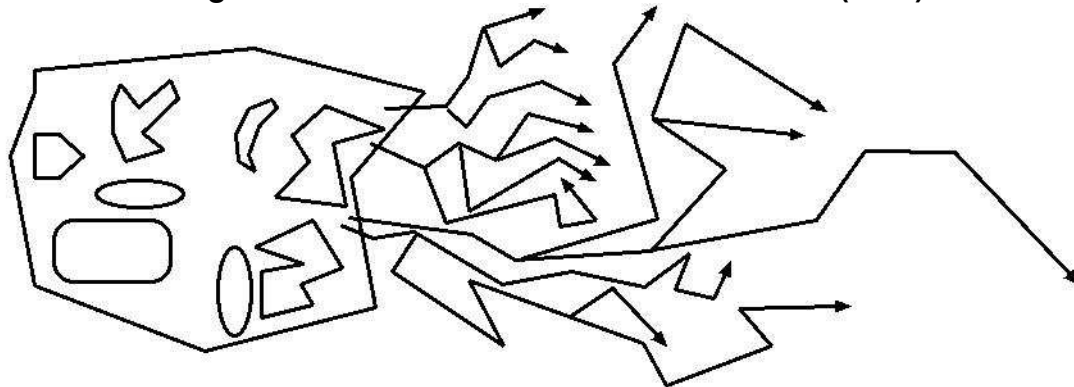
[BACK TO CONTENTS.](#)

CONJECTURE: Evolution's uses of evolved construction kits

A possible defence of Darwinian evolution would enrich it to include investigation of

- (a) the Fundamental Construction Kit (FCK) provided by physics and chemistry before life existed,
- (b) the many and varied 'Derived construction kits' (DCKs) produced by combinations of natural selection and other processes, including asteroid impacts, tides, changing seasons, volcanic eruptions and plate tectonics.

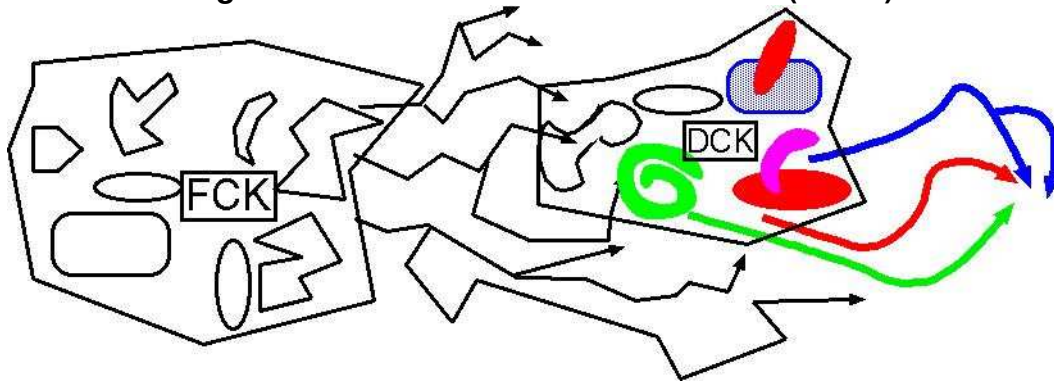
Figure FCK: Fundamental Construction Kit (FCK)



The Fundamental Construction Kit (FCK)

Evolutionary and other trajectories from the FCK through the space of possibilities

Figure DCK: Derived Construction Kits (DCKs)



As new, more complicated, life forms evolved, with increasingly complex bodies, increasingly complex changing needs, increasingly broad behavioural repertoires, and richer branching possible actions and futures to consider, their information processing needs and opportunities also became more complex.

Somehow the available construction kits also diversified, producing new, more complex **derived** construction kits, that allowed

construction not only of new **biological materials and body mechanisms**, supporting new more complex and varied behaviours

but also

construction of new more sophisticated information-processing mechanisms, enabling organisms, either alone or in collaboration, to deal with increasingly complex challenges and opportunities;

including both **concrete** and **abstract** construction kits.

For more on evolution of and use of construction-kits see [Sloman\[2017\]](#) (work in progress).

[VIDEO LINK] The part of the accompanying video recording that discusses construction kits and the resulting epigenetic mechanisms mentioned below is available [here](#).

Deep design discoveries

Many deep discoveries were made by evolution, including designs for DCKs that make possible new forms of information processing.

These have important roles in animal intelligence, including perception, conceptual development, motivation, planning, and problem solving, including

- topological reasoning about properties of geometrical shapes and shape-changes.
- reasoning about possible continuous rearrangements of material objects (much harder than planning moves in a discrete space).

Different species, with different needs, habitats and behaviours, use information about different topological and geometrical relationships, including

- birds that build different sorts of nests,
- carnivores that tear open their prey in order to feed,
- human toddlers playing with (or sucking) body-parts, toys, etc.

Later on, in a smaller subset of species (perhaps only one species?) new meta-cognitive abilities gradually allowed previous discoveries to be noticed, reflected on, communicated, challenged, defended and deployed in new contexts.

Such 'argumentative' interactions may have been important precursors for [chains](#) of reasoning, including the proofs in Euclid's Elements.

Why is this important?

This is part of an attempt to explain how it became possible for evolution to produce mathematical reasoners.

New deep theories, explanations, and working models should emerge from investigation of preconditions, biological and technological consequences, limitations, variations, and supporting mechanisms for biological [construction kits](#) of many kinds.

For example, biologists (e.g. [Coates et al. 2011](#)) have pointed out that specialised construction kits, sometimes called 'toolkits', supporting plant development were produced by evolution, making upright plants possible on land (some of which were later found useful for many purposes by humans, e.g. ship-builders).

Specialised construction kits were also needed by vertebrates and others by various classes of invertebrate forms of life.

[BACK TO CONTENTS.](#)

CONSTRUCTION KITS FOR INFORMATION PROCESSING

Construction kits for biological information processing have received less attention.

One of the early exceptions was Schrödinger's little [1944 book](#), *What is life?* (read by James Watson, before he worked with Crick on DNA).

More general construction kits that are tailorable with extra information for new applications can arise from discoveries of parametrisable sub-spaces in the space of possible mechanisms

e.g. common forms with different sizes, or different ratios of sizes, of body parts, different rates of growth of certain body parts, different shapes or sizes of feeding apparatus, different body coverings, etc.

Using a previously evolved construction kit with new parameters (specified either in the genome, or by some aspect of the environment during development) can produce new variants of organisms in a fraction of the time it would take to evolve that type from the earliest life forms.

Similar advantages have been claimed for the use of so-called Genetic Programming (GP) using evolved, structured, parametrised abstractions that can be re-deployed in different contexts, in contrast with Genetic Algorithms (GAs) that use randomly varied flat strings of bits or other basic units.

Evolution sometimes produces specifications for two or more different designs for different stages of the same organism, e.g. one that feeds for a while, and then produces a cocoon in which materials are transformed into a chemical soup from which a new very different adult form (e.g. butterfly, moth, or dragon fly) emerges, able to travel much greater distances than the larval form to find a mate or lay eggs.

These species use mathematical commonality at a much lower level (common molecular structures) than the structural and functional designs of larva and adult, in contrast with the majority of organisms, which retain a fixed, or gradually changing, structure while they grow after hatching or being born, but not fixed sizes or size-ratios of parts, forces required, etc.

Mathematical discoveries were implicit in evolved designs that support parametrisable variable functionalities, such as evolution's discovery of homeostatic control mechanisms that use negative feedback control, billions of years before the Watt centrifugal governor was used to control speed of steam engines.

Of course, most instances of such designs would no more require awareness of the mathematical principles being used than a Watt-governor, or a fan-tail windmill (with a small wind-driven wheel turning the big wheel to face the wind) does.

In both cases, one part of the mechanism acquires information about something (e.g. whether speed is too high or too low, or the direction of maximum wind strength) while another part does most of the work, e.g. transporting energy obtained from heat or wind power to a new point of application.

Such transitions and decompositions in designs could lead to distinct portions of genetic material concerned with separate control functions, e.g. controlling individual development and controlling adult use of products of development, both encoded in genetic material shared across individuals. There may already be deep relevant work unknown to me already done by the Biosemiotics community referenced [above](#).

[BACK TO CONTENTS.](#)

Metacognition evolves

Very much later, some meta-cognitive products of evolution allowed individuals (humans, or precursors) to attend to their own information-processing (essential for debugging), thereby 'rediscovering' the structures and processes, allowing them to be organised and communicated -- in what we now call mathematical theories, going back to Euclid and his predecessors (about whose achievements there are still many unanswered questions).

If all of this is correct then the physical universe, especially the quantum mechanical aspects of chemistry discussed by Schrödinger provided not only

a construction kit for genetic material implicitly specifying design features of individual organisms,

but also

a 'Fundamental' construction kit (FCK) that can produce a wide variety of 'derived' construction kits (DCKs)

some used in construction of individual organisms, others in construction of new, more complex DCKs, making new types of organism possible.

Moreover, as Schrödinger and others pointed out, construction kits that are essential for micro-organisms developing in one part of the planet can indirectly contribute to construction and maintenance processes in totally different organisms in other locations, via food chains, e.g. because most species cannot synthesise the complex chemicals they need directly from freely available atoms or subatomic materials. So effects of DCKs can be very indirect.

Functional relationships between the smallest life forms and the largest will be composed of many sub-relations.

Such dependency relations apply not only to mechanisms for construction and empowerment of major physical parts of organisms, but also to mechanisms for building information-processors, including brains, nervous systems, and chemical information processors of many sorts. (E.g. digestion uses informed disassembly of complex structures to find valuable parts to be transported and used or stored elsewhere.)

So far, in answer to Graham Bell ([quoted above](#)), I have tried to describe the need for evolutionary selection mechanisms to be supported by enabling mechanisms.

A few others have noticed the problem denied by Bell. E.g. [Kirschner and Gerhart](#) added some important biological details to the theory of evolved construction-kits, though not (as far as I can tell) the ideas (e.g. about abstraction and parametrisation) presented in this paper. Work by Ganti and Kauffman is also relevant -- and probably others unknown to me!

[BACK TO CONTENTS.](#)

Biological uses of abstraction

As organisms grow in size, weight and strength, the forces and torques required at joints and at contact points with other objects change.

So the genome needs to use the same design with changing forces depending on tasks. Special cases include forces needed to move and manipulate the torso, limbs, gaze direction, chewed objects, etc. 'Hard-wiring' of useful evolved control functions with mathematical properties can be avoided by using designs that allow changeable parameters -- a strategy frequently used by human programmers.

Such parametrisation allows both for changes in size and shape of the organism as it develops, and for many accidentally discovered biologically useful abstractions that can be parametrised in such designs -- e.g. allowing the same mechanism to be used for control of muscular forces at different stages of development, with changing weights, sizes, moments of inertia, etc. But the parameters need not be numerical, e.g. a grammatical structure accepting different noun phrases.

Even more spectacular generalisation is achievable by re-use of evolved construction kits ([above](#)).

-- not only across developmental stages of individuals within a species,

-- but also across different species that share underlying physical parametrised design patterns,

-- with details that vary between species sharing the patterns

(as in vertebrates, or the more specialised variations among primates, or among birds, or fish species).

Such shared design patterns across species can result either from species having common ancestry or from convergent evolution 'driven' by common features of the environment,

e.g. re-invention of visual processing mechanisms might be driven by aspects of spatial structures and processes common to all locations on the planet, despite the huge diversity of contents.

Such use of abstraction to achieve powerful re-usable design features across different application domains is familiar to engineers, including computer systems engineers.

'Design sharing' explains why the tree of evolution has many branch points, instead of everything having to evolve from one common root node.

Symbiosis also allows combination of separately evolved features.

Similar 'structure-sharing' often produces enormous reductions in search-spaces in AI systems.

It is also common in mathematics: most proofs build on a previously agreed framework of concepts, formalisms, axioms, rules, and previously proved theorems. They don't all start from some fundamental shared axioms.

If re-usable abstractions can be encoded in suitable formalisms (with different application-specific parameters provided in different design contexts), they can enormously speed up evolution of diverse designs for functioning organisms.

This is partly analogous to the use of memo-functions in software design (i.e. functions that store computed values so that they don't have to be re-computed whenever required, speeding up computations enormously, e.g. in the Fibonacci function).

Another type of re-use occurs in (unfortunately named) 'object-oriented' programming paradigms that use hierarchies of powerful re-usable design abstractions, that can be instantiated differently in different combinations, to meet different sets of constraints in different environments, without requiring each such solution to be coded from scratch: 'parametric polymorphism' with multiple inheritance.

This is an important aspect of many biological mechanisms. For example, there is enormous variation in what information perceptual mechanisms acquire and how the information is processed, encoded, stored, used, and in some cases communicated. But abstract commonalities of function and mechanism (e.g. use of wings) can be combined with species specific constraints (parameters).

Parametric polymorphism makes the concept of consciousness difficult to analyse: there are many variants depending on what sort of thing is conscious, what it is conscious of, what information is acquired, what mechanisms are used, how the information contents are encoded, how they are accessed, how they are used, etc.

[BACK TO CONTENTS.](#)

Mathematical consciousness

Mathematical consciousness, still missing from AI, requires awareness of [possibilities](#) and [impossibilities](#) not restricted to particular objects, places or times -- as Kant pointed out.

Mechanisms and functions with mathematical aspects are also shared across groups of species, such as phototropism in plants, use of two eyes with lenses focused on a retina in many vertebrates, a subset of which evolved mechanisms using binocular disparity for 3-D perception.

That's one of many implicit mathematical discoveries in evolved designs for spatio-temporal perceptual, control and reasoning mechanisms, using the fact that many forms of animal perception and action occur in 3D space plus time, a fact that must have helped to drive evolution of mechanisms for representing and reasoning about 2-D and 3-D structures and processes, as in Euclidean geometry.

In a search for effective designs, enormous advantages come from (explicit or implicit) discovery and use of mathematical abstractions that are applicable across different designs or different instances of one design.

For example a common type of grammar (e.g. a phrase structure grammar) allows many different languages to be implemented including sentence generators and sentence analysers re-using the same program code with different grammatical rules.

Evolution seems to have discovered something like this.

Likewise, a common design framework for flying animals may allow tradeoffs between stability and manoeuvrability to be used to adapt to different environmental opportunities and challenges.

These are mathematical discoveries implicitly used by evolution.

Evolution's ability to use these discoveries depends in part on the continual evolution of new DCKs providing materials, tools, and principles that can be used in solving many design and manufacture problems.

In recently evolved species, individuals e.g. humans and other intelligent animals, are able to replicate some of evolution's mathematical discoveries and make practical use of them in their own intentions, plans and design decisions, far more quickly than natural selection could.

Only (adult) humans seem to be aware of doing this.

Re-usable inherited abstractions allow different collections of members of one species, (e.g. humans living in deserts, in jungles, on mountain ranges, in arctic regions, etc.) to acquire expertise suited to their particular environments in a much shorter time than evolution would have required to produce the same variety of packaged competences 'bottom up'.

This flexibility also allows particular groups to adapt to major changes in a much shorter time than adaptation by natural selection would have required. This requires some later developments in individuals to be delayed until uses of earlier developments have provided enough information about environmental features to influence the ways in which later developments occur, as explained later.

This process is substantially enhanced by evolution of metacognitive information processing mechanisms that allow individuals to reflect on their own processes of perception, learning, reasoning, problem-solving, etc. and (to some extent) modify them to meet new conditions.

Later, more sophisticated products of evolution develop meta-meta-cognitive information processing sub-architectures that enable them to notice their own adaptive processes, and to reflect on and discuss what was going on, and in some cases collaboratively improve the processes,

- e.g. through explicit teaching
- at first in limited social/cultural contexts, after which the activity spread
- using previously evolved learning mechanisms.

As far as I know only humans have achieved that, though some other species apparently have simpler variants.

These conjectures need far more research!

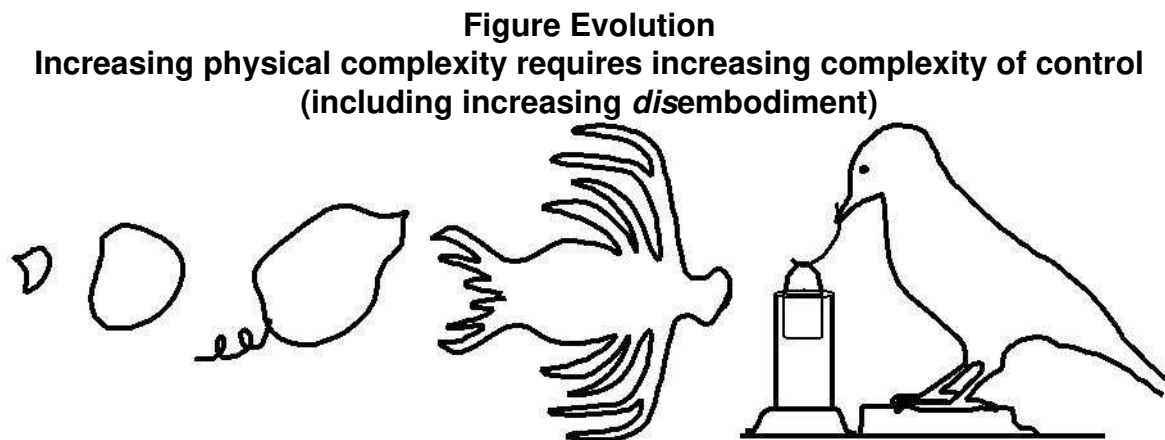
Human AI designs for intelligent machines created so far seem to have far fewer layers of structural/parametrised abstraction, and are far more primitive, than the re-usable designs produced by evolution. Studying the differences is a major sub-task facing the M-M project (and AI).

This requires a deep understanding of what needs to be explained.

[BACK TO CONTENTS.](#)

Designing designs

Just as the designer of a programming language cannot know about, and does not need to know about, all the applications for which the programming language will be used, so also can the more abstract products of evolution be instantiated (e.g. by setting parameters) for use in contexts in which they did not evolve.



A schematic depiction of some of the transitions in biological evolution, including changes in physical form, size, habitat, capabilities and, by implication, information processing abilities required.

Many discontinuities in physical forms, behavioural capabilities, environments, types of information acquired, types of use of information and mechanisms for information-processing are still waiting to be discovered.

A highly intelligent crow

Added 18 Aug 2017; expanded 21 Aug 2017

The bird alluded to in my sketch above, on the right, is Betty a New Caledonian crow, one of

two brought to Oxford University by Dr Jackie Chappell for her research, Betty and Abel (female and male) [Weir et al. \(2002\)](#).

In 2002, Betty made headlines by demonstrating that she could make hooks from a straight piece of wire in order to get a bucket of food out of a vertical glass tube, unlike Abel, who could recognize the potential of and make use of a previously made hook, but never made a new hook. This video showed Betty's performance on her seventh trial with a straight piece of wire.

<https://www.youtube.com/watch?v=OYZnsO2ZgWo>

The video that became famous was one of several on the Oxford Ecology web site, showing Betty making hooks in several different ways, all without any hesitation, despite the fact that the material for the hook, and the laboratory environment, were both very different from the materials and environments available to her species.

[C. Rutz, et al., 2016](#), make somewhat disparaging remarks about Betty's achievements, as if the important questions are about how clever crows are rather than about what they can actually do, what information processing mechanisms, are required to make that possible, and how those mechanisms (or at least the relevant genetic information) were produced by evolution. As far as I know very little is understood about the forms of representation and mechanisms of information processing used by human and non-human animals in such tasks.

The early research reports on Betty did not point out something I discovered by looking at the videos on the Oxford Ecology lab web site, namely that Betty made hooks in at least four very different ways and in every case, seemed to recognize and make use of the strategy without the need for any trial and error. This required her to perceive very complex causal structures in the environment and somehow work out in advance how to use them, perhaps using mechanisms similar to those involved in the toddler's experiments with a pencil, shown in the [Toddler/pencil video](#), despite differences between bird brains and mammal brains.

Moreover, although there is (so far) no reason to think that either a crow or a 17.5 month toddler has the meta-cognitive ability to *notice* that an impossibility or necessary connection has been discovered, that is consistent with it being noticed and used in selecting actions.

Brains develop many capabilities before developing the meta-cognitive mechanisms required to become aware of using those capabilities, e.g. abilities to discern and use grammatical structures in heard utterances.

[BACK TO CONTENTS.](#)

[VIDEO LINK]

Toddler with pencil video

The recorded presentation for this workshop discussed a video presentation of a 17.5 month old (largely pre-verbal) human toddler who was able to grasp the possibility of a complex 3-D trajectory for getting a pencil from pointing through a hole in a sheet of paper in one direction, to entering the same hole from the opposite direction (apologies for sound quality):

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/#infant-pencil>

also discussed here in connection with "Toddler theorems":

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html#pencil>

There are many unknowns about such human and non-human abilities: what information is acquired perceptually and used, how the information is represented in the animal's mind, what prior information was used, what information processing mechanisms are required at various stages in the organism's life, which aspects of the genome contribute to these abilities, how previous experience of individuals interacts with genetic information to provide the mature abilities to perceive, reason about, and make use of unrealised possibilities and constraints on possibilities in the environment.

I suggest that all these aspects of practical intelligence concerned with spatial structures, possibilities, constraints, and actions produced are pre-cursors of the kinds of competence that eventually led to the extraordinary discoveries made by ancient mathematicians. And unlike McCarthy and Hayes (1969) I think future robots replicating such competences will need something different from purely logical (Fregean) forms of representation and reasoning, though the arguments and examples in [Sloman \(1971\)](#) and chapter 7 of [Sloman 1978](#), leave most of the design questions unanswered.

Evolution of human language capabilities

One of the most spectacular cases is reuse of a common collection of language-creation competences in a huge variety of geographical and social contexts, allowing any individual human to acquire any of several thousand enormously varied human languages, including both spoken and signed languages.

Serious errors in wide-spread beliefs about human linguistic competences, how they evolved and how individuals acquire them can be challenged using the famous case of deaf children in Nicaragua, [referenced below](#).

The deaf children spontaneously, and cooperatively, created a new sign language because their teachers had not learned sign languages early enough to develop full adult competences. So the children who started learning in collaboration with the teachers went on without the help of the (mystified) teachers, and created a rich new sign language. This suggests that what is normally regarded as language learning is really cooperative language creation, demonstrated in this video:

<https://www.youtube.com/watch?v=pjtioIFuNf8>

Re-use can take different forms, including

- re-use of a general design across different species by instantiating a common pattern,
- re-use based on powerful mechanisms for acquiring and using information about the available resources, opportunities and challenges during the development of each individual.
- "recursive"(?) reuse: using a general mechanism to get information at one stage in development to provide "parameters" or other structured information required during expression of more abstract genetic information later on -- e.g. providing parameters that influence gene expression (not to be confused with acquisition of data in an already developed learning mechanism). The most striking evidence for this comes from language development, but there are other cases, some documented in [Karmiloff-Smith\(1992\)](#). I think mathematical development is full of unrecognized examples (including "toddler theorems", [Sloman \(2013c\)](#).)

One of the implications of this is that since individuals cannot have any conception of the "rewards" that will come much later on from what they do, they need to have forms of motivation that are **not** reward-based. I call that "architecture-based motivation" (ABM vs RBM) [Sloman\(2009\)](#).

The first process happens across evolutionary lineages.

The second happens within individual organisms in their lifetime

Social/cultural evolution requires intermediate timescales.

Evolution seems to have produced multi-level design patterns, whose details are filled in incrementally, during creation of instances of the patterns in individual members of a species.

If all the members live in similar environments that will tend to produce uniform end results.

However, if the genome is sufficiently abstract, then environments and genomic structures may interact in more complex ways, allowing small variations during development of individuals to cascade into significant differences in the adult organism, as if natural selection had been sped up enormously.

A special case is evolution of an immune system with the ability to develop different immune responses depending on the antigens encountered. Another dramatic special case is the recent dramatic cascade of social, economic, and educational changes supported jointly by the human genome and the internet!

[BACK TO CONTENTS](#).

Changes in developmental trajectories

As living things become more complex, increasingly varied types of information are required for increasingly varied uses.

The processes of reproduction normally produce new individuals that have seriously under-developed physical structures and behavioural competences.

Self-development requires physical materials, but it also requires information about what to do with the materials, including disassembling and reassembling chemical structures at a sub-microscopic level and using the products to assemble larger body parts, while constantly providing new materials, removing waste products and consuming energy.

Some energy is stored and some is used in assembly and other processes.

The earliest (simplest?) organisms can acquire and use information about (i.e. sense) only internal states and processes and the immediate external environment, e.g. pressure, temperature, direction of gravity, presence of chemicals in the surrounding soup, and perhaps arrival of photons, with all uses of information taking the form of immediate local reactions, e.g. allowing a molecule through a membrane.

Changes in types of information, types of use of information and types of biological mechanism for processing information have repeatedly altered the processes of evolutionary morphogenesis that produce such changes: a positive feedback process.

An example is the influence of mate selection on evolution in intelligent organisms: mate selection is itself dependent on previous evolution of cognitive mechanisms. Hence the prefix 'Meta-' in 'Meta-Morphogenesis'.

This is a process with multiple feedback loops between new designs and new requirements (niches), as suggested in [Figure EPI](#) below and [Sloman\[2017\]](#), enlarging on themes in [Sloman\(2000\)](#).

[BACK TO CONTENTS.](#)

Online vs offline intelligence

As the previous figure suggests, evolution constantly produces new organisms that may or may not be larger than predecessors, but are more complex both in the types of physical action they can produce and also the types of information and types of information processing required for selection and control of such actions.

Some of that information is used immediately and discarded (online perceptual intelligence) while other kinds are stored, possibly in transformed formats, and used later, possibly on many occasions (offline perceptual intelligence) -- a distinction often mislabelled as 'where' vs 'what' perception.

This generalises Gibson's theory that perception mainly provides information about 'affordances' rather than information about visible surfaces of perceived objects.

These ideas, like Karmiloff-Smith's Beyond Modularity suggest that one of the effects of biological evolution was fairly recent production of more or less abstract construction kits that come into play at different stages in development, producing new more rapid changes in variety and complexity of information processing across generations as explained below (See fig 2)

It's not clear how much longer this can continue: perhaps limitations of human brains constrain this process. But humans working with intelligent machines may be able to stretch the limits.

At some much later date, probably in another century, we may be able to make machines that do it all themselves -- unless it turns out that the fundamental information processing mechanisms in brains cannot be modelled in computer technology developed by humans.

Species can differ in the variety of types of sensory information they can acquire, in the variety of uses to which they put that information, in the variety of types of physical actions they can produce, in the extent to which they can combine perceptual and action processes to achieve novel purposes or solve novel problems, and the extent to which they can educate, reason about, collaborate with, compete against conspecifics, and prey or competitor species.

As competences become more varied and complex, the more disembodied must the information processing be, i.e. disconnected from current sensory and motor signals (while preserving low level reflexes and sensory-motor control loops for special cases).

This may have been a precursor to mathematical abilities to think about transfinite set theory and high dimensional vector spaces or complex modern scientific theories.

E.g. Darwin's own thinking about ancient evolutionary processes. was detached from his particular sensory-motor processes at the time! This applies also to affective states, e.g. compare being startled and being obsessed with ambition.

The fashionable emphasis on "embodied cognition" may be appropriate to the study of organisms such as plants and microbes, and perhaps insects, but evolved intelligence increasingly used disembodied cognition, most strikingly in the production of ancient mathematical minds. This led to new complexities in processes of epigenesis (gene-influenced development).

[BACK TO CONTENTS.](#)

Epigenesis in organisms and in species

Epigenetics is the study of processes and mechanisms by which genes influence the development of individual members of a species.

Figure WAD: Waddington's view of epigenesis: The Epigenetic Landscape

A ball rolling (passively) down a fixed landscape

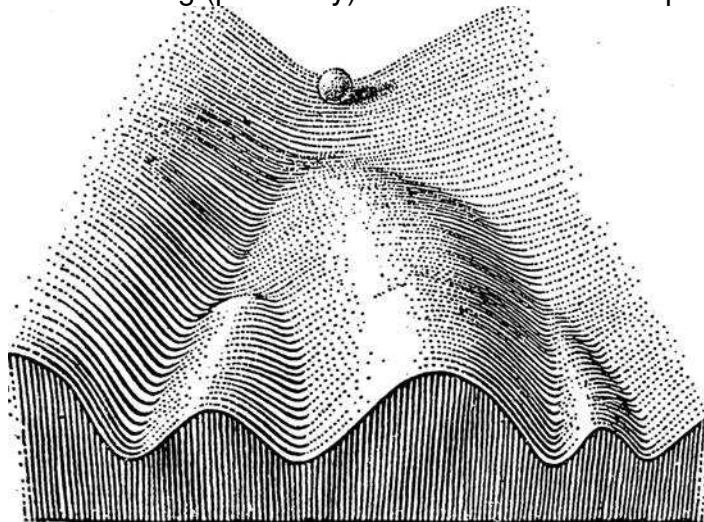
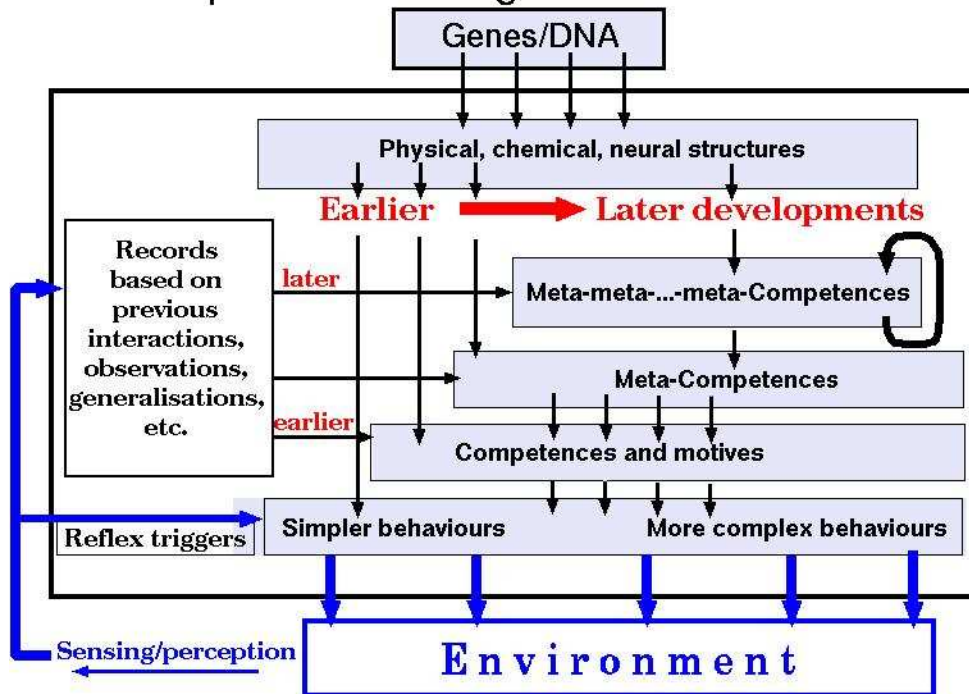


Figure EPI: The meta-configured genome
A more complex theory of epigenesis (beyond Waddington)
Multiple routes from genome to behaviours



Cascaded, staggered, developmental trajectories, with later processes influenced by results of earlier processes in increasingly complex ways. Proposed by Chappell and Sloman (2007)

In [Figure EPI](#), early genome-driven learning from the environment occurs in loops on the left. Downward arrows further right represent later gene-triggered processes during individual development instantiating generic patterns on the basis of results of earlier learning via feedback on left e.g. syntactic structures found, in the case of language development. Chris Miall helped with the original diagram in the Journal paper. Alan Bundy suggested the feedback loop extending meta-...competences: indicating that individual learning can expand the scope of genetic influences on acquired structures during epigenesis.

A possible name for the Chappell-Sloman theory is "The Meta-Configured Genome" (producing a meta-configured epigenetic landscape?). Could a similar diagram represent physical/chemical evolution of the universe? Genes would have to be replaced by hitherto unknown aspects of fundamental physics.

[BACK TO CONTENTS.](#)

Variations in epigenetic trajectories

The description given so far is very abstract and allows significantly different instantiations in different species, addressing different sorts of functionality and different types of design, e.g. of physical forms, behaviours, control mechanisms, reproductive mechanisms, etc.

At one extreme the reproductive process produces individuals whose genome exercises a fixed pattern of control during development, leading to 'adults' with only minor variations.

At another extreme, instead of the process of development from one stage to another being fixed in the genome, it could be [created during development](#) through the use of more than one level of design in the genome.

E.g. if there are two or more levels then results of environmental interaction at earlier levels could determine how generic potential is later instantiated at higher levels -- e.g. what sorts of verbal morphology, grammatical structure or semantic structure develop at later stages. If there are multiple levels then what happens at each new level may be influenced partly by results of earlier developments.

If the abstract structures are not mere numeric formulae whose instantiation involves insertion of numeric values, but also allow previously acquired structures (e.g. spatial structures, process structures, grammatical structures) to be inserted at later stages of development, then the results can differ far more than if numerical values are used. The same is true if different chemical structures are inserted into the same chemical framework. This may have been one of the insights driving [Schrödinger\(1944\)](#).

In a species with such multi-stage development, at intermediate stages not only are there different developmental trajectories due to different environmental influences, there are also selections among the intermediate level patterns to be instantiated, so that in one environment development may include much learning concerned with protection from freezing, whereas in other environments individual species may vary more in the ways they seek water during dry seasons.

As the development of linguistic competence shows clearly these processes can produce new construction kits within individuals, including construction kits for creating new (to the learner) forms of representation and reasoning. The theory of "architecture-based motivation" [Sloman\(2009\)](#) implies that similar structural variation can occur in patterns of motivation.

Then differences in adults come partly from the influence of the environment in selecting patterns to instantiate. E.g. one group may learn and pass on information about where the main water holes are, and in another group individuals may learn and pass on information about which plants are good sources of water. Further details may be results of different individual experiences, or complex interactions with products of previously instantiated genetic patterns.

If these conjectures are correct, patterns of individual development in a species will automatically be varied because of patterns and meta-patterns picked up by earlier generations and instantiated in cascades during individual development.

So different cultures produced jointly by a genome and previous environments can produce very different expressions of the same genome, even though individuals share similar physical forms.

The main differences are in the kinds of information acquired and used, and the information processing mechanisms developed. Not all cultures use advanced mathematics in designing buildings, but all build on previously evolved understanding of space, time and motion.

All this implies that evolution has found how to provide rich developmental variation by allowing information gathered by young individuals not merely to select and use pre-stored design patterns, but to create new patterns by assembling fragments of information during earlier development, then using more abstract processes to construct new abstract patterns, partly

shaped by the current environment, but with the power to be used in new environments.

Developments in culture (including language, science, engineering, mathematics, music, literature, etc.) all show such combinations of data collection and enormous creativity, including creative ontology extension (e.g. the Nicaraguan children mentioned above).

Unless I have misunderstood her, this is the type of process Annette Karmiloff-Smith called 'Representational Re-description' (RR).

Genome-encoded previously acquired abstractions 'wait' to be instantiated at different stages of development, using cascading alternations between data-collection and abstraction formation (RR) by instantiating higher level generative abstractions (e.g. meta-grammars), not by forming statistical generalisations.

This could account for both the great diversity of human languages and cultures, and the power of each one, all supported by a common genome operating in very different environments.

Jackie Chappell noticed the implication that instead of the genome specifying a fixed 'epigenetic landscape' (as proposed at one stage by [Waddington](#)) it provides a [schematic](#) landscape and mechanisms that allow each individual (or in some cases groups of individuals) to modify the landscape while moving down it (e.g. adding new hills, valleys, channels, barriers, taps that can be turned on or off, etc.).

Though most visible in language development, the process is not unique to language development, but occurs throughout childhood (and beyond) in connection with many aspects of development of information processing abilities, construction of new ontologies, theory formation, etc.

This differs from forms of learning or development that use uniform statistics-based methods for repeatedly finding patterns at different levels of abstraction.

Instead, [Figure EPI](#) indicates that the genome encodes increasingly abstract and powerful creative mechanisms developed at different stages of evolution, that are 'awakened' (a notion used by Kant) in individuals only when appropriate, so that they can build on what has already been learned or created in a manner that is tailored to the current environment.

For example, in young (non-deaf) humans, processes giving sound sequences a syntactic interpretation develop after the child has learnt to produce and to distinguish some of the actual speech sounds used in that location.

In social species, the later stages of [Figure EPI](#) include mechanisms for discovering non-linguistic ontologies and facts that older members of the community have acquired, and incorporating relevant subsets in combination with new individually acquired information.

Instead of merely absorbing the details of what older members have learnt, the young can absorb forms of creative learning, reasoning and representation that older members have found useful and apply them in new environments to produce new results.

In humans, this has produced spectacular effects, especially in the last few decades.

The evolved mechanisms for representing and reasoning about possibilities, impossibilities and necessities were essential for both perception and use of affordances and for making mathematical discoveries, something statistical learning cannot achieve.

[BACK TO CONTENTS.](#)

Evolution as epigenesis

Added 18 Aug 2017

The theory of evolved construction kits mentioned above (see [Sloman\[2017\]](#)), can be understood as claiming that something loosely similar to the above model of epigenesis applies also to the processes of biological evolution, except that the initial state is not a product of biological evolution. The physical state of a newly formed earth-like planet, like a genome, has enormous potential for supporting "epigenetic" processes that eventually lead to billions of different forms of life.

Evolution and use of derived construction kits (DCKs), both concrete and abstract (as required for many forms of information processing), uses a process of multi-layer feedback partly like the Chappell-Sloman conjectures depicted in [Figure EPI](#).

The repeated use by evolution of increasingly complex and abstract designs, shows that the theory of evolution by design is basically correct except that there is no single initial design: new designs are repeatedly created and put to use. That seems to include designs used in ancient mathematical minds, still available in the human genome, that have so far eluded AI researchers, and everyone else.

Space-time

An invariant for all species in this universe is space-time embedding, and changing spatial relationships between body parts and things in the environment.

The relationships vary between water-dwellers, cave-dwellers, tree-dwellers, flying animals, and modern city-dwellers.

Representational requirements depend on body parts and their controllable relationships to one another and other objects.

So aeons of evolution will produce neither a tabula rasa nor geographically specific spatial information, but a collection of generic mechanisms for finding out what sorts of spatial structures have been bequeathed by ancestors as well as physics and geography, and learning to make use of whatever is available [McCarthy \(2008\)](#): that's the main reason why embodiment is relevant to evolved cognition. However, as organisms grow more complex, with more complex behavioural alternatives and more options to choose between, requiring actions extended over space and time, cognitive processing must become increasingly disembodied, i.e. disconnected from sensory and motor subsystems. For example, if some plan unexpectedly goes badly wrong, working out how that happened may require reflecting in detail on the processes that occurred and the un-tried alternatives. The results of that process of analysis may prove life-saving at some future time.

Enactivist, anti-cognitivist, anti-computational theories of cognition are rightly given short shrift in [Rescorla \(2015\)](#). (Most of the achievements of human mathematics, science and engineering, and even philosophy, would have been impossible if enactivist/embodied theories of cognition were close to the truth. Developing an enactivist/disembodied theory of cognition is itself a mostly disembodied process!)

Kant's ideas about geometric knowledge are relevant though he assumed that the innate apparatus was geared only to structures in Euclidean space, whereas our space is only approximately Euclidean.

Somehow the mechanisms conjectured in Figure 2 eventually (after many generations) made it possible for humans to make the amazing discoveries recorded in Euclid's Elements, still used world-wide by scientists and engineers.

If the parallel axiom is removed what remains is still a very rich collection of facts about space and time, especially topological facts about varieties of structural change, e.g. formation of networks of relationships, deformations of surfaces, and possible trajectories constrained by fixed obstacles.

If we can identify a type of construction-kit that produces young robot minds able to develop or evaluate those ideas in varied spatial environments, we may find important clues about what is missing in current AI.

Long before logical and algebraic notations were used in mathematical proofs, evolution had produced abilities to represent and reason about what Gibson called 'affordances', including types of affordance that as far as I know he did not consider, namely reasoning about possible and impossible alterations to spatial configurations and necessary consequences of possible alterations. How brains represent possibilities, impossibilities, and necessary consequences (as opposed to learnt associations) is, as far as I know, still open.

Example:

The (topological) impossibility of solid linked rings becoming unlinked is usually obvious even to children without any mathematical education. E.g.

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/rings.html>

This rubber-band example is harder to understand:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/rubber-bands.html>

I suspect brains of many intelligent animals make use of topological reasoning mechanisms that have so far not been discovered by brain scientists or AI researchers. Weaver birds are among the most obvious candidates.

Addition of meta-cognitive mechanisms able to inspect and experiment with reasoning processes may have led both to enhanced spatial intelligence and meta-cognition, and also to meta-metacognitive reasoning about other intelligent individuals.

[BACK TO CONTENTS.](#)

Other species

I suspect that further investigation will reveal varieties of information processing (computation) that have so far escaped the attention of researchers, but which play important roles in many intelligent species, including not only humans and apes but also elephants, corvids, squirrels, cetaceans and others.

In particular, some intelligent non-human animals and pre-verbal human toddlers seem to be able to use mathematical structures and relationships (e.g. partial orderings and topological relationships) unwittingly. Mathematical meta-meta...-cognition seems to be restricted to humans, but develops in stages, as Piaget found ([1952](#)). E.g. it seems that recognition of transitivity of one-one correspondence is not achieved until about the 5th or 6th year. (What would have to change in brains to make that occur?)

However, I suspect that (as Kant seems to have realised) the genetically provided mathematical powers of intelligent animals make more use of topological and semi-metrical geometric reasoning, using analogical, non-Fregean, representations, than the logical, algebraic, and statistical capabilities that have so far dominated AI and robotics. ([Sloman \(1971\)](#) and chapter 7 of [Sloman 1978](#))

For example, even the concepts of cardinal and ordinal number are crucially related to concepts of one-one correspondence between components of structures, most naturally understood as a topological relationship rather than a logically definable relationship.

<http://www.cs.bham.ac.uk/research/projects/cogaff/crp/#chap8.html>

(NB 'analogical' does not imply 'isomorphic' as often suggested. A typical 2D picture (an analogical representation) of a 3D scene cannot be isomorphic with the scene depicted. A 3D to 2D projection is not an isomorphism except in special cases. There is a deeper distinction between [Fregean](#) and [Analogical](#) forms of representation [Sloman \(1971\)](#), concerned with the relationships between representation and what is represented.

[BACK TO CONTENTS.](#)

Disembodiment of cognition evolves (Epigenesis of evolutionary mechanisms)

All this shows why increasing complexity of physical structures and capabilities, providing richer collections of alternatives and more complex internal and external action-selection criteria, requires increasing disembodiment of information processing -- as the information required for particular episodes of control, motive generation, or planning becomes more abstract, more concerned with remote locations more concerned with past events or future possibilities, less concerned with reference to physical objects, actions and other processes, and more concerned with things like past percepts, memories, possible future decisions, modes of reasoning, and mental states of other individuals (what they perceive, want, intend, know, can do, etc.)

Such transitions occur both in individual development in intelligent species and also in evolution of complex organisms.

The fact that evolution is not stuck with the Fundamental Construction Kit (FCK) provided by physics and chemistry, but also produces and uses new 'derived' construction-kits (DCKs), including abstract construction kits needed for intelligent organisms (e.g. grammar construction kits in humans), enhances both the mathematical and the ontological creativity of evolution, which is indirectly responsible for all the other known types of creativity.

Although I have not developed the idea in this paper, the work on construction kits and their **essential** role in evolution on this planet, suggests that there are weak but important analogies between epigenetic processes in individual humans illustrated in [Figure EPI](#), and some evolutionary processes. In both cases, the development depends on discovery of powerful abstractions ("moving upwards") that can be instantiated in different ways in different species or the same species at different times ("moving downwards"), instead of all evolution being simply "sideways" movement at a fixed level of abstraction in design.

(This distinction is ignored by theories of mathematical discovery by humans that emphasise use of [metaphor](#) and [analogy](#) instead of use of [abstraction](#) and multiple re-instantiation.)

The fact that many evolved construction kits, and their products, depended on natural selection "discovering" enormously powerful re-usable mathematical abstractions, whose re-use involved not just copying, but instantiation of a generic schema with new parameters, illustrates a partial analogy between epigenesis in intelligent organisms and epigenesis in evolution. To that extent I am proposing that evolution needs intelligent design, but all the intelligence used in the design processes was previously produced by evolution.

This counters both the view that mathematics is a product of human minds, and a view of metaphysics as being concerned with something unchangeable.

The notion of 'Descriptive Metaphysics' presented by Strawson (1959) needs to be revised, to include '[Meta-Descriptive Metaphysics](#)'.

[BACK TO CONTENTS.](#)

Are non-Turing forms of computation needed?

I also conjecture that filling in some of the missing details in this theory (a huge challenge) will help us understand both the evolutionary changes that introduced unique features of human minds and why it is not obvious that Turing-equivalent digital computers, or even asynchronous networks of such computers running sophisticated interacting virtual machines, will suffice to replicate the human mathematical capabilities that preceded modern logic, algebra, set-theory, and theory of computation.

In fact, I once argued ([Sloman \[IRREL\]](#)) that Turing machines as originally specified by Alan Turing, are irrelevant to AI, and that most of the ideas about computation construed as a form of information processing that have informed AI did not *depend* on the theory of turing machines or the existence of physical Turing machines, although the theory can help with discussion of a subset of properties of AI programs and problems, such as their efficiency, or predictability. I suspect Turing would have agreed with this, given the direction his thought seemed to be taking in 1952.

In particular, for a machine (or animal) M consisting of a Turing machine informationally coupled to an extended physical environment (e.g. via a sensor-driven device that alters parts of the tape, or even the 'machine table' (set of rules) of the machine) the behaviour of the system M will not be subject to the mathematical limits of turing machines. (Turing himself explored the idea of a TM connected to an "oracle".)

Although isolated Turing machines form a theoretically important special subset of the set of possible information processing mechanism, the general notion of information processing or computation is not restricted to computation using a Turing machine, nor computation using bit patterns or numbers, nor computation based on running symbolic programs of the sorts that run on our computers.

Those are all special cases of the more general notion of a mechanism that can produce and store enduring and discretely or continuously changeable structures, with mechanisms that can operate on those structures by creating them, testing their forms, modifying them, producing new ones, and basing control decisions on them, or, in other words using them as instructions (but with no instructor presupposed). This merely means that the machine uses the structures as specifications of types of action to be performed -- either on such structures themselves, or on other things, including parts of organisms, parts of the environment, or contents of distributed virtual machines (whose practical importance has only been understood in detail by human engineers in the last 80 years or so, although evolution made that discovery much earlier, and probably uses sophisticated mechanisms that human engineers have not yet thought of).

The scope and limits of information processing competences of life forms on this planet will depend on the precise forms of **virtual** information processing machinery that evolution has managed to produce, for use in perception, reasoning, learning, motive formulation, planning, metacognition, other-related metacognition, and generalising use of affordances to include thinking about possibilities and impossibilities. I suspect current methods of neuroscientific investigation cannot yield deep information about these problems, especially if we have underestimated the importance chemical information-processing in brains [Grant \(2010\)](#).

Moreover, it seems that current AI mechanisms (including deep learning mechanisms -- which lack the ability to discover impossibilities and necessary connections, among other things) cannot produce reasoners like Euclid, Zeno, Archimedes, or even reasoners like pre-verbal toddlers, weaver birds and squirrels. Some of the reasons have been indicated [above](#).

In particular, the examples of mathematical discovery suggest that there are serious gaps in current AI, despite many impressive achievements. I see no reason to believe that uniform, statistics-based learning mechanisms will have the power to bridge those gaps: in particular it is impossible for statistical reasoning based on empirical evidence to establish **necessary** truths (as Kant argued against Hume).

What about logic?

Whether increased sophistication of logic-based reasoners will suffice (as suggested by McCarthy and Hayes)[\(1969\)](#) is not clear.

The discoveries made by ancient mathematicians preceded the discoveries of modern algebra and logic, and the arithmetisation of geometry by Descartes. So they were definitely not consciously using **only** logical/algebraic reasoning

Evolved mechanisms that use previously acquired abstract forms of meta-learning with genetically orchestrated instantiation triggered by developmental changes (as suggested earlier in explaining [Figure EPI](#)), may do much better.

Those mechanisms depend on rich internal languages that evolved for use in perception, reasoning, learning, intention formation, plan formation and control of actions before communicative languages.

This generalises claims made in [Chomsky \(1965\)](#), and his later works, which focused only on development of human spoken languages, ignoring the extent to which language and non-linguistic cognition evolve in a species, and develop in individuals, with mutual support. For more on the importance of use of rich internal languages for perception, intentions, reasoning, planning, and control of actions, before evolution of communicative languages see [this presentation](#). David Mumford (mathematician) is one of the few people who seem to agree with this [Mumford\(2016\)](#).

[BACK TO CONTENTS.](#)

The importance of virtual machinery

Building a new computer for every task was made unnecessary by allowing computers to have changeable programs.

Initially each program, specifying instructions to be run, had to be loaded (via modified wiring, switch settings, punched cards, or punched tape), but later developments provided more and more flexibility and generality, with higher level programming languages providing reusable domain specific languages and tools, some translated to machine code, others run on a task specific virtual computer provided by an interpreter.

Later developments provided time-sharing operating systems supporting multiple interacting programs running effectively in parallel performing different, interacting, tasks on a single processor.

As networks developed, these collaborating virtual machines became more numerous, more varied, more geographically distributed, and more sophisticated in their functionality, often extended with sensors of different kinds and attached devices for manipulation, carrying, moving, and communicating.

These developments suggest the possibility that each biological mind is also implemented as a collection of concurrently active nonphysical, but physically implemented, virtual machines interacting with one another and with the physical environment through sensor and motor interfaces.

Such 'virtual machine functionalism' could accommodate a large variety of coexisting, interacting, cognitive, motivational and emotional states, including essentially private qualia as explained by Sloman and Chrisley ([2003](#)).

Long before human engineers produced such designs, biological evolution had already encountered the need and produced virtual machinery of even greater complexity and sophistication, serving information processing requirements for organisms, whose virtual machinery included interacting sensory qualia, motivations, intentions, plans, emotions, attitudes, preferences, learning processes, and various aspects of self-consciousness.

[BACK TO CONTENTS.](#)

The future of AI

As far as I know, we still don't know how to make machines able to replicate the mathematical insights of ancient mathematicians like Euclid e.g. with 'triangle qualia' that include awareness of mathematical possibilities and constraints, or minds that can discover the possibility of extending Euclidean geometry with the neusis construction, mentioned [below](#).

For further discussion of roles of 'triangle qualia' in discoveries made by ancient mathematicians see the other web pages linked from this one.

NOTE

It is not clear whether we simply have not been clever enough at understanding the problems and developing the programs, or whether we need to extend the class of virtual machines that can be run on computers, or whether the problem is that animal brains use kinds of virtual machinery that **cannot** be implemented using the construction kits known to modern computer science and software engineering. As Turing hinted in his 1950 paper: aspects of chemical computation may be essential.

Biological organisms also cannot build such minds directly from atoms and molecules. They need many intermediate DCKs, some of them concrete and some abstract, insofar as some construction kits, like some animal minds, use virtual machines.

Evolutionary processes must have produced construction kits for abstract information processing machinery supporting increasingly complex multi-functional virtual machines, long before human engineers discovered the need for such things and began to implement them in the 20th Century.

Studying such processes is very difficult because virtual machines don't leave fossil records (though some of their products do). Moreover details of recently evolved virtual machinery may be at least as hard to inspect as running software systems without built-in run-time debugging 'hooks'. This could, in principle, defeat all known brain scanners.

'Information' here is not used in Shannon's sense (concerned with mechanisms and vehicles for storage, encoding, transmission, decoding, etc.), but in the much older sense familiar to Jane Austen and used in her novels e.g. *Pride and Prejudice*, in which how information content is used is important, not how information bearers are encoded, stored, transmitted, received,

etc. The primary use of information is for control.

Communication, storage, reorganisation, compression, encryption, translation, and many other ways of dealing with information are all secondary to the use for control. Long before humans used structured languages for communication, intelligent animals must have used rich languages with structural variability and compositional semantics internally, e.g. in perception, reasoning, intention formation, wondering whether, planning and execution of actions, and learning.

We can search for previously unnoticed evolutionary transitions going beyond the examples here (e.g. Figure 1), e.g. transitions between organisms that merely react to immediate chemical environments in a primaevial soup, and organisms that use temporal information about changing concentrations in deciding whether to move or not.

Another class of possible examples include new mechanisms required after the transition from a liquid based life form to life on a surface with more stable structures (e.g. different static resources and obstacles in different places), or a later transition to hunting down and eating mobile land-based prey, or transitions to reproductive mechanisms requiring young to be cared for, etc.? Perhaps after analysing enough such transitions we'll be able to guess at some of previously unnoticed information processing mechanisms discovered by evolution and then understand how to use them to extend AI in deep ways. It may, or may not, depend on some previously unnoticed feature of fundamental physics, or perhaps just some well known fact about chemistry that is important for operation of neurones (e.g. computations in synapses [Grant \(2010\)](#)), [Gallistel&Matzel, 2012](#) and [Trettenbrein\(2016\)](#).

Compare Schrödinger's discussion, in [1944](#), of the relevance of quantum mechanisms and chemistry to the storage, copying, and processing of genetic information, and [Sloman \(2013b\)](#). I am suggesting that questions about evolved intermediate forms of information processing are linked to philosophical questions about the nature of mind, the nature of mathematical discovery, and deep gaps in current AI.²⁷

NOTES:

Boden ([1990](#)) distinguishes H-Creativity, which involves being historically original, and P-Creativity, which requires only personal originality. The distinction is echoed in the phenomenon of convergent evolution, illustrated in

<https://en.wikipedia.org/wiki/List%20of%20examples%20of%20convergent%20evolution>

The first species with some design solution exhibits H-creativity of evolution. Species in which that solution evolves independently later exhibit a form of P-creativity.

Why did Turing write in his [his 1950 paper](#) that chemistry may turn out to be as important as electricity in brains?

[BACK TO CONTENTS.](#)

Links to examples

I have many examples related to reasoning in geometry, topology, one-to-one correspondences, and consequences for arithmetic, e.g.:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/torus.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/rubber-bands.html>

A possible route to discovery of prime numbers:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html#primes>

Further examples of perception of spatial impossibility:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html>

and the discussion of "toddler theorems" in [Sloman \(2013c\)](#).)

Related document

An extended abstract for a closely related invited talk at the AISB Symposium on computational modelling of emotions is also available online at:

<http://www.cs.bham.ac.uk/research/projects/cogaff/aisb17-emotions-sloman.pdf>

[BACK TO CONTENTS.](#)

Angle trisection

A striking example is angle trisection, mentioned earlier. It is well known (though not easy to prove) that although [bisecting](#) an arbitrary angle is easy in Euclidean geometry [trisecting](#) an arbitrary angle is impossible. However, there is a simple extension to Euclidean geometry, known to Archimedes, the "neusis" construction, that makes it possible to trisect an arbitrary angle easily, as explained here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/trisect.html>

The possibility of extending Euclidean geometry by adding this construction and showing its power, was done without the use of modern logic, algebra, set theory, proof theory etc.

As far as I know, there is no current AI reasoning system capable of discovering such a construct, or considering whether it is an acceptable extension to Euclid's straight-edge and compasses constructs, checking whether it does provide a way to trisect any angle. Similar comments can be made about Mary Pardoe's proof of Euclid's triangle sum theorem, discussed [above](#). It uses a construction that is related to the neusis construction, insofar as both allow a straight-edge to be rotated.

One of the requirements for an adequate AI model or replica or theory of human intelligence is that it should be able to model those ancient discovery processes, as well as more recent forms of mathematical reasoning.

Perhaps future AI will overcome the current limitations, with deep implications for psychology and neuroscience, as well as philosophy of mathematics.

I suspect that very detailed examination of different forms of reasoning actually used by mathematicians of various levels of sophistication, along with some informed guesswork about the forms of reasoning leading up to all the discoveries assembled by Euclid, will one day enable future AI researchers to implement mathematical reasoners that are more like the earliest human mathematicians than current theorem provers are.

Subsets of those capabilities will turn out to be relevant to explaining competences shared with many other intelligent species. I also suspect the mechanisms will turn out to be crucially related to perception and use of spatial affordances by ancient humans and many other animals, though only humans seem to have the meta-cognitive reflective capabilities and capabilities to communicate what they have learnt and argue about it.

I am not suggesting that the abilities are innate in humans. On the contrary many forms of information processing (computation) in humans instead of being directly specified in the genome seem to be products of complex interactions between the genome and the environment, across several stages of development, illustrated in the sketch of epigenetic mechanisms [Figure EPI](#) above. This is close to Kant's claim: the knowledge is not *derived from* experience, but is *awakened by* experience: i.e. he seems to have had a deep epigenetic theory though he lacked the conceptual tools and biological knowledge required to develop it.

Likewise any adequate future neuroscience should be able to explain which features of brains (neural/sub-neural ...) make those discovery processes possible, including how those mechanisms enable/support understanding of the mathematical proofs that are involved.

That kind of understanding should go beyond blind reconstruction or "parroting" of the proofs, as sometimes happens when children are taught mathematics by inadequate teachers.

One extreme mode of "discovery" could be provided by a systematic generator of all possible sequences of characters, or words in a human language (up to some maximum length). But that would not provide the ability to *understand* mathematical proofs, or to use them to solve novel problems or discover deep new questions, as good mathematicians typically do. (An example is the question about motion of vertex of a triangle presented [above](#) and discussed in more detail in [a separate document](#).)

[BACK TO CONTENTS](#).

ACKNOWLEDGEMENTS

My main debt is to Mr Adendorf (not sure about spelling) who was my mathematics teacher and inspirer at South African College School (SACS) in Cape Town before I went to university in 1953. I was introduced to Euclidean geometry by him. My ideas about evolution and epigenesis were extended radically by Jackie Chappell when she came from Oxford to Birmingham in 2004. I have learnt from many colleagues and students. Recently I have had useful challenges and corrections from students, including [Auke Booij](#) (Birmingham) who several times found counter-examples to my proposed geometric/topological examples and [Aviv Keren](#) (Jerusalem) who also works on mathematical cognition.

REFERENCES

To be re-formatted ... one day.

Graham Bell, *Selection The Mechanism of Evolution*, OUP, 2008. Second Edition.

M. A. Boden, *The Creative Mind: Myths and Mechanisms*, Weidenfeld & Nicolson, London, 1990. (Second edition, Routledge, 2004).

M. A. Boden, 2006, *Mind As Machine: A history of Cognitive Science (Vols 1--2)*, OUP, Oxford
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/boden-mindasmachine.html>

[Pythag] Alexander Bogomolny, (2017) Pythagorean Theorem and its many proofs from
Interactive Mathematics Miscellany and Puzzles

<http://www.cut-the-knot.org/pythagoras/index.shtml>

Accessed 15 August 2017

Euclid and John Casey, *The First Six Books of the Elements of Euclid*, Project Gutenberg, Salt Lake City, Apr, 2007, <http://www.gutenberg.org/ebooks/21076>

Also see "The geometry applet"

<http://aleph0.clarku.edu/~djoyce/java/elements/toc.html> (HTML and PDF)

Jackie Chappell and Aaron Sloman, (2007) "Natural and artificial metaconfigured altricial information-processing systems", *International Journal of Unconventional Computing*, 3(3), 221-239.

<http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#717>

N. Chomsky, 1965, *Aspects of the theory of syntax*, MIT Press, Cambridge, MA.

Shang-Ching Chou, Xiao-Shan Gao and Jing-Zhong Zhang, 1994, *Machine Proofs In Geometry: Automated Production of Readable Proofs for Geometry Theorems*, World Scientific, Singapore,

<http://www.mmrc.iss.ac.cn/~xgao/paper/book-area.pdf>

Juliet C. Coates, Laura A. Moody, and Younousse Saidi, "Plants and the Earth system - past events and future challenges", *New Phytologist*, 189, 370-373, (2011).

Alan Turing - His Work and Impact, eds., S. B. Cooper and J. van Leeuwen, Elsevier, Amsterdam, 2013. ([contents list](#)).

Kenneth Craik, 1943, *The Nature of Explanation*, Cambridge University Press, London, New York,

Deaf Studies Trust, 'Deaf Children Beginning to Sign', A paper prepared for Frank Barnes School, London, 17th February 1995

D. C. Dennett, 1978 *Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press, Cambridge, MA.

D. C. Dennett, 1995, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Penguin Press, London and New York,

D.C. Dennett, 1996

Kinds of minds: towards an understanding of consciousness,
Weidenfeld and Nicholson, London, 1996,

<http://www.amazon.com/Kinds-Minds-Understanding-Consciousness-Science/dp/0465073514>

T. Froese, N. Virgo, and T. Ikegami, Motility at the origin of life: Its characterization and a model', *Artificial Life*, 20(1), 55-76, (2014).

Igor Ganti, 2003 *The Principles of Life*, OUP, New York, Eds. Eors Szathmary & James Griesemer, Translation of the 1971 Hungarian edition.

Gallistel, C.R. & Matzel, L.D., 2012(Epub), The neuroscience of learning: beyond the Hebbian synapse, *Annual Review of Psychology*, Vol 64, pp. 169--200,

<https://doi.org/10.1146/annurev-psych-113011-143807>

H. Gelernter, 1964, Realization of a geometry-theorem proving machine, in *Computers and Thought*, Eds. Feigenbaum, Edward A. and Feldman, Julian, pp. 134-152, McGraw-Hill, New York, Re-published 1995 (ISBN 0-262-56092-5),

<http://dl.acm.org/citation.cfm?id=216408.216418>

Also at

<https://pdfs.semanticscholar.org/2edc/8083073837564306943aab77d6dcc19d0cdc.pdf>

J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, MA, 1979.

Seth G.N. Grant, 2010, Computing behaviour in complex synapses: Synapse proteome complexity and the evolution of behaviour and disease, *Biochemist* 32, pp. 6-9,

<http://www.biochemist.org/bio/default.htm?VOL=32&ISSUE=2>

M. M. Hanczyc and T. Ikegami, 'Chemical basis for minimal cognition', *Artificial Life*, 16, 233-243, (2010).

John Heslop-Harrison, *New concepts in flowering-plant taxonomy*, Heinemann, London, 1953.

Immanuel Kant, *Critique of Pure Reason*, Macmillan, London, 1781. Translated (1929) by Norman Kemp Smith.

Various online versions are also available now.

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

Stuart Kauffman, 1995 *At home in the universe: The search for laws of complexity*, Penguin Books, London.

M.W. Kirschner and J.C. Gerhart, *The Plausibility of Life: Resolving Darwin's Dilemma*, Yale University Press, Princeton, 2005.

D. Kirsh, "Today the earwig, tomorrow man?", *Artificial Intelligence*, 47(1), 161-184, (1991).

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

John McCarthy and Patrick J. Hayes, 1969, "Some philosophical problems from the standpoint of AI", *Machine Intelligence 4*, Eds. B. Meltzer and D. Michie, pp. 463--502, Edinburgh University Press,

<http://www-formal.stanford.edu/jmc/mcchay69/mcchay69.html>

John McCarthy, "The well-designed child", *Artificial Intelligence*, 172(18), 2003-2014, (2008). (Written in 1996).

Tom McClelland, (2017) AI and affordances for mental action, in Computing and Philosophy Symposium, *Proceedings of the AISB Annual Convention 2017* pp. 372-379. April 2017.

<http://wrap.warwick.ac.uk/87246>

Tom McClelland, 2017 "The Mental Affordance Hypothesis" MindsOnline 2017

<http://mindsonline.philosophyofbrains.com/2017/2017-session-1/the-mental-affordance-hypothesis/>

Video presentation <https://www.youtube.com/watch?v=zBqGC4THzqg>

Kenneth Manders (1998) The Euclidean Diagram, reprinted 2008 in *The Philosophy of Mathematical Practice*, OUP, pp.80--133 Ed Paolo Mancosu,
<http://doi.org/10.1093/acprof:oso/9780199296453.003.0005>

Kenneth Manders (2008) "Diagram-Based Geometric Practice", In Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*. OUP, pp.65--79

<http://doi.org/10.1093/acprof:oso/9780199296453.003.0004>

Nathaniel Miller, 2007, *Euclid and His Twentieth Century Rivals: Diagrams in the Logic of Euclidean Geometry*, Center for the Study of Language and Information, Stanford Studies in the Theory and Applications of Diagrams,

<https://web.stanford.edu/group/cslipublications/cslipublications/site/9781575865072.shtml>

David Mumford(Blog), Grammar isn't merely part of language, Oct, 2016, Online Blog,

<http://www.dam.brown.edu/people/mumford/blog/2016/grammar.html>

Jean Piaget, 1952, *The Child's Conception of Number*, Routledge & Kegan Paul, London.

W. T. Powers, *Behavior, the Control of Perception*, Aldine de Gruyter, New York, 1973.

Michael Rescorla, (2015) The Computational Theory of Mind, in *The Stanford Encyclopedia of Philosophy*, Ed. E. N. Zalta, Winter 2015,

<http://plato.stanford.edu/archives/win2015/entries/computational-mind/>

Philippe Rochat, 2001, *The Infant's World*, Harvard University Press, Cambridge, MA,

C. Rutz, S. Sugawara, J E M van der Wal, B C Klump, & J St Clair, 2016, 'Tool bending in New Caledonian crows' in *Royal Society Open Science*, Vol 3, No. 8, 160439.

DOI: 10.1098/rsos.160439

<https://research-repository.st-andrews.ac.uk/handle/10023/9283>

A. Sakharov (2003 onwards)

Foundations of Mathematics (Online References)

Alexander Sakharov, with contributions by Bhupinder Anand, Harvey Friedman, Haim Gaifman, Vladik Kreinovich, Victor Makarov, Grigori Mints, Karlis Pdnieks, Panu Raatikainen, Stephen Simpson,

"This is an online resource center for materials that relate to foundations of mathematics (FOM). It is intended to be a textbook for studying the subject and a comprehensive reference. As a result of this encyclopedic focus, materials devoted to advanced research topics are not included. The author has made his best effort to select quality materials on www."

<http://sakharov.net/foundation.html>

NOTE: some of the links to other researchers' web pages are out of date, but in most cases a search engine should take you to the new location.

Dana Scott, 2014, Geometry without points. (Video lecture, 23 June 2014, University of Edinburgh)

<https://www.youtube.com/watch?v=sDGnE8eja5o>

Erwin Schrödinger, *What is life?*, CUP, Cambridge, 1944.

Commented extracts available here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/schrodinger-life.html>

Claude Shannon, (1948), A mathematical theory of communication, in *Bell System Technical Journal*, July and October, vol 27, pp. 379--423 and 623--656,

<https://archive.org/download/pdfy-nl-WZBa8gJF18QNh/shannon1948.pdf>

Stewart Shapiro, 2009 We hold these truths to be self-evident: But what do we mean by that? *The Review of Symbolic Logic*, Vol. 2, No. 1

<https://doi.org/10.1017/S175502030909011X>

A. Sloman, 1962, *Knowing and Understanding: Relations between meaning and truth, meaning and necessary truth, meaning and synthetic necessary truth* (DPhil Thesis), PhD. dissertation, Oxford University, (now online)

<http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-1962>

A. Sloman, 1971, "Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence", in *Proc 2nd IJCAI*, pp. 209--226, London. William Kaufmann. Reprinted in *Artificial Intelligence*, vol 2, 3-4, pp 209-225, 1971.

<http://www.cs.bham.ac.uk/research/cogaff/62-80.html#1971-02>

An expanded version was published as chapter 7 of [Sloman 1978](#), available [here](#).

A. Sloman, 1978 *The Computer Revolution in Philosophy*, Harvester Press (and Humanities Press), Hassocks, Sussex.

<http://www.cs.bham.ac.uk/research/cogaff/62-80.html#crp>

A. Sloman, 1984, The structure of the space of possible minds, in *The Mind and the Machine: philosophical aspects of Artificial Intelligence*, Ed. S. Torrance, Ellis Horwood, Chichester,

<http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#49a>

A. Sloman, 1996, Actual Possibilities, in *Principles of Knowledge Representation and Reasoning* (Proc. 5th Int. Conf on Knowledge Representation (KR '96)), Eds. L.C. Aiello and S.C. Shapiro, Morgan Kaufmann, Boston, MA, pp. 627--638,
<http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15>

A. Sloman, (2000) "Interacting trajectories in design space and niche space: A philosopher speculates about evolution", in *Parallel Problem Solving from Nature (PPSN VI)*, eds. M.Schoenauer, et al. Lecture Notes in Computer Science, No 1917, pp. 3-16, Berlin, (2000). Springer-Verlag.

A. Sloman, 2001, Evolvable biologically plausible visual architectures, in *Proceedings of British Machine Vision Conference*, Ed. T. Cootes and C. Taylor, BMVA, Manchester, pp. 313--322,
<http://www.cs.bham.ac.uk/research/projects/cogaff/00-02.html#76>
<http://dx.doi.org/10.5244/C.15.33>},

A. Sloman, 2002, The irrelevance of Turing machines to AI, in *Computationalism: New Directions*, Ed. M. Scheutz, MIT Press, Cambridge, MA, pp. 87--127,
<http://www.cs.bham.ac.uk/research/cogaff/00-02.html#77>

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

A. Sloman, 2008, The Well-Designed Young Mathematician, *Artificial Intelligence*, 172, 18, pp. 2015--2034, Elsevier.
<http://www.cs.bham.ac.uk/research/projects/cogaff/08.html#807>

A. Sloman, (2008a). Architectural and representational requirements for seeing processes, proto-affordances and affordances. In A. G. Cohn, D. C. Hogg, R. Moller, & B. Neumann (Eds.), *Logic and probability for scene interpretation*. Dagstuhl, Germany: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany.
<http://drops.dagstuhl.de/opus/volltexte/2008/1656>

A. Sloman, 2009, Architecture-Based Motivation vs Reward-Based Motivation, *Newsletter on Philosophy and Computers*, American Philosophical Association, 09,1, pp. 10--13, Newark, DE, USA
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/architecture-based-motivation.html>

A. Sloman, 2011, What's information, for an organism or intelligent machine? How can a machine or organism mean? In *Information and Computation*, Eds. G. Dodig-Crnkovic and M. Burgin, World Scientific, pp.393--438,
<http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#905>

A.Sloman (2010-2017), A (Possibly) new kind of (non?) Euclidean geometry, based on an idea by Mary Pardoe.
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/p-geometry.html>

A. Sloman, 2013a "Virtual Machine Functionalism (The only form of functionalism worth taking seriously in Philosophy of Mind and theories of Consciousness)", Research note, School of Computer Science, The University of Birmingham.
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-func.html>

A. Sloman, 2013b "Virtual machinery and evolution of mind (part 3) Meta-morphogenesis: Evolution of information-processing machinery', in *Alan Turing - His Work and Impact*, eds., S. B. Cooper and J. van Leeuwen, 849-856, Elsevier, Amsterdam.

<http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1106d>

Project Web page: <https://goo.gl/9eN8Ks>

A. Sloman, (2013c), Meta-Morphogenesis and Toddler Theorems: Case Studies, Online discussion note, School of Computer Science, The University of Birmingham,

<http://goo.gl/QgZU1g>

Aaron Sloman (2007-2014)

Unpublished discussion Paper: Predicting Affordance Changes: Steps towards knowledge-based visual servoing. (Including videos).

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/changing-affordances.html>

A. Sloman (2015). What are the functions of vision? How did human language evolve? Online research presentation.

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk111>

A. Sloman 2017, "Construction kits for evolving life (Including evolving minds and mathematical abilities.)" Technical report (work in progress)

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/construction-kits.html>

(An earlier version, frozen during 2016, was published in a Springer Collection in 2017:

https://link.springer.com/chapter/10.1007%2F978-3-319-43669-2_14

in *The Incomputable Journeys Beyond the Turing Barrier*

Eds: S. Barry Cooper and Mariya I. Soskova

<https://link.springer.com/book/10.1007/978-3-319-43669-2>)

Aaron Sloman, Jackie Chappell and the CoSy PlayMate team, 2006, Orthogonal recombinable competences acquired by altricial species (Blankets, string, and plywood) School of Computer Science, University of Birmingham, Research Note COSY-DP-0601,

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/orthogonal-competences.html>

Aaron Sloman and David Vernon. A First Draft Analysis of some Meta-Requirements for Cognitive Systems in Robots, 2007. Contribution to euCognition wiki.

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-requirements.html>

P. F. Strawson, *Individuals: An essay in descriptive metaphysics*, Methuen, London, 1959.

Max Tegmark, 2014, *Our mathematical universe, my quest for the ultimate nature of reality*, Knopf (USA) Allen Lane (UK), (ISBN 978-0307599803/978-1846144769)

Arnold Trehub, 1991, *The Cognitive Brain*, MIT Press, Cambridge, MA,

<http://people.umass.edu/trehub/>

Trettenbrein, Patrick C., 2016, The Demise of the Synapse As the Locus of Memory: A Looming Paradigm Shift?, *Frontiers in Systems Neuroscience*, Vol 88,

<http://doi.org/10.3389/fnsys.2016.00088>

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

A. M. Turing, (1952) "The Chemical Basis Of Morphogenesis", *Phil. Trans. Royal Soc. London B* 237, 237, 37-72.

Note: A presentation of Turing's main ideas for non-mathematicians can be found in Philip Ball, 2015, "Forging patterns and making waves from biology to geology: a commentary on Turing (1952) 'The chemical basis of morphogenesis'", <http://dx.doi.org/10.1098/rstb.2014.0218>

Barbara Vetter (2011), Recent Work: Modality without Possible Worlds, *Analysis*, 71, 4, pp. 742--754, <https://doi.org/10.1093/analys/anr077>

C. H. Waddington, 1957 *The Strategy of the Genes. A Discussion of Some Aspects of Theoretical Biology*, George Allen & Unwin, 1957.

R. A. Watson and E. Szathmary, "How can evolution learn?', *Trends in Ecology and Evolution*, 31(2), 147-157, (2016).

Weir, A A S and Chappell, J and Kacelnik, A, (2002) Shaping of hooks in New Caledonian crows, *Science*, vol 297, p 981,

L. Wittgenstein (1956), *Remarks on the Foundations of Mathematics*, translated from German by G.E.M. Anscombe, edited by G.H. von Wright and Rush Rhees, first published in 1956, by Blackwell, Oxford. There are later editions. (1978: VII 33, p. 399)

About this document

Revised versions of this document will be available here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/ijcai-2017-cog.html>

and a PDF version.

The 42 minute video prepared for the IJCAI workshop presentation on 19th August 2017 is available here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/ai-cogsci-bio-sloman.webm>

A partial index of discussion notes is in

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/AREADME.html>

Installed: 10 Aug 2017

Last updated: 18 Aug 2017; 22 Aug 2017; 27 Aug 2017; 29 Aug 2017; 12 Sep 2017; 18 Sep 2017; 12 Oct 2017

Maintained by [Aaron Sloman](#)
[School of Computer Science](#)
[The University of Birmingham](#)
