Mathematical cognition is primarily about necessities and impossibilities, recognized through understanding of structural relationships

\textbf{NOT}

regularities and probabilities derived by collecting statistical evidence.

E.g. if spatial volume $V_1$ contains $V_2$, and $V_2$ contains $V_3$, then $V_1$ \textbf{must} contain $V_3$. 
I.e. $V_1$ contains $V_2$, and $V_2$ contains $V_3$ and $V_1$ does not contain $V_3$ is \textbf{impossible}.

Similarly $3+5$ \textbf{must} equal 8. It \textbf{cannot} equal 9. 
(Immanuel Kant pointed out this feature of mathematics in 1781.)

Therefore, neural nets that merely derive probabilities from statistical data cannot explain those key kinds of mathematical cognition.

Surprisingly many neuroscientists, psychologists, philosophers and AI researchers are blind to this limitation, including some much admired for research on mathematical cognition!

Euclid’s Elements is full of examples but stupid 20th century educational decisions have deprived a high proportion of intelligent learners of opportunities to experience discovering geometrical proofs and refutations.

Alternate titles:
-- Impossibility: the dark face of necessity.
-- What don’t we know about spatial perception/cognition? especially the perceptual abilities of ancient mathematicians?
-- Aspects of mathematical consciousness,
-- Possible impossible contents of consciousness!

Current AI vision systems lack mathematical qualia, experienced by Euclid, Archimedes and many of their contemporaries, predecessors and successors!
This is an attempt to understand requirements for removing that gap in AI, inspired in part by Immanuel Kant, Oscar Reutersvard, James Gibson, Roger Penrose and especially Max Clowes (1933-1981), who introduced me to AI work on scene analysis around 1969. See Appendix 2 of this \textit{memorial tribute}. Roger Penrose has thought and written about such
matters including the well known Penrose impossible triangle also discussed below. Our analyses are very different, however!

Wittgenstein, *Tractatus* 3.0321: "Though a state of affairs that would contravene the laws of physics can be represented by us spatially, one that would contravene the laws of geometry cannot."

But what about:

Here’s a blue square seen from the edge:

Here’s a red circular disc seen from the edge:

Here’s a green circular square seen from the edge:
NOTE: This is work in progress
This is part of the Turing-inspired Meta-Morphogenesis project, concerned with identifying and explaining the many transitions in types of information-processing in the course of biological evolution on Earth:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html

This document has been through several major reorganisations, which may have led to internal inconsistencies, duplications that need to be removed, and poor formatting, to be fixed later. Some of the ideas go back to my 1962 DPhil thesis defending Kant’s philosophy mathematics, written before I had learnt about AI or computers.

THANKS

My thanks to Dima Damen http://www.cs.bris.ac.uk/~damen/, for the invitation to talk about vision at Bristol University, 2nd Oct 2015, which launched this document.

Also colleagues, students, and friends over many years, who introduced me to AI and a new way of thinking about minds, including vision. Thanks to Aviv Keren, for useful comments on earlier versions: https://www.researchgate.net/profile/Aviv_Keren

FORMATS:
This document is available in html and pdf
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html
(http://goo.gl/Zz2O1l)
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.pdf

A partial index of discussion notes is in
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/AREADME.html

NOTE The Meta-Morphogenesis project:
An introduction to the Turing-inspired Meta-Morphogenesis project can be found here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html

It includes a large, growing and messy collection of draft papers on evolution of biological information processing mechanisms, partly inspired by the work of Alan Turing. Recurring themes in this work include the role of implicit mathematical discoveries made by biological evolution (natural selection as a "Blind Theorem-Prover") Around November 2014 the project began to emphasise ‘construction kits’ of many sorts, including the fundamental construction kit (FCK) provided by physics and chemistry, and increasingly complex and more specialised derived construction kits (DCKs).
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/construction-kits.html
CONTENTS

- CONTENTS (This list)
- Background
- Some Kantian examples (Added 29 Oct 2018)
- Personal history
- What are the functions of perception?
- Introduction: Gibson's notion of affordance
- Beyond Gibson: Possibilities, impossibilities, necessities
- Online and offline uses of spatial information
- Note on the irrelevance of "possible world" semantics
- Requirements for the online and offline functions of vision
- Biological examples of offline use of vision
  - Planning and deliberation by portia spiders
  - Proto-affordances
- Representing possible processes
- Possibility perception
- Possibilities vs Probabilities
- Temporally chained possibilities
- Perceiving vs understanding
- Pictures of possible and impossible object configurations
- What can you do with a collection of eight similar cubes?
- What if you had nine similar cubes?
- More possible moves
- More challenges to psychology, neuroscience and AI
- Reutersvard 1934
- Pre-20th Century pictures of impossible scenes
- Mis-use of the word "illusion"
- Dice example (Video)
- Tangled bodies (Added 16 Oct 2017)
- Impossible transitions on a grid
- Figure Numerical
- Figure Accessibility
- Impossible number relationships
- Discovering primeness geometrically
- Physical constraints vs rule-based constraints (Added: 18 Jan 2016)
- Domestic possibilities and impossibilities (furniture)
- Constructive, supportive impossibilities
- Impossibilities concerned with cardinality
- Non-human reasoners about possibilities and impossibilities
- Toddler Topology
- The Side Stretch Theorem
- The Side Containment Theorem (SCT)
- Properties of triangles and other abstract shapes
Background

The original version of this document was intended for a presentation on vision at Bristol University, in October 2015. After the event I kept on adding examples, and attempting to clarify their theoretical significance.

The nature of the subject matter, including the variety of examples that kept turning up, forced the document to grow considerably beyond the first draft. At some future date it may become a collection of separate documents. As indicated above, this is part of the Turing-inspired Meta-Morphogenesis project, begun in 2012: http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html

My original interest in the topics presented here was sparked about 60 years ago, when I attempted to clarify, defend and extend Immanuel Kant’s claims in Kant (1781), about the nature of mathematical discovery, as reported in my Oxford University DPhil thesis(1962).
That defence requires, among other things, analysing requirements for perceptual systems, especially spatial perception in animals and future intelligent machines. At that time, I knew nothing about computers or AI. I had heard of Turing and Turing machines, but did not study any details, despite the fact that, for a short time my supervisor was Hao Wang, who was then working on a logical theorem prover for IBM.


This paper was originally intended to be a much simpler document focused mainly on lessons to be learnt from our ability to perceive, think about, and reason about not only what currently exists or is happening in the environment but also what could, could not, or must, exist or happen. These can be described as modal features of the environment, in contrast with categorical features, that are restricted to what is the case and statistical features that summarise regularities, in the form of countable or measurable ratios of frequencies of occurrence of various types of object, state, or process.

The ability to perceive and reason about modal features can be described as modal competences. As Kant recognised, Kant (1781), these competences are essential to ancient mathematical discoveries, which are concerned with what is possible, impossible, or necessarily the case. Also, as Kant pointed out, they are not all derivable from explicit definitions using logic -- i.e. they are synthetic, not analytic, in his terminology. I think that was his way of saying that they were not merely examples of definitions, or purely logical deductions from definitional truths. In contrast this is analytic (in English): "No bachelor uncle is an only child".

The competences involved in discovering modal features (e.g. impossibility, necessary truth) are totally different from competences concerned with perception of and reasoning about what occurs more or less frequently, or what is more or less probable.

This means that statistical evidence combined with probabilistic reasoning cannot explain the competences: statistical evidence cannot prove that something is impossible (e.g. a 7-sided regular polyhedron) or that something is necessarily the case (e.g. a regular polyhedron has faces with at most 5 edges).

It follows that statistics-based neural nets, however deep, cannot make such ancient mathematical discoveries, despite the huge, and steadily growing, collection of partial successes (many of which are inherently deceptive).

I have the impression that many researchers in psychology, neuroscience, and neural-net inspired AI (e.g. Deep Learning mechanisms) have never noticed that key features of mathematical discovery, closely related to important kinds of practical spatial reasoning and perception of positive and negative affordances, involve discovering facts about impossibilities and necessary connections. Such facts cannot be expressed in terms of probabilities.

I hope at some point to produce a first draft characterisation of the examples below in terms of the mechanisms required to become aware of the relevant kinds of impossibility (or necessity). However that is a very difficult task, and may require new forms of computation.

For example, a particular formula or inference in propositional calculus (boolean algebra) will have a finite number of variables each capable of being true or false, so the total set of possibilities within which something can be found to be necessary or impossible is finite
(though it may be very large, since adding a new variable doubles the set of possibilities). It is easy to program a computer to check such collections of possibilities in order to decide whether an inference is valid or invalid, or a proposition is necessarily true, or necessarily false, or neither.

In contrast considering whether two simple closed curves on a surface (e.g. two curves on a torus) can be continuously deformed into each other requires consideration of an infinite set of spatial configurations, in which both the shapes and the locations of the curves vary. Discovering impossibility (or necessity) in that sort of context is much more difficult. I am not aware of any serious contenders for explanatory mechanisms (although there are theorem provers that can handle "arithmetized" geometrical reasoning, e.g. using Cartesian coordinates, which were not required or used by the ancient mathematicians (or squirrels and toddlers solving spatial problems).

Some examples include both discrete and continuous variation. For example it is possible to generate 2D polygons on a surface by joining chains of straight line segments. If a polygon is composed of only three lines of fixed sizes the shape can be translated, or rotated, but angles cannot change. Adding a fourth side produces a structure capable of infinitely many variations in shape, with angles changing. How can you be sure that any planar quadrilateral, no matter what its size and shape, can be continuously deformed into another shape without altering the lengths of its sides, only the angles at which they meet?

For example, if one side (any side) of the quadrilateral is fixed in place, then, by moving the other sides it is possible to change the angles at the corners of the quadrilateral without changing the length of any side, but it is not possible to change any angle while the other three angles remain fixed. What brain mechanisms make it possible to recognize such spatial possibilities and impossibilities, without any actual changes of length or angle occurring?

Many organisms have body parts that are not rigidly connected, so that they have infinitely many possible configurations in which angles vary. Some organisms can also inflate or deflate parts of their bodies continuously (e.g. mouth cavity and chest cavity volume in humans, or the configuration of tongue, lips, cheeks and teeth), adding to the kinds of infinite variability of which they are capable -- which organisms perceiving them may need to be able to understand and reason about. What brain mechanisms make it possible to reason about such configuration changes without changing the configurations, as squirrels seem to do in choosing actions to achieve difficult goals? (Squirrel intelligence and slug intelligence are compared in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/squirrel-intelligence.html.)

What additional mechanisms are required to be able to reason about the limits of such reasoning abilities in other individuals, or oneself?

Many species also build external structures, including nests, webs, sand-structures, tunnels, etc., that allow both continuous and discrete spatial changes. How much they understand about what they are doing and how they interact with the results is unclear, as opposed to merely having evolved reactions that suffice for their needs. In contrast human engineers, architects and users of their products all need deep understanding of spatial and causal variation that involves combinations of discrete and continuous change. Spatial intelligence includes abilities to reason about some of the limits of such variation.
The possibility-spaces just mentioned typically include processes extended over time, and in many cases collections of interacting concurrent physical processes with discrete and continuous sub-processes. In contrast, all processes in a digital computer (at least at the level of the digital electronics) are discrete, although they may be used to approximate continuous processes, e.g. in video displays of moving objects. However, the ability to generate complex continuous processes (or detailed simulations of them) is totally different from the ability to understand the possibility sub-spaces and their constraints (impossibilities and necessities). E.g. a tropical sandstorm has no understanding of tropical sandstorms, and what they can and cannot do.

The existence of human (and non-human) spatial reasoning capabilities of so many types is currently unmatched in AI and unexplained in neuroscience (and mostly unnoticed in psychology -- Piaget being one of the few exceptions).

The collection of examples below needs to be reorganised in such a way as to indicate different collections of requirements for mechanisms that understand why some things are impossible or necessary. Statistical learning and probabilistic reasoning cannot achieve that.

As far as I know there is nothing in current psychology, neuroscience, or AI that gives any indication of how animal brains can discover or represent impossibilities or necessary connections, a requirement for explaining human mathematical abilities. Piaget, who studied Kant, is one of the few psychologists I’ve encountered who understood the need for such explanations. But he lacked conceptual tools capable of formulating explanations. It is not clear that current AI provides such tools, except for limited classes of logic-based reasoning and discovery. For more on this see the Meta-Morphogenesis project mentioned above.

Researchers who have never previously noticed or learnt about this (Kantian, modal) feature of mathematical knowledge, seem to find it hard to understand at first. (Or perhaps I am bad at explaining...)

This prevents them grasping the deep limitations of statistics-based intelligence, i.e. intelligence based on abilities to acquire, reason about and use statistical information and probabilities. (Impossibility and necessity are not the same concepts as 0% and 100% probability.)

These scientific/philosophical failures of observation (or analysis) can lead to very shallow seriously mistaken descriptions and explanations of mathematical competences, when researchers don’t realise that what they have explained is something much simpler/shallower than mathematical competences. (A study of Kant is not normally part of a developmental psychology degree, or training in AI or robotics, unfortunately.)

The perceptual abilities I’ll draw attention to can be thought of as extending James Gibson’s ideas, summarised below. He regarded the function of perception as being primarily to provide perceivers with information about affordances -- that is, information about actions that perceivers can or cannot perform in their current situation, among the actions that might be relevant to their needs or interests Gibson(1979).

This also includes information about how to vary actions, e.g. when to decelerate while approaching a target or obstacle.
Mechanisms for perception of possible actions between which the perceiver can choose must have evolved later than evolution of the much simpler reflex actions triggered without any consideration of alternatives.

Evolutionary developments also provided abilities for some organisms to acquire and use information about things that are impossible, or which could obstruct, or fail to be be useful for a particular goal, i.e. negative affordances. (This can save a lot of time when confronted with difficult tasks.)

These abilities to detect and use positive and negative affordances can be seen as constituting a subset of the phenomena Immanuel Kant thought about in connection with the nature of mathematical knowledge. Some of the connections between Kant and Gibson were pointed out in [Mace,2005], though not the connection with mathematical discovery.

Gibson seems not to have noticed that the perceived affordances he discussed are a subset of a broader collection of modal perceptual competences. An example is the ability to identify possible and impossible structures and processes in the environment, and the ability to think about necessary consequences of possible events or actions that need have nothing to do with the perceiver’s current needs or interests. E.g. seeing that what someone else is trying to do is impossible. I suspect that even some pre-school children will realise not only that it is possible to remove a shoelace from a shoe by pulling one end of the lace, and impossible to remove it, without breaking anything, by pulling both ends at the same time.

It is not hard to grasp that it is possible to add exactly three new buttons to a collection of exactly five buttons, without changing anything else: e.g. one can try it out, or imagine trying it out. However, it not easy to explain how such a process will necessarily produce a collection of eight buttons. Moreover, this is not a fact specifically about buttons, but about any collection of distinguishable enduring countable objects.

Other examples include abilities to discover necessary truths in topology and geometry, for example that containment is transitive. Examples of such necessary connections can be relevant to a perceiver’s actions but not all need be. For example, if an event happens on Mars and Mars is part of the solar system then the event happens within the solar system. Noticing the necessity has nothing to do with finding practical uses or limitations of the practical use of such information.

Some Kantian examples
(Added 29 Oct 2018)

An example given by Kant is that it is impossible for two straight lines in the same plane to completely enclose a finite portion of the plane. A more complex geometric insight is that is impossible for three plane surfaces to completely enclose a finite portion of 3D space. Those impossibilities could be relevant as negative affordances if someone wished to fence off part of a field by using only two straight fences and no other pre-existing barriers, or wished to create a completely closed container for tools or for an animal that might attempt to escape, made of exactly three flat pieces of material. What brain mechanisms allow humans to recognize such impossibilities?

These examples are closely related to well known facts of Euclidean geometry, though one does not need to have studied mathematics to recognise the claimed necessities and impossibilities:
A closed finite-sized planar polygon must have at least three sides. Why?
A polyhedron (3D space bounded by plane surfaces) must have at least four sides. Why?
A polyhedron bounded by surfaces meeting only at right angles must have at least six sides and at least 12 edges. Why?

Reformulated 12 Mar 2020. Imagine an arbitrary convex polyhedron (i.e. a finite convex solid, bounded entirely by a number of planar surfaces). Any vertex on that polygon could be removed by a single planar slice through the polygon that removes no other vertex. That removal will leave a new planar polygon, containing all the remaining vertices. What will such a vertex-removal process do to the numbers of vertices, edges, and surfaces? How will those numbers differ between the original polygon and the new polygon? I.e., after such a slicing operation will the total number of vertices $V$, the total number of edges $E$, and the total number of planar faces $F$, be the same, or go up or go down? What can you say about the changes in numbers that will occur? How can you be sure?

There are many more illustrations of the fact that, in humans (and possibly several other intelligent types of animal), the functions of vision include perception of modalities (i.e. what is possible, impossible, or necessarily the case). This has nothing to do with discovering probabilities or combining sensory modalities (touch, sound, sight, etc.), though it can use any or all of those sensory modalities.

These discoveries can be about exosomatic information, for example discoveries about what is or is not possible in the environment -- i.e. outside the skin. That contrasts with learning sensory-motor and other somatic relationships (correlations inside the organism’s skin). Evolution made use of many such implicit mathematical discoveries long before there were human mathematicians, but that’s another sub-topic.

I recently learnt that in 1938 Alan Turing had noticed a distinction between mathematical intuition and mathematical ingenuity, claiming that only the latter could be implemented in computers. Most, if not all, of the examples in this document, seem to be illustrations of powers of human mathematical intuition, especially spatial intuition providing mathematical knowledge of geometry and topology. Turing’s ideas are summarised and discussed in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/turing-intuition.html also (Pdf) (Still work in progress.)

As far as I know nobody in AI knows how to replicate these abilities (involving intuition, or insight) in machines, and no psychologists or neuroscientists can explain how brains make such discoveries possible. I hope new answers will eventually emerge from the Meta-Morphogenesis (M-M) project, summarised in:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html

The claims about human abilities to perceive possibilities, impossibilities and necessities are illustrated below using modified versions of a picture drawn by Oscar Reutersvard in 1934, as a key example. However, many other examples are presented. In particular I’ll offer examples related to proto-mathematical discoveries made by pre-verbal human toddlers presented and discussed in this (also very messy) document on “Toddler theorems”:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html
Personal history

I first attempted to explicate the modal concepts used here in chapter 7 of my 1962 DPhil thesis defending Kant’s philosophy of mathematics against attacks by philosophers who had no personal experience of discovering and proving geometrical truths. Kant’s distinctions were also summarised very briefly in Sloman(1965). A digitised, searchable, version of the 1962 thesis was made freely available online in 2016:
http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-1962/ (HTML and PDF)

Some of the material in this document is re-visited in the context of my long term attempts to understand the kinds of reasoning required by such discoveries, and the evolved biological mechanisms that made the reasoning possible -- these abilities remain generally unexplained (as indicated in Turing’s contrast between mathematical intuition and mathematical ingenuity, mentioned above). In 2017 I started trying to spell out requirements for what I’ve temporarily labelled a “super-Turing membrane machine” able to reason about possible and impossible deformations of shapes (e.g. triangles) and the consequences, discussed explicitly, though tentatively, below and in these (mostly draft) documents:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/super-turing-geom.html
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/deform-triangle.html
and implicitly in:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/sloman-ptai17.html

Expanded abstract for PTAI Conference Nov 2017

Some readers may wish to skip the preliminary remarks below and jump straight to examples, e.g. the section on Representing possible processes.

What are the functions of perception?

Many researchers assume that the function of perception is to find out what IS the case, what WAS the case and what WILL BE the case in the environment -- including the immediately perceived environment and the extended environment. Examples include perceiving objects as space-filling solids, even when only their external surfaces are perceived. Sometimes, for various reasons, the inferences from sensory and other information are not totally reliable. Some researchers therefore regard the functions of vision (or more generally perception) as including finding out what is PROBABLY the case.

However, there are important functions of vision that are not included in these obvious functions. Vision also provides information about possibilities, impossibilities, and necessary consequences. That can also include acquiring conditional information, e.g. about what WOULD HAVE BEEN the case if ..., or what WOULD BE or WILL BE the case if... . E.g. perceiving an apple hanging in a tree as supported by its stalk provides information about what would happen if the stalk were to break. As in the previous cases the information available in some cases is incomplete or unreliable or in some other way less than perfect. What is derived is then thought to require inferences about probabilities of various alternatives. However this paper is not concerned with those cases, but cases where a change will make something possible, or impossible, or necessarily the case. E.g. it is possible for me to move closer to an open doorway to another room, and if I do I shall necessarily see more of the room if information travels in straight lines.
These un-noticed or inadequately understood functions of vision are concerned with obtaining information about what is POSSIBLE or IMPOSSIBLE, or NECESSARILY or CONTINGENTLY the case in the environment. (What is contingent is possibly true and possibly false and neither necessarily true nor necessarily false.) For more on these "alethic" modal concepts see https://en.wikipedia.org/wiki/Modal_logic#Alethic_logic.

The spatial perceptual functions described in terms of modal concepts of possibility, impossibility, necessity and contingency, have nothing to do with PROBABILITIES (although probabilities presuppose possibilities). In particular, the concepts "possible", "impossible", and "necessarily true" are totally different from notions of a non-zero, zero, or 100% probability. Probabilities are essentially ratios of numbers produced by counting or measuring.

The functions of vision related to perception of possibilities and impossibilities (constraints on possibilities) seem rarely to be noticed by vision researchers, although some researchers interested in perception have investigated at least some of them, including Immanuel Kant Kant (1781) and more recently James Gibson Gibson (1979)) and his followers. But Gibson and most psychologists, unlike Kant, typically fail to address relationships between mathematical competences and these spatial competences, the main topic of this document. Piaget was an exception, especially in his last book, and his 1952 book e.g. on gradual development of understanding of 1-1 correlations and cardinality.

Probability concepts presuppose concepts of possibility, since probabilities are comparisons among sets of possibilities. These are often partial orderings, sometimes with numerical comparisons added. However, the (alethic) modal concepts of possibility, impossibility, and necessity used here do not presuppose probabilities. In particular, they are totally different from numerical probability concepts.

There are deep unanswered questions about whether and how the alethic modal concepts, "possible", "impossible", "necessary", and "contingent" (= neither impossible nor necessary) are used by other animals, and about how they can be represented in information-processing systems (e.g. in minds of animals or robots). The roles these concepts play in intelligence tend to be mis-described, or ignored by perception researchers, especially their connection with mathematical knowledge.

I shall use a variety of examples to illustrate some of the ways these modal concepts work, why they are important for intelligent animals or machines, how the functions of vision (and more generally perception) involve them, and how they are connected with mathematical discoveries.

**Note:**
A feature of the analysis presented here is rejection of "Possible worlds semantics" for the modal concepts relevant to intelligent agents (including non-human intelligent agents, such as squirrels, and early humans who made ancient mathematical discoveries). For background information on possible world semantics see, for example, http://plato.stanford.edu/archives/sum2015/entries/possible-worlds
http://plato.stanford.edu/entries/logic-modal-origins/.

The modal concepts used here are based on the analysis of Kant’s intentions in Sloman(1962).
One of the important functions of vision is to obtain information about how the environment relates to abilities, risks, needs, or intentions of other agents -- i.e. "vicarious" affordances. Gibson (1966 and 1979) discussed some special cases of this, but I don’t think he saw all the important implications or pre-requisites of being able to see what is relevant to the desires, intentions, preferences, beliefs, of other agents, or oneself. These are topics that need full discussion on another occasion, though they will be briefly mentioned below. (See also Sloman (2009a).)

We’ll see that information about what is or is not possible is relevant both to the immediate practical uses of vision and also to the roles of vision, and meta-cognition, in some types of mathematical discovery. I’ll try to indicate, in very crude outline, how the earliest mathematical discoveries might have been concerned with meta-theories about possibilities for action. The need for theories about possibilities arises naturally for intelligent agents choosing and acting in a structured environment. The need for meta-theories arises if those agents have the ability to detect and reflect on their own theories.

Meta-meta theories are required for reflecting on or discussing the properties of those meta-theories, and how they can be found to be true. The evolutionary changes making that possible also made possible the kinds of mathematical presentation found in Euclid’s work.

Later forms of mathematics (based on formal systems developed in the last two centuries) have other functions, which will not be discussed here. So I’ll ignore most mathematics based on the developments in logic, set-theory and formal systems since around 1900. The view of some mathematicians that what happened earlier was not really mathematical discovery and reasoning is just false, unless the label "mathematics" is re-defined to make it true.

There can be enormous variations between the spatial capabilities and performances of individual humans, individual monkeys, individual crows, etc. For the abilities to use vision to acquire modal information are not all innate: they develop under the influence of the environment. Individuals in different environments will therefore develop different spatial competences. (There may also be genetic differences.)

Learning to perceive different sets of possibilities and constraints can be compared with differences between learning to read French and learning to read Chinese. In both cases there are considerable individual and cultural differences. It follows that requiring experiments on vision to provide reliable repeatable data about humans will rule out many experiments that provide information about important visual/spatial capabilities, since what is true of one person may not be true of another. Yet facts about individuals, even unusual or unique individuals, are part of what science needs to explain, e.g. by explaining how the individuals process information, including how and why they differ.

A good theory of vision should explain how the individual competences work (preferably demonstrated in working AI models) AND how a (mostly) common genetic heritage can produce differences in competences of individuals that share the heritage. That requires a model of individual development, some features of which are sketched in the section on Evo-Devo Issues.

My complaints about wide-spread neglect of important functions of vision apply both to theories of human vision and theories of animal vision, and to statistics-based models and theories of vision that have been successfully applied in special purpose robots and other machines with useful, but very limited functionality. The main theme here is the need for a good theory of vision to be part of
a good theory explaining important types of mathematical discovery.

This point can be generalised: a good theory of mind, or of evolution of minds, or of development of minds needs to explain the abilities of at least one sort of mind to discover and use mathematical truths about what is and is not possible. Such information is very different from abilities to acquire and use information about **probabilities** that many current AI systems focus on. In particular, a good theory of what minds are and how they evolved needs to explain what made it possible for Euclid and other ancient mathematicians to make the discoveries reported in *Euclid's Elements*. Additional examples are presented in these papers and papers they reference:

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-sum.html
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/trisect.html

Although no detailed explanation (or working model) exists, we can discuss requirements for future candidates. Some incomplete conjectures are presented below.

---

**Introduction: Gibson’s notion of affordance**

Some of the earliest AI vision work (and some research in psychology) focused on perception and recognition of 2-D patterns in Images. But it was soon realised that human and animal vision goes beyond that, e.g. because the visible part of the three-dimensional world is projected into a two-dimensional image by the sensing apparatus. So a visual system needs to "reverse" this process: the original three-dimensional reality must be inferred from the two-dimensional image (plus some background knowledge, where necessary). This (obviously) does not involve building a new 3-D structure inside the brain. It requires building an information structure that provides spatial information about the 3-D structure in the environment. Often some or all of the same spatial information can be acquired through other sensory/motor subsystems, including tactile, haptic and vestibular (semi-circular canal) mechanisms.

(Exactly what that information is, and how it is represented is a complex topic: different researchers make different assumptions about this. One assumption many researchers make that I specifically reject, is that 2-D or 3-D spatial information needs to be represented in terms of 2-D or 3-D coordinates of objects or object parts. I think that in many cases it is more like a collection of partial orderings, e.g. relative distance, relative height, relative orientation -- sloping more or less, relative curvature, etc.)

This claim that seeing is "reconstructing" is often attributed to David Marr (1982), though it was taken for granted by AI researchers much earlier, e.g. Roberts, 1965 and others surveyed in Ballard and Brown, 1983, though they proposed different theories about the details.

On that view, the main functions of vision should be the same across all species, though Marr acknowledged that for some species, e.g. insect species, the functions might be different, and of course some human visual capabilities, such as reading text or musical scores, understanding sign languages, and interpreting maps and engineering drawings, are unique to humans.
In opposition to these views, James Gibson ([1966] and [1979]) criticised researchers who thought the function of vision in animals was simply to produce some sort of representation (or collection of representations) of the objects visible in the environment, including information about distances to their surfaces, orientations of visible surfaces, illumination, and a variety of other geometrical relationships.

He pointed out that there is a completely different function for perception in general, and vision in particular, which he called perception of affordances. In that sort of perception, the information acquired is not about the actual contents of the environment in a form that is independent of the perceiver’s capabilities and interests, but is information relevant to potential actions of the perceiver: e.g. what actions are possible for the perceiver in the current environment, given the perceiver’s physical capabilities and current needs (or goals, preferences, etc.). Actions produce changes, so the perceived information is about possible changes the perceiver can bring about. We’ll generalise that below.

A more extreme version of Gibson’s view, treats the information content derived from visual sensory information as heavily dependent on the viewer’s anatomy and physiology, and current or possible needs, preferences, dislikes, etc.

An even more extreme theory could claim that there is no explicitly describable spatial content, only an unintelligible mass of conditional causal connections between sensor neurones and motor neurones, modulated by signals from internal sensors concerned with the organism’s current needs. I think some researchers who emphasise embodied cognition and who deny the use of representations, are implicitly committed to such an extreme position. But that view will be ignored here. (I expect any experienced engineer will easily see its flaws.) Most of this paper presents and analyses cases of perception of what is and is not possible. Some parts discuss implications for mathematical cognition.

**Beyond Gibson: Possibilities, impossibilities, necessities**

Most models and theories of perception seem to restrict the functions of perception, including spatial perception, to detecting and recording how things are (or probably are) in the environment, classifying them and predicting what will happen or what will be the case (with or without probability estimates), and in some cases retrodicting or diagnosing causes of, or predecessors of, what is the case.

For more intelligent species, perception, and especially vision, can also be used for acquiring information about what is and is not possible: i.e. modal information. So theories, models, and robotic implementations of vision systems that ignore perception of and reasoning about possibilities, impossibilities and necessities are seriously impoverished. That criticism of standard theories of vision is closely related to Gibson’s criticisms, but he does not go far enough in the direction proposed below (elaborating on Sloman (2009a)).

So, vision (and to a lesser extent other modes of perception) can be used not only to gain information about what is the case in the environment, but also information about possibilities, and relations between possibilities. This generalises Gibson’s ideas about perception of positive and negative “affordances” for the agent. In particular possibilities and constraints on possibilities that are perceived visually need not concern actions or needs or preferences of the perceiver; the visually acquired information can go not only beyond spatial structures, and immediately useful
information about possible actions for the perceiver, but can also include future possibilities, explanations of previously realised possibilities or failures, and also discovering possible or impossible events or processes that have nothing to do with the perceiver’s intentions, plans, or actions.

The need to generalise Gibson’s ideas
Most "Gibsonian" theories of perception (especially visual perception) that I am aware of fail to do justice to the variety of functions of vision, the variety of types of contents of visual experience, and consequently the variety of requirements for explanatory mechanisms, or mechanisms needed to give robots human-like (or even squirrel-like, crow-like, etc.) visual or more generally spatial perceptual capabilities. This is also true of theories of intelligence or cognition that (over-)emphasise embodiment. Focusing on too few examples of what needs to be explained leads to bad theories in both science and philosophy. It can also lead to impoverished engineering.

In particular, theorists emphasising embodiment often ignore the distinction between "online" and "offline" uses of visual information, discussed below, and the more subtle division between different "offline" uses of perception of what is the case, including perception of what is and is not possible and how those possibilities and impossibilities can change if some current possibility is realised. Processes of predicting planning, designing and explaining may all use chains of alterations in what is and is not possible.

Note added 22 Feb 2020
I previously failed to make clear that the possibilities and impossibilities mentioned here are not possible and impossible sensory or perceptual contents (in the minds of perceivers) but possible and impossible spatial contents in the physical environment --- including some that are relevant to choices in engineering or architectural design, for example.

Moreover, some of the perceptually available kinds of information about possibilities and impossibilities illustrated below are also essential to some ancient mathematical discoveries, for example in geometry and topology, recorded in Euclid’s Elements, which I contrasted with logical discoveries in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/ijcai-2017-cog.html

I also include below some examples that were not part of ancient mathematics, but require similar abilities to discover spatial possibilities and impossibilities. One example that as far as I know has not previously been discussed is the necessary connection between two aspects of a process of triangle deformation:
deform-triangle.html http://www.cs.bham.ac.uk/research/projects/cogaff/misc/ deform-triangle.html

The connection between the functions of visual perception in humans and other animals, and mathematical discoveries made by Euclid and his predecessors is the main topic driving the construction of this paper, but the connections involve somewhat long and tortuous links.

In other documents I’ll focus on some visual capabilities of mathematicians: not modern mathematicians reading logical and algebraic formulae and proofs (requiring a related, but different, set of competences), but the ancient mathematicians whose discoveries I suspect led, eventually, to Euclid’s Elements. As far as I know, these abilities are still unexplained and have not been replicated in AI systems.
A full investigation will require understanding how evolution of biological functions of human vision, including visual competences shared with other species, led to capabilities that were able to support mathematical discoveries, even though that was not why they evolved.

Some of those evolutionary changes seem to be recapitulated in child development, as described in the Meta-Configured genome theory. XXXX Understanding the details may be essential for high quality mathematics teaching, but that's a topic that will not be addressed here. (See the sections on Toddler Topology, and Toddler Theorems below, and the epigenetic schema in the section on Evo-Devo Issues.)

In my youth it was still customary to teach geometric mathematical competences at school, but most current youngsters (at least most young students and researchers I meet in Universities), seem to be deprived of that privilege.

Many potential readers of this document will therefore unfortunately have no prior experience of some of the phenomena under discussion. Links are provided to web pages presenting various more or less elementary fragments of Euclidean geometry and a subset of topology concerned with continuous deformations in space. I shall try to present examples that are intelligible to non-mathematicians, all of whom have mathematical competences, whether recognized or not.

The work presented here implicitly presents requirements for some of the construction kits that build human visual systems. We need open minds as to whether well-known forms of computation and physical assembly suffice.

Online and offline uses of visual information

A crucial first step in understanding the connections between vision and mathematics, and the roles of visual mechanisms (in contrast with separate cognitive mechanisms) in making mathematical discoveries, is to distinguish uses of perceptual information in (a) online intelligence and (b) offline intelligence:

(a) Online information about affordances is used immediately in triggering new behaviours or modifying existing behaviours (e.g. blinking reflexes, swerving to avoid something, changing direction while chasing something, closing a fist around something seen to be graspable).

(b) Offline information about affordances is used in considering possibilities, comparing possibilities, understanding relationships between possibilities, selecting possibilities to be achieved at some later time, or deciding between alternative possibilities that could explain past events or states. More sophisticated cases involve use of information about impossibilities. (Examples are given below, and in Sloman(2007-2014).)

This is not necessarily a sharp dichotomy: there may be processes/activities that use both online and offline functions of vision, sometimes in succession and sometimes in combination. However, in the extreme cases the types of information-processing mechanism required are very different, even if intermediate cases arise from use of both types of mechanism in combined tasks.
I suspect some of the enthusiasm for "embodied cognition" and "extended mind" theories is based partly on recognition of the importance of online intelligence coupled with blindness concerning offline intelligence, and partly on ill-founded anti-computational prejudices in some cases. But I shall not pursue those points here.

**Note on the online/offline distinction**

I have recently learnt that other writers use the online/offline distinction in partly related ways. (I may have picked it up from one of them.)

I think I first encountered the phrase "online intelligence" in a talk by Karen Adolph in 2007. But the online/offline distinction is closely related to the distinction between "reactive" and "deliberative" sub-systems familiar in AI long before that, and much used in the CogAff Project: http://www.cs.bham.ac.uk/research/projects/cogaff/#overview

In The Computer Revolution in Philosophy (1978), Chapter 6 used the labels "executive" and "deliberative" for a related distinction: http://www.cs.bham.ac.uk/research/projects/cogaff/crp/#6.11.

Sloman(1983) makes closely related distinctions using different terminology, e.g. comparing the (online) use of vision to control painting the edge of a table, or to guide a familiar grasping action, with more descriptive (offline) uses focused on by AI researchers. Offline perceptual intelligence involves using perception to acquire and (at least temporarily) store information in a form that can be used immediately or later for various purposes, including novel purposes, such as working out how to use an unfamiliar tool. (People who fail to understand this sometimes talk of "where" vs "what" perceptual functions, or "unconscious" vs "conscious" perception, e.g. because they have not learnt to think like designers of working machines.)

These distinctions were elaborated following a discussion with Dean Petters, in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/fully-deliberative.html

Many of the details are ignored here, though they should all be seen as part of a larger investigation linking modes of representation, types of perception, modes of reasoning, and modes of learning and discovery.

**Note on the irrelevance of "possible world" semantics**

There are many philosophers who have worked on an idea (with a long history, but sharpened in the last quarter century or so by philosophers like David Lewis and Saul Kripke, among many others), namely the notion that our ideas of possibility and necessity depend on a prior idea of a set of possible worlds. (It is not usually expressed so baldly.) I think that analysis is completely misguided, and that the ideas of Gibson about the possibilities for change in particular contexts considered by intelligent agents (including young children, and other intelligent animals) point to a deeper, more 'local', basis for modal concepts, allowing simpler versions to be used by other intelligent species and pre-verbal children.

Instead of possible whole worlds we use possible alternative fragments of the world, usually restricted to an accessible part of space time, though one aspect of cognitive development is increasing ability to consider larger extensions, in space and time (past and present).
Ultimately this will relate to the combinatorial powers supported by physics (including the structure of space-time) and chemistry. But that is a topic for another discussion. Some of the ideas were presented in my DPhil thesis in 1962, and in Sloman, (1996), which introduced the idea of physical objects or mechanisms being "possibility transducers". (E.g. possible voltages applied to a fixed resistor are associated with possible currents: [[Add note on how this connects with John Barnden’s ATT-META mechanism.]]

**Requirements for the online and offline functions of vision**

What sorts of visual mechanisms give organisms online and offline forms of intelligence?

(a) Online use of visual information requires fast-acting information stores (memory mechanisms) whose contents constantly influence forms of behaviour, and which are constantly overwritten as new information comes in, so that any use of the information has to be fast.

In many (most?) cases the mechanisms using such information are fast-acting (i.e reflexes) and either innate or produced by extended learning or training, e.g. in many sporting activities, musical competences, linguistic competences, and others. Some may use evolutionarily very old mechanisms (e.g. blinking), others newer, more sophisticated, mechanisms (e.g. musical sight-reading).

(b) Offline use of visual information requires longer term forms of storage, so that information acquired at a particular time can be used at different times, for multiple purposes, usually in combination with other forms of information, new and old, often on the basis of temporarily assembled structures -- using what are often referred to as "deliberative" mechanisms, discussed in more detail in Sloman (delib).

**NOTE:**

Among some psychologists, neuroscientists and even philosophers, a failure to understand this distinction has led to deep muddles about "What" vs "Where" visual processing pathways in brains. Both online and offline visual processing can include identification/categorisation mechanisms ("what") and inferences about location ("where"). And each of those two can use the other. I’ve never understood how anyone took the What/Where idea seriously.

See Sloman (1982).

Another common muddle seems to involve the assumption that online uses of visual information to initiate or control action are somehow incompatible with the use of the same information to provide content for visual consciousness, so that the process cannot be reflected on, talked about, evaluated, etc. ([REFS needed -- E.g. Milner and Goodale ????]). This assumption both underrates the sophistication of some of the engineering designs produced by biological evolution and also underrates what might one day be achieved by robot designers -- if it has not already been achieved in robot visual learning mechanisms, that use repeated trial and error learning to "re-shape" control algorithms.

In the case of many human online skills, e.g. in athletics, playing a musical instrument, painting pictures, and many craft skills, apprentices depend on the ability of experts not only to perform skilfully in reactive mode but also to be aware of what’s going on and use that information to help learners.
I don’t know of any AI robot that can learn and teach in this way, but in simple cases it should be feasible soon.

### Biological examples of offline use of vision

An example of offline use is an animal seeing some fruit in a tree and being motivated to climb the tree to get to the fruit. The fruit need not, and typically will not, remain visible to the animal during the process. A more complex use would be an animal recording information about the location of the fruit and not using the information until later, when it is hungry. Stored information in combination with the need for food can trigger a process of planning a route back to the tree, and use of the route to get to where the fruit was. This can happen even if the fruit is no longer there, because it has fallen to the ground, or been consumed by something else.

The ability to do use information offline, in forming and executing multi-step plans is often thought to be restricted to a small subset of vertebrates, but there is evidence of such abilities in other species, including the Portia Spider.

**Planning and deliberation by portia spiders**

The portia spider works out a route to its prey then follows it even when it can no longer see the prey, making detours if necessary and avoiding branches that would not lead to the prey.

“By visual inspection, they can select, before setting out, which detour routes do and do not lead to prey, and successfully perform a detour with no further visual contact with the prey”.

M. Tarsitano, 2006,

> Route selection by a jumping spider (Portia labiata) during the locomotory phase of a detour, *Animal Behaviour*, 72, Issue 6, pp. 1437--1442, [http://dx.doi.org/10.1016/j.anbehav.2006.05.007](http://dx.doi.org/10.1016/j.anbehav.2006.05.007)
>
> See also: [https://en.wikipedia.org/wiki/Portia_%28genus%29](https://en.wikipedia.org/wiki/Portia_%28genus%29) and [https://en.wikipedia.org/wiki/Portia_fimbriata#Hunting_and_feeding](https://en.wikipedia.org/wiki/Portia_fimbriata#Hunting_and_feeding)

Another kind of offline use involves passing information to another agent: e.g. pointing at where the fruit is, or telling someone where it is, or explaining how to get to it. (These are three different cases.) It is easy to think of other cases of offline use of perceptual information: left as an exercise for readers.

**Proto-affordances**

Gibson’s idea that the main function of vision is to provide information about affordances, can be further generalised to include the role of vision not only in acquiring information about possible actions of the perceiver (used in either online or offline intelligence), but also information about possible changes in the environment, and constraints on those changes, irrespective of whether the changes are produced by the perceiver, and irrespective of whether the changes are known to be relevant to the current or future needs or interests of the perceiver. An example would be noticing the possibility of the fruit falling and hitting a branch below it. Every physical configuration of objects has multiple possibilities of change that can be understood by perceivers who have no interest in whether the changes occur or not. I call those "proto-affordances". See (e.g. Hartson)
Examples of scenes with multiple proto-affordances are presented below.

Representing possible processes

Information about actual or possible processes can take many forms. At one extreme all processes are represented in terms of rates of change of measurable quantities, or vectors (usually represented by coordinates). If the information available is not sufficiently precise to provide numerical values an alternative is to encode possibilities either in terms of enclosing intervals (perhaps with fuzzy boundaries) or by probability distributions over possible values, or by ... (need to enlarge list of options).

A quite different approach to representing possibilities is to switch to topological, or more generally relational and structural descriptions. Such descriptions can specify parts, and relationships between parts, possibly parametrised relationships; e.g. A and B meet at an angle that is smaller than the angle between C and D, or the distance between A and B is less than the distance between C and D, or A, B, C, and D are parallel with gaps of increasing size along the sequence, or the vertex at which A and B meet lies on C, and many more.

The availability of such non-numerical representations of structure can make it inappropriate (and wasteful) to use the most powerful mathematical methods for representing processes, e.g. using differential and integral calculus. It can be especially wasteful and inappropriate if the precision of those methods is overkill for the information needs of an animal (e.g. answering "Am I getting closer to my prey?"). Use of qualitative, topological, comparative, imprecise descriptions may allow far greater generality at the cost of some ad-hocery: e.g. learning about special important cases (structures) and describing others in terms of them.

The use of grammars and parse trees (or nets) in linguistics and in compiler design illustrates the power and versatility of non-numerical forms of representation: for some purposes. In the 1960s, Clowes (see Clowes Tribute) and others proposed similar techniques for visual systems, including claiming that, like sentences, pictures/images could have (two-dimensional) syntactic structures representing (in many cases three-dimensional) contents, though finding appropriate generalisations of the notions of "grammar" and "semantic content" was not easy Kanef(1970). Clowes and others pointed out that pictures, like sentences, could have inconsistent semantics.

Chapter 9 of Sloman (1978) provided a demonstration of multi-layer semantics, as a proof of principle, showing how images, like spoken, written and signed sentences, can be interpreted as having several distinct levels of structured (syntactic/semantic) content, detected in messy images by the POPEYE program. (At that time AI theories of vision as requiring a mixture of bottom up and top down -- and middle out-- concurrent processing were unfashionable. An application for funds to continue the research was refused, and a paper reporting the work proved unpublishable, partly because, as the main reviewer pointed out, it was inconsistent with David Marr’s ideas regarding vision, which I thought then, and still think, were oversimplified. (His admirers were more narrow minded than he was!)
Such multi-layered semantic content is related to the use of multi-layered genomes, in which complex structures (usually developed later in evolution) are expressed by the genome at later stages of individual development, but in ways that make use of the structures/competences produced at earlier stages, which can vary according to physical and cultural environments. (This is how a particular genome has the ability to produce a very wide variety of types of mind, using information acquired from the environment during development. This is totally unlike any standard learning mechanism fixed at an early stage, and producing increasingly complex products, by repeatedly using statistical regularities found earlier.)

Possibility perception

To recapitulate: besides the processes actually occurring in any situation, other processes are possible, and if they were to occur the results would include perception of processes that were not perceived previously. One thing that should be obvious, but apparently is not obvious to many vision researchers is that perception of processes is not a matter of finding depictions of moving boxes in an image and tracking them. Watching a ballet performance in which there are not only groups of dancers performing on the stage, but also other animals, furniture, perhaps books and toys, provides opportunities for very rich visual processing, in some ways (but not all ways) more complex than videos of varied vegetable matter.

Could challenges for online intelligence lead to a new kind of offline intelligence

Is it possible that mechanisms that originally evolved to serve fast-acting behavioural reflexes, may later have been modified to serve fast-acting mechanisms for building new temporary internal information structures triggered by the contents of fast-acting sensory buffers developed for online intelligence (type (a) above)?

During speech understanding these extended online control mechanisms would construct intermediate information structures representing phonemes, morphemes, words, phrases, clauses, sentences and other linguistic entities.

But before that, evolution seems to have have produced older visual mechanisms shared across more species: mechanisms that that rapidly construct a variety of information structures about aspects of the environment, including parts of familiar objects, combinations of familiar objects, possible action trajectories, possible consequences of such actions, some of which are necessary consequences, constraints on such possibilities (i.e. impossibilities), possible non-action processes, and combinations of all the above to form information structures about a complex situation, structure or process.

All of these modal features (types of possibility and impossibility) perceived, are distinct from, though dependent on, perception of actual structures and processes.

Many contents of such perceptions are not used at the time for any practical purposes, but some may be used as a source of inferences for decision making later on, e.g. perceiving a person eating a complex object, might be a source of information relevant to the perceiver assembling ingredients for making a sandwich, locations where ingredients might be stored, and various partially
assembled sandwich states of changing complexity. For some of the entities thus recorded at high speed, e.g. various objects, spaces, spatial relations, and motions, the relevance to possible future moves by the viewer may also be derived and recorded. That could be the birth of certain kinds of affordance perception discussed by Gibson.

I suggest that many such perceived situations and processes involve unavoidable constraints some of which were noticed and thought about by long dead ancestors of humans, using ancient possibility/impossibility perception mechanisms -- laying foundations for later evolution and development of sophisticated mathematical discovery and reasoning mechanisms.

If perception of a static scene can trigger rapid construction on varying spatial scales and temporal scales, with varying combinations of concreteness and abstractness, then perception of a complex moving scene, or a complex static scene perceived by a moving viewer will require mechanisms that can rapidly modify the information structures, driven by information about changes in receptor information contents in combination with other information, including information about the viewer’s actions, and additional background knowledge about the type of environment.

If all that apparatus for motion perception is already available to deal with a wide variety of types of motion, whether motion of the viewer, or motion of perceived objects or both, then perhaps the same apparatus can also play a role in a new kind of perception of static scenes, by implicitly representing widely varying possibilities that cover things that could happen in such a situation.

If the mechanisms for abstraction are available for dealing economically with actual processes they may also allow representation and reasoning about possible processes.

These generalisations of Gibson’s ideas seem to be crucial for understanding mathematical cognition in humans, other animals and possible future robots. That’s because key forms of mathematical discovery are concerned with what is possible and what is impossible, and how the set of possibilities and impossibilities relevant to a situation can change if some of the possibilities are realised. Your possibilities for action and perception outside a doorway are different from the possibilities just inside the doorway.

**Possibilities vs Probabilities**

The possibilities and impossibilities mentioned above have nothing to do with information about probabilities. Possibilities are more fundamental than probabilities: probabilities can exist only in relation to sets of possibilities, but sets of possibilities, e.g. possible actions that an animal can perform in some situation, or possible ways in which three coins could be arranged on a table, have nothing to do with probabilities. The possibilities can be noticed, reasoned about, and in some cases made actual, without any consideration of probabilities. If you know what coins are and you see four coins on the table there are many possibilities for changing the configuration on the table including adding more coins, removing coins turning coins over, rotating the coins without turning them over, and moving them into new locations.

**Temporally chained possibilities**

A great deal of attention has given to probabilities of various things happening when one or more coins are flipped, or one coin is flipped several times. Here’s one of many online tutorials:

http://gwydir.demon.co.uk/jo/probability/info.htm
What has not been noticed is that there are also many impossibilities associated with coins -- not when they are flipped but when they are turned over in a controlled way. Repeatedly turning over only one coin gives a very boring predictable result. If there are N coins, with N > 1, and only one coin is turned over at a time, then the resulting sequences of patterns can have very different structures. E.g. here are five sequences of patterns formed by three coins, using H or T to show whether each coin is head up or tail up.

Pattern Elements:

|   | 1  | 2  | 3  | 4  | 5  | 6  | 7  | ...
|---|----|----|----|----|----|----|----|---|
| P1 | [HHH] | [HHT] | [HTH] | [THH] | [HHT] | [THH] | [HHH] | ...
| P2 | [HHH] | [HHT] | [HTH] | [THH] | [HHH] | [HHH] | [TTT] | ...
| P3 | [TTT] | [HHT] | [HTH] | [THH] | [HHH] | [HHH] | [TTT] | ...
| P4 | [HTT] | [HHT] | [HHH] | [THH] | [THT] | [TTT] | [TTH] | ...
| P5 | [HTT] | [HHT] | [HTH] | [THH] | [THT] | [TTT] | [TTH] | ...

Four of the sequences shown are impossible. One of the sequences is possible. Which?

(Correction 6 Jan 2019: Previously the description erroneously said four are possible and one impossible. My thanks to Marcel Kvassay for pointing out the slip.)

Now consider what happens if you combine a pattern of controlled coin turning while counting, using only a single coin?

There's this possible sequence:

|   | 1  | 2  | 3  | 4  | 5  | 6  | 7  | ...
|---|----|----|----|----|----|----|----|---|
|   | [H] | [T] | [H] | [T] | [H] | [T] | [H] | ...

and this:

|   | 1  | 2  | 3  | 4  | 5  | 6  | 7  | ...
|---|----|----|----|----|----|----|----|---|
|   | [T] | [H] | [T] | [H] | [T] | [H] | [T] | ...

If you think about a flat coin showing only either H or T, and you keep counting as you turn a coin, are any other combinations possible?

Now suppose you have two coins, and as you count you turn only one coin over, but sometimes you turn over the first coin and sometimes the second: i.e. there is not a regular alternation. Here are sequences combining counting (numbers) and pairs of coins

|   | 1  | 2  | 3  | 4  | 5  | 6  | ...
|---|----|----|----|----|----|----|---|
|   | [HH] | [HT] | [HH] | [TH] | [HT] | [TH] | ...
|   | 1  | 2  | 3  | 4  | 5  | 6  | ...
|   | [HH] | [HT] | [TT] | [TH] | [TT] | [TH] | ...
|   | 1  | 2  | 3  | 4  | 5  | 6  | ...
|   | [HH] | [HT] | [TH] | [HH] | [TH] | [TT] | ...

The number of possible sequences is far greater than if you have only one coin and turn it over for each number counted. If you have more than one coin, only one of which is turned over at a time, then you cannot work out in advance what coin pattern will be reached at the 11th step, since it will depend on the earlier sequence of choices.

Question: If you turn over exactly one coin at a time as you count, with no constraints on which coin, are there any constraints on the resulting possible sequences, e.g. any possible patterns of count number and states of the two coins that cannot occur? What about combinations of types of number, e.g. odd numbers, even numbers, multiples of 3, powers of 3, with coin-face
Some of the possibilities, if realised, will necessarily have certain consequences concerning what can happen next. For example if there are several coins on the table then each one will either be H up or T up. Then turning any of them over will switch to the other state. If you are counting, then each number will have certain mathematical properties, e.g. which other numbers divide into it with no remainder (i.e. what its factors are), and when you move to the next number the properties (the list of factors) will change. Can you find some combinations of coin states nd count numbers that cannot occur in such a sequence?

Apologies to number theorists who find this too easy. My point is not to display deep new mathematical results, but to raise questions about what sort of mind a machine, e.g. a future robot, needs to have in order to make discoveries about which static combinations are impossible, or which sequences of combinations are impossible.

More importantly, are the same kinds of mechanisms required for discovering the coin-plus-counting impossibilities as the other possibilities presented in this document, e.g. geometrical impossibilities and topological impossibilities. In other words, could there have been a stage in our evolutionary history (or in the development of an individual human) before which none of these impossibilities could be discovered followed by evolutionary change to a new state that allowed our ancestors to discover all of these impossibilities?

Or were different evolutionary transitions in information processing required to enable the different impossibility discoveries to be made?

One of the simplest examples is the fact that if in the two coin example, two turning-over processes occur in sequence: involving either only one coin turned twice or first one coin turned then the other, then if they initially had the same state they will end up with the same state, and if they initially had different states they will end up with different states -- after two such turns. Why must one of those be the result? How can we design machines that are able to answer these questions? (What sorts of machines could think of these questions without being tested by humans? What makes you think of new questions? What mechanisms allow you to respond to such triggers?)

Can these discoveries be made using statistical/probabilistic reasoning?
My experience of talking to many researchers about examples presented in this document is that often they assume that a robot could make these discoveries by experimenting with lots of coin turning sequences and making statistical inferences. And some psychologists or neuroscientists will assume that brains make the discoveries by using such reasoning.

But it is impossible to use statistical evidence to discover that something is impossible, or that something is necessarily the case. Our ancestors may have used empirical evidence to conclude that it is impossible for a person in Paris and a person in Toronto to have a normal conversation, at the same speed as a face to face conversation. But we now know that that conclusion would have been false.

To make the mathematical discoveries about possibilities, impossibilities and necessary connections discussed in this document (and many others) no amount of empirical evidence will suffice, though it may be suggestive.
As Immanuel Kant pointed out over two centuries ago, the discovery of necessity or impossibility requires something different: something that mathematicians have been doing for millennia, and I believe other humans and many other intelligent animals who were not explicitly doing mathematics were also able to make and use such discoveries, but without the meta-cognitive abilities to notice and think about their reasoning.

An example is the 3D topological reasoning about possible 3D trajectories in this 17.5month (pre-verbal) child: http://www.cs.bham.ac.uk/research/projects/cogaff/movies/ijcai-17/small-pencil-vid.webm

Of course, discovering examples can provide knowledge of probabilities, and that can in some cases trigger investigation leading to a deeper insight into what is impossible or necessarily the case, in addition to discoveries about probabilities.

For example if you have a collection of coins and a chess board, and place one coin in each square on the board randomly or systematically placing them with head or tail up, you can do experiments to find out the probability of all four corner locations being occupied by coins in the same state, e.g. all H or all T. However, using mathematical, not observational reasoning, we can find the probability without sampling the space of possible configurations, which is just as well because getting a fair subset of the space would take a long time, since there are 18,446,744,073,709,551,616 possibilities. With that size of space, no human could ever collect adequate statistical evidence to settle the probability of all for corners having the same colour.

Isn’t it fortunate that evolution produced brain mechanisms that are far more powerful than probabilistic learners? What mechanisms? When will robots have them? Do some AI machines already have them? Does anyone know how brains do such things?

Perceiving vs understanding

It is possible to perceive convincing evidence of a regularity without understanding why the regularity holds. For example, Morley’s amazing theorem states “The three points of intersection of the adjacent trisectors of the angles of any triangle form an equilateral triangle.” This wonderful web site http://www.cut-the-knot.org/triangle/Morley/ includes a Javascript applet that allows the corners of a triangle to be moved arbitrarily. When that happens, the applet automatically adjusts the trisectors of the angles and the points of intersection of adjacent trisectors. Playing with this demonstration may convince someone that theorem is true, or that it has a very high probability, but that’s not the same as understanding why it must be true, or seeing that there cannot be any counter examples. Without that insight you cannot be sure that Morley’s claim isn’t merely a very close approximation to the truth: e.g. if the angles of the central triangle remain very close to 60° while changing imperceptibly. A person who has not grasped a proof lacks the "mathematical qualia" for that theorem.

Contrast Mary Pardoe’s proof of the Triangle Sum Theorem, in the form: the interior angles of any planar triangle must sum to a half rotation (180°):
In a plane surface, rotating the blue arrow through the three internal angles (i.e. A, then B, then C) always brings it back to the starting line, pointing in the reverse direction, without ever crossing over its original orientation, and this (obviously?) doesn't depend on the shape of the triangle.

The above example illustrates ways in which possibilities and necessities or impossibilities can be closely related: realisation of some possibilities may necessarily have certain consequences. What they are and why they are inevitable differs from case to case. The ability to notice possibilities and impossibilities (necessities) and the consequences of realising some possibilities in a situation is an important aspect of human development, as Piaget noticed. His last two books (1981-1983) discuss many examples.

However, although Piaget realised that these are important aspects of human cognition, and some of ways of probing children’s minds are based on deep insights about varieties of cognitive function, I am not sure that his theories regarding the cognitive mechanisms (which I found hard to follow) were sufficiently well developed to be useful, e.g. in explaining mathematical cognition, or in designing intelligent machines with human-like powers of mathematical discovery. Piaget’s work on cardinality is mentioned below below..

One of the problems of discussing such issues is that there are so many different types of case, and we need to understand the variety in order to come up with good theories about what’s going on. In particular there are some cases where the cognitive competences involved are purely logical reasoning capabilities, whereas in other cases more varied mathematical abilities are required, e.g. concerned with reasoning about spatial structures and processes, as in topology and Euclidean geometry.

Pictures of possible and impossible object configurations

In order to illustrate some of these points I shall discuss some examples of pictures of impossible objects, for which the mathematical physicist Roger Penrose and the artist Maurits Cornelis Escher are famous. Less well known is the Swedish artist Oscar Reutersvard, whose pictures of impossible objects came earlier (1934), and have some interesting features more directly relevant to mathematical reasoning about possibilities and impossibilities, because the pictures invite changes of various sorts, or, in other words, display multiple affordances. The pictures also help to shed light on requirements for artificial vision systems, and theories of vision. Indirectly they also challenge theories about what brains can do and how they work. (There are also some well known paintings of impossible scenes by much older artists.)
Many of Escher’s pictures have far more complex examples of this, e.g. his Waterfall picture (https://en.wikipedia.org/wiki/Waterfall_(M._C._Escher))

Instead of jumping straight to impossible objects, it is illuminating to consider pictures of scenes containing items in different arrangements, some possible, and some not and try to understand what exactly changes, both in the scenes and in the cognitive/perceptual processes.

What can you do with a collection of eight similar cubes?

**Fig: 8 cubes (A)**

8 Blocks in space: how else could they be arranged?

Alternative configurations of the blocks are presented below. You can probably imagine a series of individual block-trajectories that would transform the above figure into the one below, and similarly for (most of) the later examples.

**Fig: 8 cubes (B, C)**
Two more possible configurations of eight blocks.

AI vision systems in the early 1970s could already (very, very slowly, because of computer speeds at the time) interpret a variety of 2D line drawings as representing 3D polyhedral objects, including deciding which picture lines corresponded to concave, convex, or occluding edges, e.g. Clowes (1971). A short time later shadows and cracks were added, and then the ability to cope (to some extent) with noisy images containing spurious line fragments, missing line fragments or lines with gaps due to image noise (for example). (Ref G.Grape, G. Hinton). A tutorial presentation on some of those techniques is referred to below.

However such programs did not have the ability to suggest, or reason about, alternative configurations of blocks: they saw only what existed. Moreover they were not able assign precise lengths or angles in all the images in which they could perceive structure. And although they could in some cases detect that one visible surface must be further from the viewer, they did not reason about whether the whole scene depicted was consistent. So they could not detect circular “further than” relationships, though I suspect that could have been added. But that’s just a special case of detecting impossibilities.

Thinking about, or imagining possible variations in a scene is a crucial ability for many intelligent animals, including nest-builders, hunters, and animals that care for their young.

Humans can do this not only for real scenes containing physical objects but also for depicted scenes: where the pictures specify physical objects in physical relationships, including relationships like adjacency, co-linearity, being above, being between, or supporting.

They presumably cannot do all this at birth. Why not? What mechanisms do they lack? How do they acquire the mechanisms that provide the new abilities later on? Is it merely a process of learning to use mechanisms they already have from birth? Or from before birth -- e.g. from month X of foetal development?), or do new brain mechanisms grow during years of physical growth (and thereafter)?

Fig: 8 cubes (D, E, F)
Here are three more possible configurations with 8 blocks in each. Ignore the problem of gravity for now: the blocks could be held in those locations or they might be in a location with zero gravity or the "suspended" blocks could be lowered to the surfaces below them. You can probably imagine several different sets of trajectories of individual blocks that would produce each of the three new scenes.

**What if you had nine similar cubes?**

Try testing your own brain mechanisms for imagining configurations in which there are not only eight, but nine cubes arranged in space? Here are some sample configurations. Each could be transformed into any of the others by moving individual blocks around.

Here are nine blocks on a surface, shown in three possible configurations, including one in which one of the blocks is suspended above (or floats above) another block. How many other configurations are possible? How else could they be arranged? What sequences of block moves could produce the new arrangements?

Note: Piaget asked children that sort of question using a few objects on a flat surface (1981). Not all realised that there is no fixed finite set of possibilities. Some answered after a while that no more arrangements were possible.

**More possible moves**

Seeing possibilities extends beyond seeing possibilities for re-arranging objects in the scene. In this case it also includes seeing locations where you might put a flat, or nearly flat object (e.g. your hand opened out) into the scene depicted. E.g. for any pair of adjacent cubes you could move your hand between them, provided that you rotated the hand into the right orientation.

Other things you can do in the perceived configurations include swapping pairs of blocks: move one onto the table, move another block to the newly emptied location, then move the first block into the new space. Or move both simultaneously using two hands.

Several more configurations of nine blocks are depicted below. You may or may not find one of them anomalous.
Yet more possible configurations of 9 blocks. Are they all really possible? See text for discussion.
(Inspired by Reutersvard’s 1934 drawing.)

As before you can visualise ways of rearranging the blocks or moving your hand between the blocks, as described above in Section Possible moves. Look closely at the differences between the last two configurations. The left and middle pictures (J and K) depict perfectly possible 3-D configurations of cubes (though in a normal gravitational field something would be required to hold in place the cubes that are not resting on the table or on other cubes).

But there are subtle 2-D features of the rightmost picture L that indicate that the 3-D configuration that it represents, if interpreted as a picture of 9 cubes, involves a collection of pair-wise relationships between the cubes that are all possible in isolation, but not all possible in the same 3-D configuration. This impossibility does not arise out of any mis-use of pictorial conventions. The image uses only examples of image fragments that occur in other pictures of configurations that are perfectly possible.

The fact that the scene is experienced as impossible only if all the blocks are included challenges theories about limitations of numbers of objects that can be attended to simultaneously.

Examining the image L you should be able to imagine ways of removing one block that would leave the object depicted impossible, and also ways of removing other individual blocks that would render the scene perfectly possible.

When a 3-D scene depicted is geometrically impossible there need be nothing impossible about the configuration of lines in the picture. The impossibility concerns which 3-D structure, if any, the picture depicts if all the parts are interpreted normally as depictions of 3-D structures and relationships.
Another view of the transition from part of Figure J to Figure L above. The image above left depicts a possible 3-D scene. Modifying it as on the right produces a picture that, if interpreted using the same semantic principles, represents an impossible 3-D scene, where blocks A, B, C form a horizontal line, blocks F, G, H form a vertical line, D and E are between and on the same level as C and F, and the new block X is co-linear with A, B, and C, and also with F, G, and H -- impossibly! Notice how the relationship between A and H has changed.

The drawing on the right (minus labels) was by Swedish artist, Oscar Reutersvard, in 1934

http://im-possible.info/english/articles/triangle/triangle.html
http://butdoesitfloat.com/A-father-to-impossible-figures

Compare the above two pictures. A complex picture made of parts representing possible 3-D configurations may have N parts such that if a certain part X is added (e.g. a picture of an extra block that is simultaneously co-linear with two other linear groups, as in the above figure on the right), then it becomes anomalous and cannot represent a 3-D configuration using the same rules of interpretation (based roughly on reversing projections from 3-D to 2-D). Notice that in this case, the addition of X required changes to the (2-D) depictions of blocks A and H that preserved the 3-D relationships between A and B, and between G and H, but altered the 3-D relationships A and H, depicting A as occluding H. That produces a contradiction even if block X is not depicted. If X were removed, more of H would be visible, but the impossibility would remain.

In other words the original N parts have a joint interpretation that entails that the situation depicted by adding the part X cannot exist, but not because of the addition of the new block, but because of a subtle change in the relationships between pre-existing blocks. If the blocks were depicted spread out more in space, so that they are not overlapping, this change would not be necessary,
but various relationships would become more ambiguous.

This is partly analogous to logical reasoning where \( N \) consistent propositions entail that an additional proposition \( X \) is false. So its negation can be inferred to be true.

This picture cannot be handled by the Huffman-Clowes line-labelling mechanism described in Clowes (1971) as it requires a richer grasp of geometry than the line-labelling provides (3-D relationships between adjoining or connected portions of an interpreted image). It requires an ontology of opaque 3-D objects with relationships between whole objects, not merely an ontology of edges, vertices and faces, with 2-D and 3-D relationships between them.

Humans can reason that the "obvious" interpretation of the 2D picture on the right represents a configuration that is impossible without knowing any of the actual distances or sizes, whereas I don’t believe any current AI vision system can do that, though it may not be very difficult to implement one to deal with the special case of opaque rectangular blocks in static scenes.

The detailed requirements for the richer ontology, if extended beyond 3-D objects bounded by plane surfaces, and beyond rigid objects (e.g. to include objects made of different "kinds of stuff"), and beyond static configurations, will vary for different species of animal, and for different developmental stages in the same species. As far as I know very little of this is in any current AI systems (or psychology, or neuroscience).

I see no reason to believe these capabilities could be acquired by any of the forms of learning currently fashionable in AI/Robotics. Much deeper epigenetic mechanisms are required e.g. as speculated in connection with Figure Evo-Devo, below.

This requires researchers themselves to develop deeper (meta-cognitive) ideas about forms of geometrical and topological perception and reasoning.
Moreover I don’t think neuroscientists have any idea how brains can support this kind of reasoning. (Please let me know if I am wrong!)

The above example is partly comparable to a collection of sentences, each of which describes a perfectly possible state of affairs, though their conjunction does not, e.g. "Tom is older than Dick", "Dick is older than Harry" and "Harry is older than Tom". Older than is a transitive relation, which means that "X is older than Y" and "Y is older than Z" implies "X is older than Z".

So the first two conjuncts above imply "Tom is older than Harry" but that contradicts the the third one because it is not possible for Tom to be older than Harry while Harry is older than Tom. ("Older than" is an anti-symmetric relation.) Why it is impossible, and how it is possible for an individual (human, other animal, or intelligent machine) to know that a relation is transitive and anti-symmetric, will not be discussed here.

The situations depicted in the pictures of blocks are more complicated than the linguistic example because there are several different relationships, including "further from the viewer", "higher than" or (further above the surface of the table), and "further along" in various directions in the scene, all of which are transitive and antisymmetric relations. It is left as an exercise for the reader to work out which 3-D relationships between blocks or between groups of blocks are depicted in the various pictures, and which combinations are inconsistent.

Aviv Keren drew my attention to a closely related paper by Roger Penrose (1992), in which the impossibility of a Penrose triangle is related to the mathematical concept of a "cohomology group". The ideas are introduced in terms of ratios of distances of objects along a viewing direction, which presupposes that distances have a metric (though I have not yet understood all the mathematical details of the paper).

In contrast with that approach, I have tried to show how familiar qualitative relationships of the form "further in direction D" that are transitive and antisymmetric can suffice to explain the perceived impossibility, without attributing to the viewer an understanding of coordinate geometry, or use of a metric for distance. (I suspect that the discussion by Penrose uses a special case of this.)

There are strong pictorial clues for partial occlusion. For example, a pictorial "T" junction is often used to indicate that a partly visible edge where two surfaces meet, represented by the stem of the "T", is occluded by another surface whose edge forms the crossbar of the "T" (a clue used since the 1960s by AI vision researchers).

Using transitivity and antisymmetry of "further" is easier in connection with the Reutersvard triangle than with the Penrose triangle (referred to as a "tribar" in his paper). The Reutersvard scene includes not only violation of antisymmetry of a collection of spatial relations, but also includes a large collection of affordances (in the sense of Gibson, discussed and illustrated above) concerning possible moves of the blocks and possible moves of other objects (e.g. a flat hand) in the spaces between the blocks that form an impossible collection. Details are left as a further exercise for the reader. (Or a future AI program demonstrating its spatial understanding!)
The pictures of impossible objects and sentences describing impossibilities both illustrate a deep and important point: if a form of representation (pictorial or linguistic) is to be suitable for expressing information about a rich domain of possible structures and processes, there may be no syntactic constraint on pictorial or verbal modes of composition, that is GUARANTEED to prevent self-contradiction AND provides the desired expressive power. In general discovery of features like transitivity and self contradiction can’t simply be based on definitions or logical forms. Other modes of discovery -- using topological or geometric insight -- are needed. As far as I know, there are no psychological, neural, or AI theories explaining such capabilities. I suspect we'll need to understand sub-neural chemical information processing, with its mixture of discrete and continuous processes, in order to fully understand how ancient brains discovered many examples of mathematical impossibility and necessity.

**Conjecture:** There is no drawing convention that allows all the possible configurations of nine blocks to be depicted, and which does not also allow the depiction of impossible scenes, like those shown above.

Compare: there is no generally useful human language that allows everything to be expressed that we might wish to express but prevents description of impossible configurations. Even the language of arithmetic allows us to formulate propositions that cannot be true, e.g.

**3 + 5 < 6**

Some mathematicians and programming language designers have attempted to design syntactic rules for languages that guarantee the impossibility of expressing something that is impossible. I believe that cannot be done for natural languages without intolerable restrictions in usefulness. Sloman(1971b) I believe the same can be said regarding languages for use in AI projects aiming to replicate human intelligence.

http://www.cs.bham.ac.uk/research/projects/cogaff/crp/crp/#postscript

The point can be expressed by the slogan:

If a language has compositional semantics rich enough to meet the requirements of intelligent thinkers and agents in a world like ours, then the generative power of the language will include the ability to generate semantic contents that are inconsistent, including logical and semantic paradoxes, and in the cases of pictorial languages, pictures of impossible objects and scenes. (I first encountered this idea when Max Clowes introduced me to AI work on image interpretation/scene analysis in the late 1960s):

The ability to notice impossibilities is an aspect of intelligence. For example, an animal thinking about possible ways to move a physical object to get it from one location to another location should be able to detect, at least in some cases, that moving the object through a particular gap is impossible, because the object is too wide. The ability of a 2-D structure to depict an impossible 3-D structure, and the ability of (some) humans to detect the impossibility have been much studied. But I think many of the examples have subtle clues about functions and mechanisms of vision that have not been appreciated. This is particularly evident in the 1934 picture produced by Reutersvard, presented below.
In the above examples, I have tried to show how to build up to Reutersvard’s example gradually, in order to get an accurate account of the phenomenon, by locating the discovery of impossibility in a space of possible actions using vision along with a visualisation of a target configuration.

Although James Gibson did not, as far as I know, ever discuss such examples (or the others mentioned below), I hope it is clear that the discovery of such impossibilities could arise in the context of perceiving affordances in the environment, i.e. possibilities for change of spatial relationships, and using those perceived affordances to construct something. The formation of an affordance-based intention generally leads to either a plan or an exploratory process that eventually culminates in construction of the intended spatial configuration. But in special cases it is possible to discover that the intention cannot be fulfilled because of negative, obstructive, affordances, which in some cases cannot be overcome using greater strength, new materials, collaboration with helpers, etc. I suspect that such discoveries could, over hundreds or thousands of years, have led our ancestors to formulate mathematical theories about spatial structures including discovering proofs of the sort presented by Euclid (Elements).

But even if these historical conjectures are correct, that still leaves us with the problem of explaining how the impossibilities are understood: what sorts of cognitive mechanisms allow proofs of impossibility to be constructed? What forms did those proofs take? How are they related to the processes in the mind of a child, or a squirrel, or a crow, who not merely tries and fails in a task, but comes to understand the failure? At present I don’t believe there is anything in psychology, philosophy, neuroscience, or AI that provides a rigorous explanation. Part of my reason for collecting a large and varied set of examples is to build up requirements for explanatory mechanisms. Combined with evolutionary investigations and exploration of designs for intelligent robots, we may be able to come up with a good theory that can be implemented and tested. I don’t think anyone has such a theory at present.

BACK TO CONTENTS

Reutersvard 1934

The above examples were inspired by a picture originally drawn by Swedish artist Oscar Reutersvard in 1934.

Fig: Reutersvard 1934
While a student, he drew the star at the centre, then added more lines, ending up with his impossible configuration, several years before the pictures of Penrose and Escher.

An important feature of this picture is the number and variety of possibilities for change implicitly depicted: e.g. all the places where you could put your hand between two of the cubes, all the cubes that could be removed leaving a gap, producing a possible 3-D configuration, all the pairs of cubes that could be swapped, etc. So we have a very rich collection of imagined structures, relationships and possibilities for change, including cases where what is imagined is geometrically impossible. Compare the following description of a collection of numbers.

There are nine numbers, a, b, c, d, e, f, g, h, i, all positive.

\[
\begin{align*}
  a + b &= c \\
  d + e &= f \\
  c + f + h &= i \\
  i &< a
\end{align*}
\]

What conclusions can you draw from the first three equations? Could the three equations and the inequality all be true? How do you know?

Reutersvard went on to produce many variants of his idea and a selection of his pictures were used in Swedish stamps. More pictures by Reutersvard are available here:
http://www.step-hen.dk/reutersvard.htm
It is arguable that Escher's pictures of impossible objects or scenes were more subtle and creative, with their rich blends of geometric and biological forms, but that feature is not relevant to our current discussion:
http://www.mcescher.com/gallery/impossible-constructions/
Pre-20th Century pictures of impossible scenes

Although 20th Century examples are often treated as original, much older artists had discovered the possibility of producing drawings or paintings of impossible scenes, for example Hogarth’s Satire on False Perspective (1754) shown here:

![Hogarth's Satire on False Perspective](http://upload.wikimedia.org/wikipedia/commons/4/4c/Hogarth-satire-on-false-perspective-1753.jpg)

In this case not all of the impossibilities are geometrical.

http://en.wikipedia.org/wiki/Satire_on_False_Perspective

Mis-use of the word "illusion"

Pictures of impossible objects are sometimes erroneously described as "illusions". They are illusions for viewers who accept all the 3-D structures depicted and do not detect the global impossibility (e.g. young children, or adults looking at very complex examples of such pictures). But not everyone is deceived, which raises the question: what visual/cognitive mechanisms enable non-mathematicians to realise that a picture is of something impossible, so that they are not deceived and suffer no illusion, though they may be entertained, or even amazed by the pictures? However special cases can be illusions, like the next Figure.
In contrast, the example above is illusory, because the 3-D printed triangle gives the appearance of having an impossible 3-D structure that it does not have. It really is a 3-D object, seen from a special viewpoint, but it is not a triangle. It uses a technique first demonstrated by Richard Gregory, many years ago. For more information on the origin of this triangle see:


This youtube video demonstrates Gregory’s non-triangle:
http://www.youtube.com/watch?v=gcw1IIGSGMM

Dice example

https://www.youtube.com/watch?v=EvzhqdM2yM4
A Youtube video showing construction of an "impossible" triangle using dice. Watch the sleight of hand just after 1min9secs. Compare this version:
https://www.youtube.com/watch?v=sEsPG0qW_nA

Tangled bodies
Added: 15 Oct 2017

In 1971 Max Clowes gave a lecture entitled ‘Man the creative machine: A perspective from Artificial Intelligence research’ at the Institute of Contemporary Arts (ICA), in London. A chapter based on the lecture was later included in a book of ICA lectures Clowes(1972). One of his examples (below) used a "Still" picture from the movie Sunday Bloody Sunday, showing tangled male and female bodies. Finding which visible parts in the picture belong to which person is non trivial, illustrating one of his favourite slogans "Perception is controlled hallucination". (E.g. how many people are in the scene? Which hands belong to the man?) Below I have made two sketches of the scene, one of which corresponds to the movie "still" used in the lecture. The other has a part misplaced to an impossible location. Which picture depicts the impossible scene, (a) or (b) and why? Which hands belong to whom? How many noses can you see? Answering the above questions in relating to the (possible) scene depicted requires detecting impossibilities in erroneous interpretations -- filtering interpretations.
Consider the differences between the two figures. Everyone depicted is only partly visible. How many people are visible in Figure (a)? How many in Figure (b)? How do you know?

What brain mechanisms support the information processing that allows you to generate interpretations of fragments of the image, and also to detect and rule out combinations of fragments that represent impossible configurations? Do you have to be trained on many examples of tangled arms, legs and bodies? Or can generic spatial reasoning abilities be combined with structural knowledge about human bodies.

I suggest that the mechanisms involved in rejecting wrong interpretations of ambiguous views of spatial structures in novel scenes are the same sorts of mechanism as were used by ancient mathematicians, in making geometrical and topological discoveries. At this stage I suspect nobody knows how biological brains can do that sort of thing. I'll turn now to a set of different, though partly related, examples.

Impossible transitions on a grid

Consider a grid on which it is possible to slide square tiles horizontally or vertically on the surface of the grid, like the one below. Here is a problem that is likely to be found insultingly easy for anyone reading this, but is a gateway into some much deeper problems, not all of which will be presented here.
In the above configuration, is it possible to move the red tiles on the right, one by one, by sliding horizontally and vertically to the locations of the grey tiles on the left, so that each occupied square has exactly one grey tile and one red tile? If not why not? Would it make any difference if those red tiles had started in a different configuration, or if the grey tiles had been in a different configuration? Is it necessary to try all possible routes through which the tiles can be slid (subject to the horizontal/vertical constraint) in order to discover that the goal cannot be achieved? Remember this is supposed to be an easy question.

A slightly harder question: what cognitive capabilities would a child, or a robot require in order to answer the question.

An even harder question: what kinds of evolutionary transitions might have produced the capabilities required for this task? (Later we shall contrast physically constrained and rule-constrained versions of the same problem.) What sorts of naturally occurring situations might have this sort of structure? What kinds of brain mechanisms might have been required in our earliest ancestors capable of answering this sort of question?

Some brain mechanisms would allow the question to be answered only for small groups of tiles, e.g. two or three of each colour. Do your brain mechanisms have such a limit, and if not why not?

Could a robot acquire the required abilities by being trained on lots of examples? If not, how could it be given those abilities -- i.e. what sort of artificial brain could deal with such problems, without being restricted to particular sizes of grid and particular numbers or configurations of the red and grey tiles?

The next example is partly similar, and partly different. Similar questions can be asked about brain mechanisms required, and their evolutionary origins.
On the left grid (a) the coloured squares can slide horizontally or vertically but not diagonally.

On the right grid (b) the coloured rectangles can be rotated, and can slide between squares only in diagonal directions, not horizontally or vertically.

Subject to those constraints can the three grey items slide to the squares containing the three red items in each case (or vice versa)?

If not why not?

Some readers will recognize this as mathematically related to a very well known puzzle that at first looks totally unrelated. If you don’t recognize it follow [*this link. (There may be better explanations on the internet.)

What kind of information processing system can allow an animal or robot to think about what transformations between configurations are and are not possible?

How is this related to mathematical discoveries in geometry, topology and arithmetic?

Do the shapes of the objects matter to the solution?

How is this related to perception of affordances in everyday life?

By what means could those discovery abilities possibly have evolved? What sorts of transitions in functions of vision, or more generally functions of perception, might have led up to such competences?

Can these mechanisms and processes be implemented using current tools and ideas for programming intelligent robots?

Clearly newborn human infants cannot answer these questions about objects sliding around on grids, or similar questions using different spatial configurations. Why not? In what ways would their brains have to change, or be changed, in order to enable them to think about such problems and not only work out the answers but understand that the answers are necessarily correct, and could
not be different for grids and tiles made of different materials, or located at different altitudes above sea level or even on another planet?

How might Gibson’s theory of perception of affordances have to be revised to cope with these questions?

While thinking about the above problems did you consider the possibility of altering the grid of squares so that they have two colours, like a chess board, with diagonally adjacent squares the same colour and horizontally or vertically adjacent squares different colours. That transformation makes the answers to the second set of questions, with two types of grid, trivial to answer. See this link[*] (same link as above).

What kind of brain mechanism in a human or a robot allows that sort of solution to be discovered and used in a proof? Curiously, many highly intelligent humans who already have the required brain mechanisms and knowledge of chess boards don’t notice the relevance -- perhaps because they have too much potentially relevant knowledge.

For example, in my experience some expert mathematicians immediately notice that assigning a coordinate frame to the grid gives each square two coordinates and their sum (or difference) is either odd or even. So they try to find answers in terms of parity-preserving operations, using their knowledge of arithmetic and algebra. This leads to a mathematically acceptable solution, but non-mathematicians who merely notice the consequences of having two colours alternating horizontal and vertically, as on a chess board, can also find a mathematically acceptable proof without mentioning coordinates of squares or division by 2, etc. They are simply using a different sort of mathematics, closer to the reasoning in Euclid’s *Elements*.

What sort of brain mechanism is required to enable a person presented with a problem to notice the relevance of apparently unrelated knowledge? [*]

*The Mutilated chessboard problem*
https://en.wikipedia.org/wiki/Mutilated_chessboard_problem

Two relevant books:


Max Wertheimer, *Productive thinking*

---

**Impossible number relationships**

Returning to our blocks on a table scenario, we can consider ways of forming neat rectangular arrays of blocks using some or all of the blocks. The figure below shows what can be done with all nine blocks, and what happens if you remove first one block then another. How do you know that after removing the second block the remaining blocks cannot be arranged in a regular NxM 2-D array?
Possible and impossible array configurations

(A) 
(B) 
(C) 

If we put all nine movable cubes depicted above on the table we can produce different configurations on the surface of the table. In (A) there’s a regular NxM array, with N and M both = 3. If we remove one cube, as in (B) we can arrange the remaining cubes in a regular NxM array, e.g. N=4 and M=2. If we remove yet another cube, as in (C), can we still produce a regular array using the remaining cubes? In which situation is it possible to arrange the remaining blocks in a 3-D configuration of NxMxO blocks, where N, M and O are all > 1?

Note that unlike the previous examples, in Case (C) we can’t draw the impossible configuration under discussion, namely a regular NxM array made up of exactly seven blocks. Case (C) would be dealt with easily by someone who has already learnt about prime numbers, and understands the unique factorisation theorem. But perhaps playing with the cube rearrangement task could lead a bright child to notice the impossibility, and eventually prove that the problem is not one of a failure to explore enough configurations, even without previously having learnt about prime numbers.

Exercise for the reader: what would have to go on in the child’s mind for this to happen?

One of the features of mathematics is the variety of interconnections between different problems. Sometimes essentially the same mathematical problem is discovered in quite different contexts. I think even human toddlers and intelligent non-human animals can make such discoveries and use them, but without being aware that they have done so, and consequently not being able to ask or even think about the question: “How do I know that no exception will be discovered on a high mountain, or at a freezing temperature, or while travelling in space?” They are incapable of noticing the epistemological features of the mathematical discoveries they use. And if they grow up to be philosophers with no understanding of computation, they may misdescribe what they have learnt.
Discovering primeness geometrically

This example is also in the discussion of toddler theorems, though I have never encountered an instance of the possibility considered here: http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html#primes

A child given a set of wooden cube-shaped blocks can do all sorts of experiments -- exploring the space of processes involving the blocks.

- Some of the experiments involve learning about the material of which the blocks are made.
- Some involve learning about types of physical interaction -- e.g. what happens when you bang two blocks together, or throw a block, or push one over the edge of the table, or what difference it makes whether the floor has a carpet or not when you are trying to build towers, or what happens if you put a block in a cup and shake it in various ways.
- Some of the experiments may lead to discoveries of properties of arrangements of various kinds. E.g. a group of blocks can always be arranged in a line (if there's space on the table, or on the floor, or in the room,...). But sometimes the blocks can be arranged into other configurations, e.g. a square frame, a rectangular frame, an rectangular array.

A child may notice that in certain cases, attempts to rearrange a configuration into a rectangle always fail: What kind of experimentation can that provoke, and what sorts of discoveries can be made?

![Fig Block-arrays](image)

How could one be sure that there is NO way of arranging the last collection into a rectangular array, apart from the straight line shown? Could a child playing with such blocks (or discs, or other movable objects) discover the concept of a prime number? I suspect it could be done using an ability to "carve up" a spatial region in a systematic way and then confirm by exhaustive analysis that no possible distribution of the blocks produces a rectangle, other than the co-linear arrangement.
When I discussed this hypothetical example (discovering theorems about factorisation and prime numbers by playing with blocks) with some people at a conference, one of them told me he had once encountered a conference receptionist who liked to keep all the unclaimed name cards in a rectangular array. However she had discovered that sometimes she could not do it, which she found frustrating. She had unwittingly discovered empirically that some numbers are prime, though apparently she had not worked out any mathematical implications.

Could the child rearranging blocks discover and articulate the fundamental theorem of arithmetic? (The unique factorization theorem.)

**Physical constraints vs rule-based constraints**

The sliding tiles puzzles above were introduced as if the routes along which the tiles could slide were physically constrained. Related problems can occur naturally when the task is to find a route from some location to another location satisfying some constraints.

However the physical world leaves open the possibility of lifting the tiles and moving them to new squares, a process that would be unconstrained in both varieties of the puzzle presented earlier. Someone wishing to achieve the specified end state could therefore ignore the constraints and make more direct moves.

But for some reason there could be a preference for moving the tiles along the constrained paths, requiring routes to be found, when possible. In that case, the preference could arise even if there were no physical constraints: one can "playfully" explore what is possible (a) when only vertical and horizontal moves are considered and (b) when only diagonal moves are considered. The same sort of mathematical reasoning would be relevant both to the situation with physical constraints and to the situation with non-physical constraints, only freely adopted rules.

One consequence of this is that an engineer who discovers "in the abstract" that certain sorts of constrained moves would have useful consequences in some situation can then construct mechanisms in which physical structures impose those constraints, so that all the permitted changes in those physical structures necessarily have the desired properties. Examples include designing channels for flow of water or other liquids, designing grooves along which balls can roll, designing rails to control motion of trucks, designing linkages (e.g. to produce bi-stable car boot (trunk)) lids, designing gears that control the relative speeds of rotation of two axles (and things attached to them), and many more. (The use of tools, a focus of much research in psychology, is a special case of this phenomenon: tools are aids to controlled matter manipulation. Some of them transform forces in addition to constraining motion, e.g. levers, gear-wheels, screw-drivers, pincers, etc.)

The fact that mathematical investigations can be addressed in contexts where structural constraints are "freely adopted" could lead (and I think has led) some philosophers to the mistaken conclusion that mathematics is a human creation, and contains only freely created constructs, with no absolute necessities. Wittgenstein famously wrote: "For mathematics is after all an anthropological phenomenon" (in Remarks on the Foundations of Mathematics). But clearly the consequences of freely adopting a precisely defined constraint are not themselves freely adopted, any more than the consequences of a strong physical constraint with the same structure are freely adopted.
Moreover, it is not only humans who discover and use mathematical structures and their properties: other intelligent animals do also. E.g. weaver birds make use of mathematical properties of knots. Moreover evolution (natural selection) has discovered and made use of many mathematical facts, e.g. that certain sorts of control systems (using negative feedback to achieve homeostasis) will produce stable temperatures or pressures or orientations. Many more subtle mathematical facts must have been used to allow control systems in developing brains to control various kinds of motion by changing their details (their parameters) during growth of an organism, a process in which absolute and relative sizes of body parts change, and their weights, moments of inertia, strength and other features change, but without requiring growth of new brains (or sub-brains) to control the physical configurations with all their new properties and relationships.

Such control mechanisms can be thought of as using primitive "grammars" for processes. This may be relevant to unanswered questions about evolution of languages for internal use and for communication. [Sloman][Vis-Lang].

This discussion raises interesting biological questions. What differences are there between brains that can solve problems involving satisfying constraints only when the constraints are externally imposed, e.g. by physical structures, and brains that can also solve similar constraint satisfaction problems where the constraints are not freely chosen, but adopted for some practical reason, e.g. using rules as constraints in reasoning about a diagram representing a physical structure with corresponding physical constraints, like an architect deriving consequences of some design decisions by reasoning with architectural drawings.

This is part of the evidence that evolution can be usefully considered to be a "blind mathematician", a view discussed in several of the Meta-Morphogenesis project papers.

In the context of the discussion of evolved construction-kits in a separate document construction-kits.html a distinction is made between physical, abstract and hybrid construction kits. A game in which players are constrained by physical structures is a kind of physical construction kit. If some of the physical constraints are removed, like the constraints restricting the trajectories of tiles on a surface, but the effects of the constraints are adopted as constraints on solutions to problems (i.e. rules restricting possible actions), then the result is a "hybrid" construction kit. Many games played by humans, such as soccer, cricket, tennis, and others are based on such hybrid construction kits.

In some cases, such as chess, or draughts (checkers) or GO it is even possible to play the game without making any use of physical pieces or a board, by simulating their effects. Such a game is then a purely abstract construction kit. A great deal of mathematics is concerned with investigation of properties of such abstract construction kits. But many of the original mathematical discoveries were based on concrete versions of those construction kits, where constraints (or rules) came from perceived physical structures not from intentions of players or social agreements. These developmental steps are nothing like the repeated use of statistical evidence in neural-net based models of learning, e.g. in "deep learning" mechanisms.

How and why and when and how many times, did evolution produce new brain mechanisms making novel problem solving (and problem recognition) capabilities possible? Which were the earliest cases in evolutionary trajectories leading to modern humans? Producing such mechanisms is itself a partly mathematical problem, of finding structures and rules that support the discovery processes. The ability to solve problems arising in the use of abstract construction kits, benefited...
from earlier evolutionary transitions that solved useful practical problems. But the new mechanisms constructed often had additional powers beyond the powers required to deal with the specific problems encountered.

Added 10 Nov 2016
Some of these ideas are taken a bit further in a paper (written October/November 2016) discussing the nature of mathematics and the argument that mathematical discoveries were made and used by natural selection long before there were human mathematicians. 
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/what-is-maths.html
(That was followed by later papers on mathematical discovery on the same web site.)

Domestic possibilities and impossibilities (furniture)
Domestic life is full of possibilities and impossibilities that you may or may not care about. Many people have discovered impossibilities involving doors and chairs illustrated in the example below. If you wished to slide the chair out of the room through the doorway, you might find it impossible because the chair is too wide. It is possible to rotate such a chair about a vertical axis (e.g. rotating it with all four legs in contact with the floor), but that does not allow the chair to pass through the door, if the width of the seat and also its depth (front to back) both exceed the width of the door.

Chair-Door (a)

Chair too wide to slide through doorway.

-----------------------------
However, in such a case it may be possible to rotate the chair in a different way, about a horizontal axis, so that it ends up on its side, with its feet pointing horizontally instead of downwards. In that orientation the chair may be able to pass through the doorway if it is slid until two feet are projecting around the door frame then rotated so that the other two feet can move through the doorway, after which the chair can be rotated again into the original vertical orientation.

![Chair-Door (b)](image)

If height of back of chair is less than width of doorway, the chair on its side can pass through doorway if the back feet are pushed through first, then the chair slid to the right, then the chair rotated around a vertical axis to get the back past the door frame after which the seat and front legs can be pushed through.

Can you see an alternative possible arrangement that allows the chair to go through the doorway (without being dismantled, folded, etc.)?

The chair example illustrates the fact that realising a new possibility (rotation) can remove some old possibilities (e.g. sitting on the chair) and introduce some new possibilities (e.g. getting through the doorway). Furthermore such removal and creation of possibilities can be chained to form a sequence of steps to bring about a new situation that could not be achieved by simpler direct motion of an object from its initial position to a desired new position.

Of course the possibility of getting the chair through the door existed from the start, but it was temporarily blocked by the orientation of the chair and the width of the door. Changing the chair’s orientation produced a new direct possibility.

Many action plans created by humans and other intelligent animals depend on the ability to recognize and reason about sets of possibilities and impossibilities, and ways in which they can be chained usefully.
Constructive, supportive impossibilities

Sometimes the impossibilities are obstructive, like the impossibility of moving the chair through the doorway in its original orientation. But other impossibilities provide essential constructive support for new possibilities. For example if you cannot reach something on a high shelf, you may be aware that there is a rigid box that you can place on the floor below the desired object. The rigidity of the box may remove the possibility of your being moved downwards by gravity if you stand on the box. In that case standing on the box creates a new possibility for you, reaching the object on the high shelf.

This is a common feature of biological processes: some new possibility once realised removes previous possibilities (moving downwards) and thereby enables further new possibilities (reaching to a new height). This depends on the rigidity and strength of the material used. The ability of plants to grow upwards (e.g. to get more light) depended on evolution of mechanisms making possible the growth of materials that reduced or removed possibilities of bending, i.e. rigid materials. Without such changes to the materials constructed during development, plants could not have evolved on dry land as we know them. Giant redwood trees are an extreme example. Roses or sunflowers held up by stalks would also have been impossible.

Terrence Deacon (2011) seems to me to be confused about achievements of biological evolution because he emphasises the negative aspects (constraints) in new developments (e.g. production of rigid materials) without noticing the positive aspects (e.g. enabling new possibilities, such as supporting heavy structures). Anyone who has played with construction kits, such as Meccano, will have made use of the fact that separate parts can be locked together (e.g. forming a hinge) thereby both constraining their independent motion and enabling new possibilities, such as building a structure with a part that can move relative to another part, such as the jib of a crane. Related points about molecular level construction kits were made in What is life?, by Schrödinger 1944. It seems to me that in that little book he came very close to recognising the practical importance to many organisms of abilities to recognise examples of impossibility and necessity. but did not quite get there.

So new developments may be both negative (constraining) and thereby also positive (enabling) at the same time. This is part of the intrinsic nature of possibility and necessity in spatial structures, relationships and processes, with rich implications in structures with non-rigid spatial relationships between rigid parts. See also Sloman (1996).

Being aware of such relationships and their implications is an important feature of (well developed) human perceptual consciousness, and apparently also consciousness of some other intelligent species.

Impossibilities concerned with cardinality

The following is a subset of the discussion on "What needs to be explained" added recently as a new Note to Chapter 8 on "Learning about numbers" in Sloman (1978).

Piaget, unlike many researchers into number-cognition, had read deeply in philosophy of mathematics (including work by Gottlob Frege and Bertrand Russell) and knew that understanding one-one correspondences is central to understanding the natural numbers (as cardinals). Consequently, whereas many researchers assume that being able to recognize and name the numerosity of small collections of objects is evidence for possession of a concept of number,
Piaget realised that far more is required: in particular a grasp of features of one-one correspondences that children acquire only gradually. In some cases they are not understood fully until close to age 6 years (Piaget, 1952). [The precise ages are not important: the possible forms of partial understanding are.]

Chapter 8 of Sloman (1978) attempted to illustrate some of the algorithmic and architectural requirements of a learner developing information about number names and how to use them in various practical tasks, all of which depend on the use of one or more one-one correspondences, including correspondences between objects or events and an initial sequence of number names. Some of the implicit themes were made explicit in the Note to Chapter 8, added in 2016. This section summarises a subset of that note.

In his (1952) Piaget used a variety of experiments to probe the ability of children to recognize and make use of one-one correspondences, and their abilities to reason about those correspondences, e.g. answering questions about whether and why the correspondences are or are not preserved by various actions. Many researchers attempted to replicate, or modify his experiments, but often labelled what they were studying as something like "understanding conservation", without any theory of what made such understanding possible. A useful summary of some of this work by Saul McLeod with videos can be found here: http://www.simplypsychology.org/concrete-operational.html

Piaget famously discovered the apparently staged development of the ability to understand that it is impossible for a one-one correspondence to be destroyed by a mere re-arrangement of the objects involved. His postulated stages of development need not concern us now.

Understanding that sort of "invariance" is essential to understanding the cardinal numbers. But before children reach full understanding many of them seem to regard a rearrangement that stretches out or compresses a collection of objects as altering numerical equality between that collection and another collection. So even if they have seen and accepted the original correspondence and have also seen objects being moved to form a perceptibly longer collection, but without addition or removal of any objects (i.e. simply creating a new one-one correspondence between initial and final elements of the groups) some of the children apparently think that this rearrangement changes the one-one correspondence between the objects as they were and the objects in the new configuration. But the questions asked do not normally explicitly refer to one-one correspondence. The experimenter may ask whether there are more objects than before the rearrangement. When children mistakenly say there are more, they may be answering the wrong question, or they may not understand the invariance. Figure Transitive below, illustrates a special case of the problem.

Figure Transitive
This diagram summarises two physical examples with a common structure.

**Case 1:** there are three groups of objects A, B, and C (not labelled) in the figure. If elements of set A (on left) are in 1-1 correspondence with elements of set B in the middle, and elements of B are in 1-1 correspondence with elements of another set C, on the right, then the two correspondences can be "joined" to form a 1-1 correspondence between elements of A and elements of C. A child with an understanding of number will see that it is impossible for the first two correspondences to exist without the third also existing. This might be based on the visual ability to see how each link in the first correspondence can be combined with a unique link in the second correspondence to form a new link from the first to the third set.

This correspondence is not affected by the way elements of the sets are distributed in space: e.g. one set may be compact and another stretched out. Likewise for sets of events with different time-intervals between the events.

**Case 2:** the diagram can also represent one group of physical objects first translated from locations on the left to locations in the middle of the diagram, then translated to the locations on the right. When this happens the two transformations can be composed to form a new transformation from the left hand group to the right hand group. What a learner needs to understand is that despite the changes in appearance of the group of objects after each transformation, no new objects are added and none are removed, and there is a one-one correspondence between the initial locations and the final locations of the objects. It is impossible to destroy a one-one correspondence simply by moving objects around, if no objects are destroyed or merged or split into smaller objects.

As far as I know, Piaget was not able to explain how a child (or adult) can see in both Case 1 and Case 2, that if the first two correspondences exist the third **must** also exist. My impression is that many psychologists who have read or heard about this work by Piaget do not understand the deep implications of the computational requirement to represent and reason about 1-1 correspondences. So the label "conservation" is used to sum up what the children have or have not understood when they succeed or fail in Piaget's tests. And, as Annette Karmiloff-Smith once remarked, they try to vary the tests to find out whether children can pass some variant at an earlier age, but without providing any analysis of requirements for passing or for mechanisms that can meet those requirements.

"Decades of developmental research were wasted, in my view, because the focus was entirely on lowering the age at which children could perform a task successfully, without concern for how they processed the information." [Karmiloff-Smith(1994)]

Frege and Russell essentially tried to show that this is merely a case of (rather complicated) logical deduction that could be expressed in the symbolism of modern logic. However the ancient Greeks and many others had already discovered and used such properties of numbers long before the invention of modern logic.

Understanding the concept of cardinal number includes understanding why a one-one correspondence between two collections of discrete items is preserved no matter how the items are re-arranged, as long as no objects are removed, merged or separated into two or more parts. Piaget's work showed that this understanding does not come automatically with being able to count or being able to answer questions correctly in special cases. I suspect that neither Piaget nor
anyone else knows how brains represent information about particular one-one mappings or acquire abstract non-empirical knowledge about general properties of transformations that involve one-one mappings. The concept of a one-one correspondence between two arbitrarily large collections of objects of any type (concrete, abstract, physical, mental, etc.) is not one that fits any mechanism I have ever heard a neuroscientist describe. Without that, our concept of cardinal number cannot be understood. I suspect animal brains, and especially human brains, use important mechanisms that have not yet been identified by neuroscientists.

Frege 1950 attempted to show that such mathematical knowledge is purely logical, but it is clear that mathematicians understood these properties of cardinality before the logical apparatus used by Frege and others had been discovered.

In order to understand how an AI system can understand the natural numbers as they were understood before the rise of logic, we shall have to explain how it can reason visually (e.g. using a diagram, or imagining a possible change in some collection of objects) and thereby discover that one-one correspondence is transitive (among many other properties). The figure illustrates what is discovered but does not explain how. I don’t think anyone knows how human brains make such discoveries, how the discovered information is stored, how the brain mechanisms allow future inferences to be made, and how all this knowledge is acquired in a form that is independent of how many objects are involved, how big or small they are, what shape they have and whether they are physical objects or locations, or places, or abstractions such as number names, or how the information that there are infinitely many possible cardinalities is represented in brains. These cannot be statistical generalisations from perceived examples, since when understood the generalisations are known to have no exceptions. Moreover they can be understood as applying not only to collections previously encountered but to arbitrarily large collections of objects or events or names, etc. How can biological brains support such competences and discoveries? So far I don’t think these abilities have been replicated in AI systems, though I suspect that may simply be because we have not yet discovered the right forms of representation and the required information processing architectures. [To be added: refer to work of Doug Lenat, Simon Colton, Alison Pease and others who have attempted to model mathematical cognition.]

Non-human reasoners about possibilities and impossibilities

A close study of many intelligent non-human animals reveals wide-spread abilities to identify and use chains of possibility and impossibility in complex actions that lead to achievements of goals. A striking example was Betty -- a New Caledonian crow studied in Oxford -- who in 2002 astounded not only the researchers but also many reporters and members of the public, when she managed to use a straight piece of wire to make a hook that she then used to fish a bucket of food out of a vertical glass tube, as reported in Weir, Chappell and Kacelnik (2002) and demonstrated in this Youtube video: https://www.youtube.com/watch?v=UDg0AKfM8EY.

There are several more videos showing Betty’s behaviour here: http://users.ox.ac.uk/~kgroup/tools/movies.shtml.
It was not reported at the time, but is clear from the videos on the project web site (a) that she makes the hooks in several different ways, with (approximately) functionally equivalent results (i.e. using a crack in the plastic tray to grip the end of the wire, using the tape at the base of the tube to grip the wire, using her foot on a horizontal rail to grip the wire, and using a hole in the wall next to a small perch to grip the wire, (b) that there appears to be no random trial and error in her hook-making or hook-using behaviour, (c) that each of the episodes of hook-making and use involves several different steps in which some possibility is identified, and then new possibilities and impossibilities are achieved on the basis of the previously realised possibilities. There is no evidence that she had had experience of bending pieces of wire, or similar materials before these experiments, although when she first bent a straight piece of wire she had previously used a bent piece of wire provided by the experimenters.

Betty seemed to be conscious that

- it is possible for the end of a hook to be passed under the handle of the bucket holding food,
- it is possible to raise the hook in that configuration
- it is impossible for the bucket to remain where it is when the wire is raised with the hook looped through the handle
- continued raising of the hook will lift the bucket past the top of the glass tube
- after rising beyond the top of the tube the bucket can be moved sideways then downwards to the table
- food can be extracted when the bucket is on the table

A detail that is easily overlooked is that when she inserts the hook into the tube she also places one foot on the rim of the tube. That provides two widely separated support points, presumably allowing more precise control of the wire when moving the end under the handle of the bucket. Later, she uses the foot on the rim to achieve sufficient height to pull the hook and the bucket out of the tube. Examining the last few seconds of the video suggests to me that without the foot on the tube she would not have been able to achieve the height required. How much of that Betty understood is not clear, but in the video of trial 7 she did not first try without grasping the rim.

Fig Betty hook
Toddler Topology

Piaget seems to have been aware that for young children, understanding and using topological relationships (relationships between structures that involve continuity, orderings and "betweenness", but without using metrical properties, e.g. length, area, volume, exact orientation, degree of curvature, etc.) played an important role. An example is understanding the advantage of pushing a drawer shut with a flat palm rather than using a hand curved over the top edge of the drawer, as explained here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html#drawer

Another sort of example is understanding some of the things that can be done with holes, an example being Betty’s ability to pass the end of a hook through the "hole" provided by the bucket and its handle. A video of a pre-verbal toddler apparently using a pencil to explore different routes through a hole in a sheet of paper is here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html#holes

She seems to be testing a hypothesis about 3-D topology (the existence of continuous deformations between two configurations) though she could not, at that age, have formulated such a hypothesis in words. As far as I know there is no current robot that can become aware of such a possibility and thereby acquire the motive to make the possibility actual, which seems to happen a great deal with human children (illustrating conjectures about "Architecture-based motivation" Sloman (2009b)).

The Side Stretch Theorem

A separate discussion document concerned with changing areas of triangles Sloman (2012) introduced the Side Stretch Theorem (SST), illustrated below

We can formulate the "Side stretch theorem" (SST) in two parts:

(SST-out) IF a vertex of a triangle is moved along an extended side away from the interior of the side (as in Figure S) THEN the area of the triangle increases.
IF a vertex of a triangle is moved along a side towards the interior of that side, THEN the area of the triangle decreases. (Draw your own figure for this case.)

This illustrates a deceptively simple example of a mathematical relationship between a type of spatial structure, types of process that can occur, and types of change necessarily associated with those processes. The simplicity is deceptive because although you are likely to find the claimed mathematical relationship obviously true, it is not at all clear what sorts of visual information processing mechanism provides the ability to think about the possibility of change when it is not actually occurring and find the relationships obvious. For example, the formulation in terms of an area increasing or decreasing presupposes the existence of a shape independent measure of area, and that any such measure exists is not obvious. Clarifying that concept for an arbitrary area was a major mathematical achievement, closely related to the discovery of integral calculus. However, there is a simpler mathematical discovery that does not require the notion of a measure, only the notion of inclusion or containment, a part-whole relationship.

**The Side Containment Theorem (SCT)**

It might be fruitful for the reader to pause here and try to formulate and prove a Side Containment Theorem (SCT) expressed in terms of which of two triangles contains the other, without assuming any measure of area or length, illustrated in Figure SCT:

![Figure Side Containment](image)

For now, I'll leave open the question whether the Side Stretch Theorem (SST-in/out), or the Side Containment Theorem can be derived from something more basic and obvious, requiring biologically simpler, evolutionarily older, forms of information processing.

Note, however, that the notion of the vertex being "moved along" a line "away from" or "towards" another point on the line implicitly makes use of a metrical or semi-metrical (ordering of lengths without a numerical measure) notion of length, which increases or decreases as the vertex moves. The concept of motion between two locations on a line also implicitly assumes the existence of intermediate locations between those locations.
Properties of triangles and other abstract shapes

Other documents on this site discuss discoveries that can be made by looking at shapes and imagining ways they can be altered, the constraints on possible alterations, and the necessary consequences of such alterations.

**Added 15 Oct 2017:** A problem was presented in Birmingham in Sept 2017, with some unanswered questions, discussed here:

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

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

Compare Freksa *et al.* (2019).

The thought experiments discussed there could all have been performed by ancient thinkers long before the development of modern logic and formal methods of proof. Such pre-logical mathematical thinking was the only kind available to Euclid and his predecessors, since modern logic was developed only in the last few hundred years.

**Conjecture:** the ability to make those mathematical discoveries, and others in Euclid’s *Elements* were dependent on abilities to notice and reason about possibilities, of the sorts discussed above, along with other social abilities necessary for publishing or teaching the materials.

Work to be done still includes identifying the precise visual functions needed and precise specifications for the mechanisms providing those functions. We may then be able to build artificial systems (e.g. robots) whose visual and mathematical capabilities are far more like those of humans than any existing robots.

Until we know how to do that, our robots and other forms of AI will all be severely limited and capable only of very simple forms of learning. They will not be capable of the forms of learning and discovery that drove mathematical discoveries in humans.

**Strings**

There are many ways flexible strings can be moved around. In particular, a string can be threaded through one or more holes in a piece of leather (as in a shoe). Suppose it goes through only two holes: how many different ways are there of removing the string from the holes? How can you be sure that you have counted them all? [Assume two removal processes are the same if the ends of the string go through the same holes in the same direction.]

You can remove the string by pulling one end, or by pulling the other end. Why can’t you remove it even faster by pulling both ends? What needs to be added to current robots to enable them to (a) discover such impossibilities, (b) understand why they are not possible?
If you pull both ends at the same time, there is a configuration that can be achieved faster: what configuration? The ability to answer that might be based on searching through a mass of data concerning previous pulling episodes. But that isn’t required. What sort of ability would enable a robot to answer the question without resorting to experiments with strings and holes, and without searching through stored records of previous such experiments? How do you answer the question?

**Strings, pulleys and levers**

In Sloman (1971) and in Chapter 7 of Sloman (1978) a comparison was made between reasoning processes using Fregean forms of representation (i.e. representations based on application of functions to arguments, e.g. logical and algebraic notations) and reasoning using "Analogical" forms of representation (often confused with representations based on isomorphism between thing represented and thing representing -- which is merely a special case).

![Figure Strings-pulleys-levers](image)

Such diagrams of mechanisms can be interpreted subject to constraints, such as that levers are rigid and can move only by rotating around their pivot points, and strings are flexible but unstretchable. Using such assumed constraints it is possible to reason that if end \(a\) of the left lever moves up then end \(f\) of the right lever will necessarily also move up, so that both levers will rotate clockwise about their pivot points (indicated here by small triangles). Whether actual strings, levers, pulleys, etc. have the postulated properties is an empirical question, but the consequences of having them can be derived by non-empirical reasoning, using the diagram.

It was argued that such reasoning could be as useful, and as reliable, as reasoning based on logical and algebraic forms of representation. In both cases the reasoner has to make assumptions about how the form of representation works, i.e. what various structures and transformations represent, and on the basis of such general assumptions draw conclusions by reasoning about particular cases. These forms of reasoning, familiar in uses of maps, in uses of diagrams in physics and electronics, and in architectural drawings are all examples of valid reasoning that is not based on statistical evidence or inferred probabilities.

In simple cases such reasoning can make use of imagined transformations of an imagined diagram: there is no need for a physical diagram to be used. (A similar comment applies to reasoning using a logical notation, or natural language: the reasoning can be done "in one’s head".)
I suspect that pre-verbal children and some highly intelligent non-human animals (e.g. squirrels and crows) are to some extent capable of such valid reasoning using non-Fregean forms of representation *internally*, but unlike adult humans they are incapable of reflecting on what they are doing, communicating it to others and defending or criticising the validity of the reasoning. That requires meta-cognitive extensions to the information-processing architecture.

**Topological additions October 2017**

**Superpositions**
There are several examples of impossibility of two object views being views of the same object after translation, rotation or viewing from a different direction. This includes:

- Detecting impossibility of superposition of pairs of 2D objects by translating and rotating them in a plane, e.g. "d" and "b", or two spirals where each is the mirror image of the other.

- Detecting similar 3D impossibilities, e.g. superimposition of a right hand shape on a left hand shape by any rigid 3D transformation. Compare two "opposite" helixes or two screws that are identical except for threads in opposite directions. (Their non-coincidence was one of the examples of synthetic necessary truth in *Kant* (1781).)

There has been considerable psychological research on human "mental rotation" abilities, initiated by Shepard and Metzler, summarised in [https://en.wikipedia.org/wiki/Mental_rotation](https://en.wikipedia.org/wiki/Mental_rotation), comparing difficulty of tasks, times required to detect rotation, etc.

The following figure (from Wikipedia) includes typical 2D and 3D examples of mental rotation challenges.

![Figure Mental Rotation](image)

There is a large freely available collection of pairs of images of 3D structures made of rectangular blocks here: [https://openpsychologydata.metajnl.com/articles/10.5334/jopd.ai/](https://openpsychologydata.metajnl.com/articles/10.5334/jopd.ai/)

For a machine vision system to perform this task the fact that the structures compared decompose into straight segments, or segments composed of cubes or rectangular blocks meeting at right angles makes it possible for a relatively simple algorithm to check whether superposition is possible in a finite number of steps (left as an exercise for the reader).

But as far as I know nobody has investigated brain mechanisms that could enable brains to detect impossibility of superposition by rotation and translation. I suspect individuals use a large collection of learnt heuristics that vary according to culture, age and individual.

Things get far more computationally explosive if more general shapes are used, as illustrated here using 2D shapes as examples. (Compare Minsky and Papert on Perceptrons. [REF])
Figure 2D Congruence
Which pairs of shapes are congruent if translation and rotation in the plane are allowed, but not flipping over or reflection across a line?

How would the answers change if reflection across a line, or flipping a shape over in 3D, were permitted?

For different pairs of images different heuristic methods can lead fairly quickly to answers, e.g. counting the number of 'ends' in each image (some have four some have two), testing whether a shape can be coloured in more than one colour without the two colours merging (a test for connectedness), imagining a feature of one image being superimposed on the other and following round both images looking for a location of mis-match, and finding what happens if attempts are made to imagine pairs of ends superimposed then checking how far the superposition extends.
(That reveals that two of the images that are reflections of each other about a line in the plane cannot be superimposed.)

**Figure: 2D Distortions**

Which pairs of shapes below are congruent if translation and rotation in the plane are allowed, but not flipping over or reflection across a line? Which can be made congruent if reflected across a line? Which images are proper parts of others, possibly after reflection and/or rotation?

By now some readers will have noticed that far from being purely mathematical challenges unrelated to anything else, these challenges are closely related to a common type of picture puzzle for children, where the task is to look at a picture and find one pictures of more more familiar objects forming undifferentiated parts of the picture. The concept of impossibility is relevant to this task in order to rule out answers based on finding a structure that could be part of something (e.g. a bicycle) but whose continuation could not be.