DRAFT (December 7, 2008): liable to be updated.

Could a Child Robot Grow Up To be A Mathematician And Philosopher?

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

These slides, and a shorter version with different examples, presented at MKM08, are in my 'talks' directory.

http://www.cs.bham.ac.uk/research/projects/cogaff/talks#math-robot

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mkm08

Two papers, one for MKM'08 and one for the AI Journal 2008, are included here

http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0802

http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0807

I apologise

- For length and poor organisation: The topic has proved more complex than I expected when I agreed to give the talk – I shall go on working on these slides in an attempt to improve them, in the light of discussions at the seminar, and comments from colleagues.
- For slides that are too cluttered: I write my slides so that they can be read by people who did not attend the presentation.

  So please ignore what's on the screen unless I draw attention to something.

No apologies

For using linux and latex

This file is PDF produced by pdflatex.

NOTE:

Although I don't give any references to G.Polya, the approach taken here is much influenced by reading his *How to Solve It*.

NOTE:

A shorter presentation of some of these ideas, with different examples, was given at the MKM08 conference in July 2008. The slides are available here `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mkm08`

Philosophical views on the nature of mathematics and mathematical knowledge

I shall not discuss most of these views: they are listed merely as reminders of the breadth of the field – and the list is incomplete.

- Mathematical knowledge is empirical (J.S.Mill?)
- It just makes explicit assumptions in our definitions, and therefore does not extend knowledge – mathematical truths are analytic (Hume?? Russell, in 1959)
- Arithmetic is reducible to logic – which does not make it trivial (Frege)
- All mathematics is reducible to logic and definitions: all mathematical reasoning is purely hypothetical and content free (Russell)
- Geometry is not reducible to Logic
  (Frege: discussed in recent publications: Teri Merrick, Ivan Welty)
- Mathematics (including meta-mathematics) is the study of formal systems (Hilbert)
- Logic is concerned only with syntax, not meaning. (Hilbert? Carnap? ....)
- Mathematics extends our knowledge but is not empirical (Kant)
- Conventionalism/Anthropologism (Wittgenstein?)
- Other issues: Platonism, Fictionalism, Mathematics is created not discovered, etc.
- Mathematics is/isn't mechanisable. (Turing, Gödel, Lucas, Penrose, Chaitin...)
- Intuitionism, (Brouwer, Heyting) constructivism, finitism, ultrafinitism ....

See also James Franklin's papers in my 'Warning' slide at the end.

Psychological views on the nature of mathematics and mathematical knowledge

In addition to philosophers and scientists in their philosophical moods, there are also many psychologists and educationalists who have attempted to develop theories about the nature of mathematical knowledge and the processes of mathematical discovery.

Some of that work is deeply muddled because of false assumptions, for example the assumption that the perceptual ability to distinguish small groups of different sizes shows a grasp of what numbers are, or, more generally, the assumption that it makes sense to ask whether a child (or other animal) has or has not grasped the concept of number – ignoring the fact that our concept of number, even just our concept of the non-zero natural numbers (1, 2, 3, 4, ...) is deep and multi-faceted, so that different subsets of the facets can be grasped by different individuals, or the same individual at different times.

Often such confusions come from the misguided operationalist drive to find some decisive experimental test for having the concept.

Great teachers understand that any mathematical idea is something that has to be explored from multiple directions, gaining structural insights and developing a variety of perceptual and thinking skills of ever increasing power.

In particular understanding a mathematical idea typically includes being able to detect cases where your understanding is partial or erroneous and to work out how to extend or remedy it.

Understanding what numbers are is a never-ending process.

This presentation does not establish the above claims, but does attempt to illustrate them with examples of some of the processes involved, and attempts to draw conclusions about the information processing architecture required by a learner of mathematics.

High level plan

My aim is to defend an updated AI-inspired version of Kant's philosophy of mathematics, extending my first attempt (in my 1962 DPhil Thesis).[*]

I shall proceed in a round-about way by contrasting

- Russell's view of mathematics as

  the investigation of implications that are valid in virtue of their logical form, independently of any non-logical subject matter,

- Feynman's view of mathematics as

  "the language nature speaks in".

Kant's view, that mathematical knowledge is synthetic and non-empirical (but not innate), can be interpreted as close to a modified version of Feynman's view, something like this:

> The ability that develops in a child, namely to see, manipulate, predict, plan and explain structures and processes in nature, requires a collection of information-processing mechanisms that turn out also to be the basis for the ability to make mathematical discoveries, and to use them.
> (And closely related to the ability to notice and investigate philosophical problems.)

This process could be replicated in a suitably designed robot.

[*]The thesis was recently digitised and is now online here:
`http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-1962`
Knowing and Understanding

The Kant-Feynman philosophy of mathematics

Consider a reformulation of Feynman's thesis:

If interrogated appropriately, nature is capable of providing information about itself (to animals, infants, robots and scientists)

The ability to acquire information from such encounters requires general biological information-processing capabilities (some of which have been explored in AI/robotics)

Some animals (and future robots) can acquire and use second-order information about the information;
e.g.

- **information about what forms of representation are good for representing certain aspects of nature**
  - **think about the power of numerical measures, maps, and differential calculus**
- **information about what can and cannot be done with the information,**
- **and information about the processes involved in all this.**

Mechanisms required for all of this provide a major part of the ability to do mathematics.

Gain some first hand experience

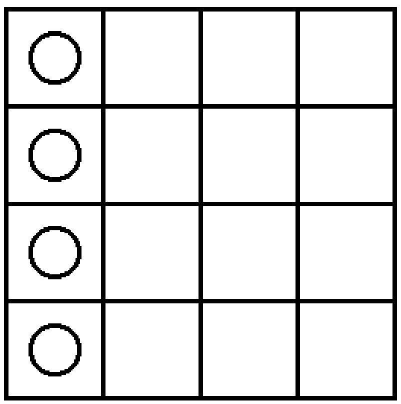
See what you can learn by thinking about these two puzzles.

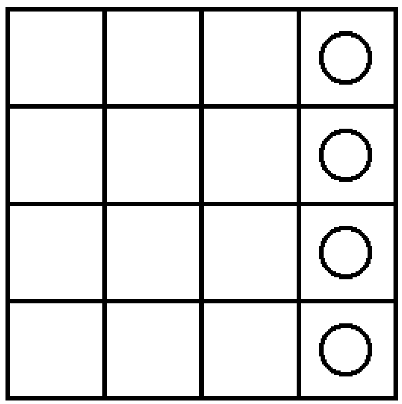
Take notes on what you do and what you learn and how you learn it.

A warm-up exercise:

Can you slide the 4 coins from column 1 to column 4 using only diagonal moves?

Using only diagonal slides, can you get from this to this?

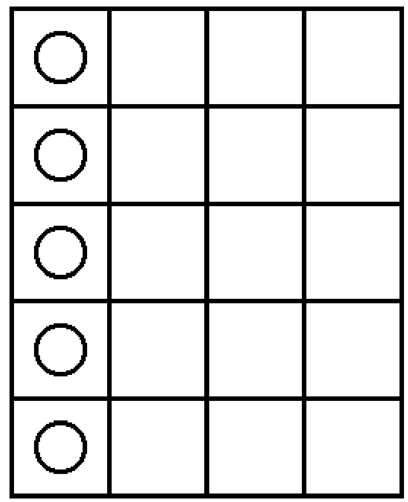


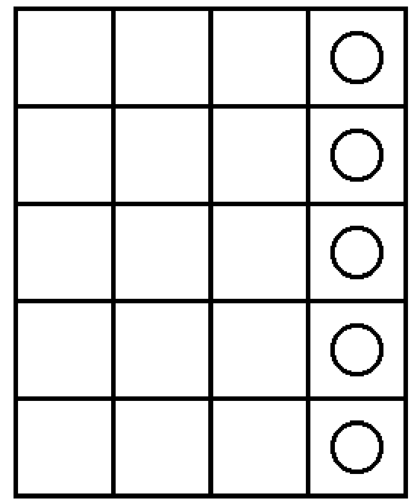


What is the minumum number of moves required

Now stretch yourself:

Can you slide the 5 coins from column 1 to column 4 using only diagonal moves?
Using only diagonal slides, can you get from this to this?





What is the minumum number of moves required?

The problems are discussed later. How people work on them varies according to prior knowledge and experience.

What can be learnt by interrogating nature

Topics for further investigation (I have time only for a few examples):

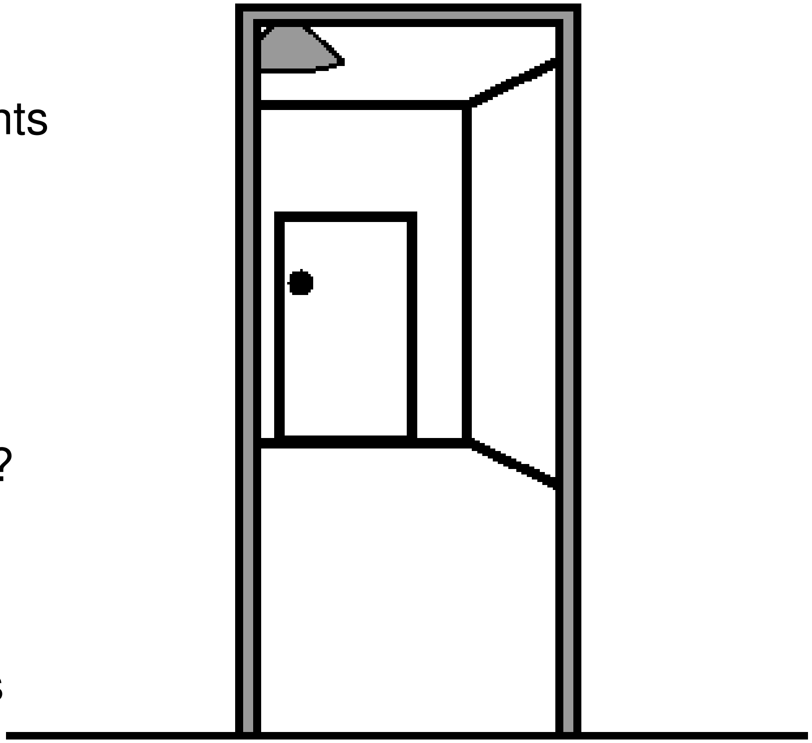
- some of the ways nature can be interrogated, e.g.
  - perceiving
  - acting and perceiving
  - getting information from others who have already acquired information
- some of the kinds of information that can be acquired by such interrogation, e.g.
  - about what particular things and what types of things exist in the environment
  - about possibilities for change and limitations of possibilities
  - generalisations about what happens when
  - limitations and benefits of particular forms of representation
  - the need to modify or extend current ontologies (CRP, 1978 Chap 2)
- some of the things that can be done with the information, e.g.
  - formulating new kinds of theory and new kinds of goal
  - achieving practical goals (changing the environment, including online control)
  - understanding causation and making correct predictions
  - explaining WHY things are as they are, in two ways:

    o Deriving consequences from theories about hypothesised mechanisms (including testing)

    o Investigating limits of what is possible in a world for which a certain form of representation is appropriate (e.g. a certain sort of geometry, a certain kind of logic).
- Information-processing architectures, mechanisms, and forms of representation required for all this to work (Including architectures that grow themselves.)

Getting information about the world from the world

Things you probably know:



- You can get more information about the contents of a room from outside an open doorway
  - (a) if you move closer to the doorway,
  - (b) if you keep your distance but move sideways.

  Why do those procedures work? How do they differ?
- Why do perceived aspect-ratios of visible objects change as you change your viewpoint?

  A circle becomes an ellipse, with changing ratio of lengths of major/minor axes.
  Rectangles become parallelograms
- In order to shut a door, why do you sometimes need to push it, sometimes to pull it?
- Why do you need a handle to pull the door shut, but not to push it shut?
- Why do you see different parts of an object as you move round it?
- When can you can avoid bumping into the left doorpost while going through a doorway by aiming further to the right – and what problem does that raise?
- How you could use the lid of one coffee tin to open the lid of another which you cannot prise out using your fingers?

The need for a self-monitoring architecture

The previous slide showed examples of your knowledge about both

1. action affordances: things you can and cannot do
2. epistemic affordances: information you can and cannot acquire
3. meta-affordances: how your actions can change both the action affordances and the epistemic affordances in various situations.

The ability to discover such things requires information-processing abilities that include not only

- the ability to perceive, form goals, make plans, and execute plans to achieve goals,

but also

- the ability to observe oneself in doing those things, and to notice recurring patterns and constraints that can be used in future to work out in advance what can and cannot happen in various situations.
  (This is one of the functions of the meta-management architectural layer in H-CogAff:
  http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#enf07
  http://www.cogs.susx.ac.uk/users/ronc/cogs.pdf)

I shall try to show that such discoveries, though they may be hard to come by without experiment and exploration, are not empirical, like the discovery that most things expand on being heated and the discovery that ice is an exception: they don't have to be tested by doing experiments on things in the environment.

The discoveries are both non-empirical, and also synthetic (knowledge-extending) as Kant claimed.

They are typically mathematical discoveries.

The child as self-programming explorer

A child who plays with toys and various parts of its body, later also learns to play with information structures.

Besides learning to manipulate objects,

a child also learns many hundreds of ways of acquiring, manipulating and using information in the first years of life;

but understanding why they work, and what their limitations are, comes later.

This requires architectural and representational changes.

Some of the books by John Holt provide many examples:

How children learn (1967), and others:

`http://www.holtgws.com/booksbyjohnholt.html`

`http://en.wikipedia.org/wiki/John_Caldwell_Holt`

I'll give examples where exploring the world and noticing patterns in the process can reveal more than just empirical generalisations.

This is part of the basis for mathematical competence.

Acquiring object-level and meta-level knowledge

Gilbert Ryle distinguished “knowing how” and “knowing that”.

We can distinguish **object-level** practical knowledge and **meta-level** practical knowledge.

- Most animals, very young children, and current robots have only object-level practical knowledge, i.e. know-how (including knowing how, knowing that, knowing who, knowing when, ...)
- This may come from evolution (the only source for most animal species), or from training, e.g. learning associations between goals+circumstances and actions that will achieve the goal in a situation, or by building up records of what’s where when.

  The latter is often labelled “episodic” memory and misleadingly contrasted with “semantic” memory.
- Such know-how is implicit in all feedback control mechanisms that achieve or maintain some state.
- But it is possible to have the object-level knowledge and lack meta-level knowledge: most animals don’t know what they can and cannot do, under what conditions they can do it, why the right actions succeed, why the wrong actions fail, etc.

  Even human self-knowledge of this sort is always limited.
- Some mechanisms that provide the meta-level knowledge can also contribute to mathematical knowledge.
- Most current AI robotic research aims only at giving object-level know-how.

I’ll return to meta-level knowledge later.

Can a robot learn the "language of nature"? (Feynman)

Very few animal species have an explicit understanding of the information available and usable in the environment, including some usable for acquiring more information (epistemic affordances).

None have it to the same extent as humans.

So far no robots have much of it.

One way of developing a theory of that understanding is to work out how to build a robot that acquires that understanding.

Most animals (precocial species) have almost all they need to know about the world "pre-compiled" by evolution (apart from parameter values calibrated or optimised after birth or hatching), whereas humans and a few other species (altricial species) learn a great deal by playing, exploring and experimenting in the environment.

Some can go on learning indefinitely:

I found out a few days ago how to disassemble a PC to enable me to replace a faulty fan held in place on the motherboard by spring-clips: and having seen the structures and relationships involved, I also now know **why** the fan cannot safely be removed without the hassle of extracting the motherboard.

As a result of searching for relevant information about the process I learnt that some motherboards have a 'North bridge' CPU chip, and why they need them, and why such a chip needs a cooling device.

Some of what I learnt I gained from my manipulation of the cables, plugs, screws, spring-clips, fan and the motherboard, while some came from other humans, via google and email.

Varieties of information from nature

Some animals (especially, but not only, humans) can indefinitely extend their understanding of structures and processes in the environment.

This includes finding out how structures and processes in the environment provide information that can be used both:

- to control actions, and also
- in forming generalisations, extending ontologies, making plans, finding explanations, testing hypotheses, debugging plans and theories, etc.

Evolution produced some learners specially equipped to learn and to use this "language" of nature, whose information is often expressed in changing perceived structures.

Much "communication" by nature depends on the fact that the environment includes a host of interacting 3-D structures and processes, including structures and processes produced by the learner's actions.

J. J. Gibson, *The Ecological Approach to Visual Perception,* 1979

Eleanor J. Gibson and Anne D. Pick,
An Ecological Approach to Perceptual Learning and Development, 2000

Nature provides information about action affordances and epistemic affordances.

Mechanisms providing the ability to go on learning more and more about the "language of nature" also provide part of the basis for human mathematical (and philosophical) competences.

Varieties of learning and development (Chappell&Sloman, 2007)

Cognitive epigenesis: Multiple routes from DNA to behaviour, some via the environment

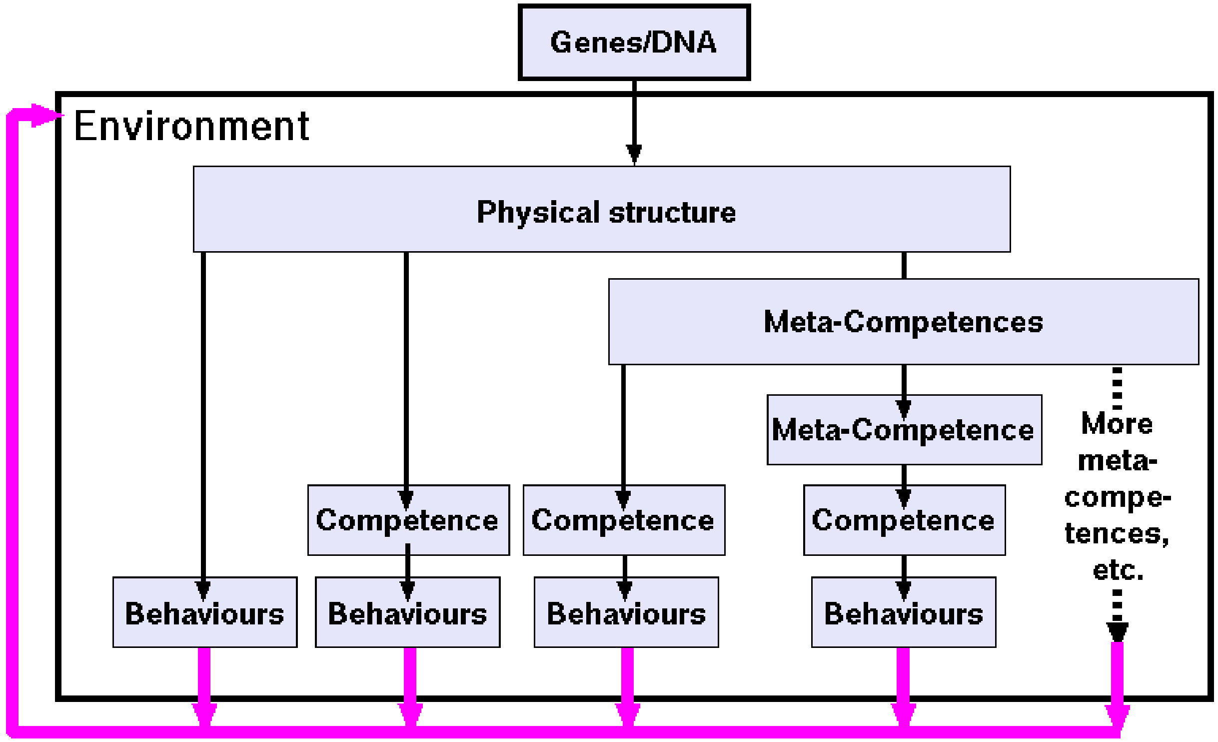
Pre-configured competences:

are genetically pre-determined, though they may be inactive till long after birth (e.g. sexual competences), and their growth may depend on standard, predictable, features of the environment, as well as on DNA.

They occur towards the left.

Meta-configured competences:

(towards the right of the diagram) are produced through interaction of pre-configured or previously produced meta-configured competences with the environment (internal or external).



(Chris Miall helped with the diagram.)

The environment changes the learning and development mechanisms.

Evolution 'discovered' that speed of learning is increased by active intervention: it produced some species that discover many facts about the environment, and themselves, through creative exploration and play, in which ontologies, theories and strategies are developed, tested and debugged.

Two very different views of mathematics

Bertrand Russell: Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true. (In *Mysticism and Logic* 1918)

Richard Feynman: To those who do not know Mathematics it is difficult to get across a real feeling as to the beauty, the deepest beauty of nature. ... If you want to learn about nature, to appreciate nature, it is necessary to understand the language that she speaks in.

In *The Character of Physical Law*

Russell's view makes it at first sight puzzling that mathematics can be relevant to the world we live in, or any other specific subject matter.

We'll remove the puzzle later, and generalise the answer.

These quotes about mathematics, many other quotes, and lots of mathematics-related fun and information, can be found in:

`http://www.cut-the-knot.org`

What Russell really meant

The 1918 aphorism above is a pithy summary of his view expressed more fully in 1903:

> "Pure Mathematics is the class of all propositions of the form p implies q, where p and q are propositions containing one or more variables, the same in the two propositions, and neither p nor q contains any constants except logical constants. And logical constants are all notions definable in terms of the following: Implication, the relation of a term to a class of which it is a member, the notion of such that, the notion of relation, and such further notions as may be involved in the general notion of propositions of the above form. In addition to these, mathematics uses a notion which is not a constituent of the propositions which it considers, namely the notion of truth."
>
> ....
>
> "The fact that all Mathematics is Symbolic Logic is one of the greatest discoveries of our age; and when this fact has been established the remainder of the principles of mathematics consists in the analysis of Symbolic Logic itself."
>
> *The Principles of Mathematics* (1903)
>
> `http://fair-use.org/bertrand-russell/the-principles-of-mathematics/`

He later wrote, in *My Philosophical Development*, (1959):

> "I have come to believe, though very reluctantly, that it [mathematics] consists of tautologies. I fear that to a mind of sufficient intellectual power, the whole of mathematics would appear trivial... I cannot any longer find any mystical satisfaction in the contemplation of mathematical truth."

Not everyone agrees with Russell!

I'll try to explain what Russell, despite his enormous achievements, failed to notice.

That includes unpacking the required kind of "intellectual power".

Russell's mistake

I shall not dispute that the subject matter and form of investigation discussed by Russell is part of mathematics (e.g. use of logic to investigate logical implications).

Rather, I shall argue that mathematical logic is only a subset of mathematics, a subset that grew out of something broader and deeper that is central to the intelligence of humans, some animals and future robots.

This partly overlaps with ideas of Kant, Frege, and various others, on kinds of reasoning in mathematics – and daily life.

This was originally the claim of my DPhil Thesis in 1962, but I did not then realise that something like AI was required to substantiate the claim.

I made the link with AI in a paper contrasting reasoning using Fregean and analogical representations at IJCAI 1971, criticising logicist AI (summarised by McCarthy and Hayes 1969) as too narrow:

`http://www.cs.bham.ac.uk/research/projects/cogaff/04.html#200407`

Applying mathematics to dividing chocolate

You have a slab of chocolate in the form of a 7 by 7 square of pieces divided by grooves, and you want to give 49 friends, each one piece.

You have a knife that can cut along a groove.

What is the minimum number of groove cuts that will divide the bar into 49 pieces?

RULES FOR CHOCOLATE CHOPPING:

Stacking or overlaying two or more pieces, or abutting two pieces, to divide them both with one cut is not allowed: each cut is applied to one of the pieces of chocolate.

I don't know where this problem originated.

I first encountered it many years ago, when it was frequently used (along with the mutilated chess-board problem and others) to illustrate the role of intelligence in finding the right representation of a problem.

The problem is included, along with many other fun things relevant to this talk in

`http://www.cut-the-knot.org`

Applying mathematics to dividing chocolate

How can you find out the answer to the question about the minimum number of cuts required to produce 49 pieces from a 7 by 7 square of pieces of chocolate?

(Without stacking/overlaying two pieces, or abutting two pieces, to divide them both with one cut.)

- You can play around with lots and lots of 7x7 bars trying different methods.
  NOTE: the number of possible ways of doing the division is pretty big!
- You can also do it for different sized slabs, e.g. 2x5, 6x3, 4x4, etc.
- You may start noticing a pattern for each size: E.g. a 2x5 array always requires 9 cuts, a 6x3 array always requires 17 cuts, etc.
- If you run out of chocolate slabs you can test the pattern out on more examples, using squared paper instead of chocolate slabs.
- You may realise that the number of cuts does not depend on what the material is that you use or what the knife is made of: only "the structure" of the process (an abstract pattern), not the material operated on, matters. (Why?)
- You may also notice that in addition to the pattern for each size there is a more general pattern that applies to all the sizes. Namely? (Answer coming soon.)
- Eventually you can see **WHY** the result generalises for all possible rectangular blocks:
  NOW you have become a mathematician – if you weren't already one.

Proof of the chocolate theorem

Seeing the proof involves seeing the pattern in a class of processes:

Suppose the number of pieces in the slab is P.

- Every cut divides one piece into two pieces.[*]
- So every cut produces one more piece than there was before.
- Initially (after no cuts) there is one piece.
- After the first cut there are two pieces.
- After the second cut there are three pieces.
- K-1 cuts produce K pieces. (A crucial, abstract, part of the **pattern** of actions and results.)
- After the last cut there are P pieces.
- Therefore
- There must have been P - 1 cuts at the end.
- Corollary: The minimum number of cuts is also the maximum number!

[*] Why is this true? Could a cut produce no new pieces? Or more than one more? We'll see later.

There is more to be said about this theorem

E.g. consider:

- generalising the theorem: it is not restricted to flat rectangular blocks made of square pieces, but works for any starting shape (e.g. a sphere) and any final collection of parts produced by a sequence of cuts each dividing a portion into two portions;
- the cognitive apparatus and forms of representation required to see the common pattern:

  At what age can a typical child see the pattern?
  What changes in the child that makes the pattern visible to the child?
  Can any other animals see such a pattern?
  What brain mechanisms are needed?
  What sort of robot could see the pattern?

- the cognitive apparatus required to notice that the patterns are independent of many features that vary between situations (e.g. colour and taste of the chocolate);
- the cognitive apparatus required to see why the theorem **must** be true given the formulation of the problem (i.e. there is no need to check it out on other planets, at high temperatures, in strong magnetic fields, etc. etc. – these cannot make a difference to the theorem: why not?)
- the possibility of counter-examples that have nothing to do with physical, or empirical facts, but about the structure of the space of geometric possibilities.

  Can you think of counter-examples? I'll present an example later.

Formalising the intuitive proof

Once the core ideas of the problem and the proof have been worked out, it becomes possible in principle to produce a logical formalisation of "the theory of chocolate chopping".

There are different ways this could be done, using different presuppositions in the formalisation.

E.g. one choice is whether to

- assume only very general logical principles, and then include an explicit formalisation of arithmetic as part of the theory
- assume the whole theory of arithmetic as "background" knowledge, to be used without requiring any justification (including the principle of mathematical induction).

The process of formalisation is not uniquely determined by the initial discovery: there may be different ways of abstracting away from details of shapes, cutting procedures, physical materials, ...

The result of the formulation would be a set of axioms using undefined predicate and function symbols, along with a proof of a theorem that could be checked against a set of formal rules without understanding anything about what is said in the theorem or any of the axioms, or any of the intermediate formulae.

Some mathematicians will not be satisfied until a formalisation is complete, while others will not care whether it is done or not.

This talk is not about formalisation, but about the earlier mathematical discovery process and how it relates to animal and robot intelligence, and requirements for modelling it.

Generalising the theorem

Mathematical thinking about the problem can continue after the answer has been discovered, and the proof of the answer discovered.

A mathematical curiosity can generate further questions about almost any discovery you have made:

- Do we really need to physically separate out the pieces after each cut?
- Isn't there something about the structure of the proof that remains unchanged if we make all the cuts without moving any of the chocolate?
- Suppose the grooves were not there to start with – why should that make a difference?
- If the chocolate pieces are not moved after each cut, what feature of the process corresponds to cutting only one portion at a time?

EUREKA: There's a theorem about what happens when you divide a closed plane region by repeatedly drawing a (loop-free) line across one of the sub-regions.

It generalises to repeatedly slicing a 3-D volume – harder to visualise, except in cases like chocolate chopping where all the slicing planes are perpendicular to one plane through the chocolate.

(And to higher dimensions, harder to visualise.)

Discovering the conditions for the theorem to hold needs some care.

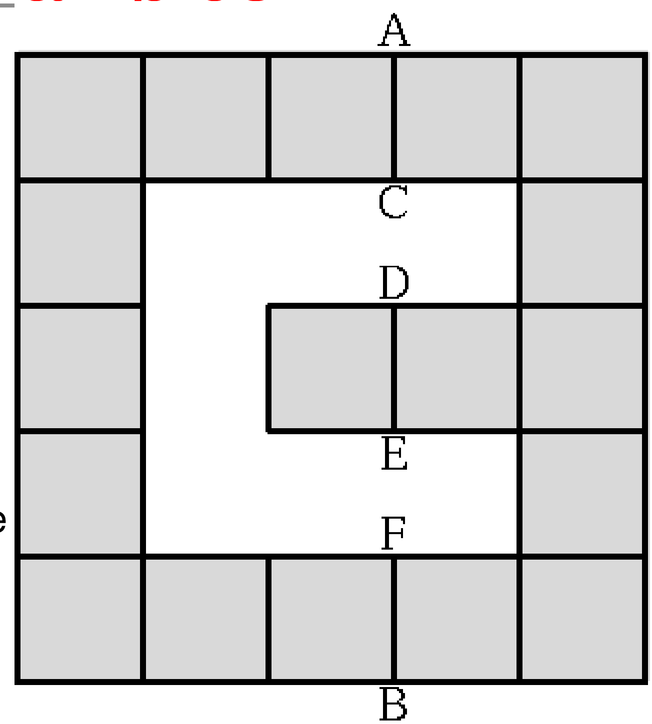
Discovering counter-examples

Previously it was stated that:

- Every cut divides one piece into two pieces.

Often a proof in mathematics that seemed valid works for a range of cases, but has counter-examples not thought of when the proof was constructed.

See Imre Lakatos: *Proofs and refutations: The Logic of Mathematical Discovery* CUP, 1976
(Analyses history of Euler's theorem about plane polyhedra.)



One of the consequences of our ability to perceive, imagine, or create instances of possibilities not previously encountered is that we can sometimes create new configurations that refute our generalisations. This is different from the empirical refutation of "All swans are white", which turned up in Australia.

Consider the "slab" of chocolate shown here.

Is it a slab or isn't it? I did not previously say anything about disallowing holes.

- If you call it a slab, and allow that a cut from A to B is consistent with the rules, then the key "pattern" that each cut produces one more piece does not apply here: there are two new pieces.
- If you say that a cut must join two boundary points and not cross a boundary, then consider the cut from A to C, or from F to B: that leaves the same number of pieces as before
- The "cut" joining boundary points C and D has no effect at all: only the DE cut produces one new piece.
- You can eliminate cases like this by stipulating that the initial slab must not contain any holes.
- However an unusual (non-convex) outer boundary for the slab can produce yet more counter-examples: finding them is left as an exercise for readers.

Another “Lakatosian” counter example

At a presentation of these slides on 21 Jan 2008 Mary Leng (Liverpool University) pointed out a possibility I had not noticed: making partial cuts – a new possibility that undermines the claim that every cut produces one new piece.

- The nice feature of this counter-example is that it does not require the invention of specially shaped slabs.
- In the original slide defining the problem, I had not noticed the need to specify that every cut must go from one boundary point to another: i.e. no cuts may begin or end at a point that is completely surrounded by chocolate.
- This example illustrates the relationship between simple everyday activities, and variations that are clearly intelligible to ordinary people with no knowledge of abstruse mathematics, and deep concepts from topology.
- Alison Sloman later pointed out that the counter-example might have been ruled out in advance by requiring portions of the slab to be broken rather than cut.

It is important not to inflate Lakatos’ argument in *Proofs and Refutations* as demonstrating that there is never any real progress in mathematics, or that mathematics is empirical.

On the contrary, every mistake that leads to a revision of a definition, or a statement of a theorem, or a proof adds to our mathematical knowledge: Learners do not need to be infallible.

See also `http://newcriterion.com:81/archive/18/may00/lakatos.htm`

Bugs in proofs and bugs in strategies

Discovering a bug in a proof, by finding a class of counter-examples is closely related to discovering a bug in a practical strategy by discovering a class of counter-examples.

An example might be a child discovering that although pulling something towards you by grasping an edge can be a good strategy, e.g. pulling a toy truck, or a blanket, there are situations where that can lead to pain, and also prevent the intended goal from being achieved.

An example is pulling a door shut by holding its edge: OUCH!

Manfred Kerber once told me about a related piece of de-bugging by a child who liked pushing drawers shut.

You can use your own creative geometrical reasoning capability to work out what the (painful) bug was, using the clues above.

You can probably also work out the revised strategy the child devised for getting the drawers shut.

Sometimes there is more than one way of debugging a faulty action strategy, or more than one way of achieving a type of goal: discovering a new way to achieve a goal is not unlike finding a new way to prove a mathematical result.

Demonstrating that will involve building working systems that share information-processing competences between the two.

Is any of mathematics non-empirical?

Since mathematical conjectures and means of proof can be triggered by experiences, and since bugs in proofs can be discovered, doesn't that show that mathematics is really empirical?

This is a tricky question: answering it requires an excursion into what we mean by 'empirical', 'depends on experience', 'experience', etc.

For now we can make a few points:

- As Popper often pointed out: where you get your theories from is irrelevant to their status as scientific theories – we can say the same about mathematical discoveries.
- What is more important is what kind of testing is relevant.
- Drawing pictures of black swans, or imagining them, is not relevant to testing the theory that all swans are white, whereas finding black swans in Australia is relevant.
- Drawing a picture of a slab of chocolate with holes, or imagining one, is relevant to testing the claim that one cut always produces one more piece.
- Imagining what you can do with a number claimed to be the largest prime number (e.g. adding one to its factorial) is relevant to refuting the mathematical conjecture that there is a largest prime.
- What you can imagine or draw or otherwise represent may show something about the space of possibilities inherent in a generative system: that's not the same as showing what physical objects or events can exist.
- There is no significant difference between constructing proofs or counter-examples in your head (in your mind's eye) and on paper: in both cases the mathematical competences required are similar.
- None of this will be understood fully until we have detailed designs for working systems: demonstrated to be designs for machines with the sorts of competences described (incompletely) here.

How can mathematics be applied to slabs of chocolate?

This would appear to be impossible on Russell's view, but not after closer inspection:

If a mathematical proposition of the form

P implies Q

has an antecedent P that matches some bit of the world,
then the consequence Q must also match the same or a related bit of the world.

This use of 'matches' needs to be clarified (and later generalised).

The general idea is that the match exists if there is a systematic way of linking parts of the formula P (the variables) to objects in some bit of the world, their properties and their relationships, so that with that interpretation P says something true. (Compare Tarski on truth).

Then the mathematical theorem would allow us to predict that the corresponding instantiation of Q necessarily also says something true.

So, being an applied mathematician requires having cognitive mechanisms that can

(a) take in and reason about the logical structures in the implication, to show that it is logically irrefutable (How? A topic for another time.)

(b) relate those logical structures to the world, including things perceived and things done in the world.

In what follows I shall try to generalise this, so that mathematical knowledge is not restricted to logical implications, nor even to things expressible in logic – e.g. since spatial structures and processes are sometimes more useful.

Feynman liked diagrams, so he might approve. Perhaps he also liked chocolate.

Seeing and playing with possibilities

It is often forgotten that perceptual apparatus can provide information not just about what exists in the environment, but also about what is possible in the environment, and, having discovered those possibilities an animal, or robot, can play with them, e.g. by trying various combinations of possibilities to find out what happens.

We can play in the environment, and we can play in our minds.

Playing can reveal both new possibilities and impossibilities. (Discovering constraints.)

That kind of experimentation can increase know-how, and support faster problem-solving, using patterns that have been learnt and stored in a fast content-addressable pattern matcher.

It is often suggested that this way of “compiling” knowledge from deliberative and meta-management mechanisms into pattern-based reactive mechanisms is a pervasive feature of human learning, e.g. learning to drive a car, to solve mathematical problems, to read text, to sight-read music, to generate sentences that say what you mean to say, to formulate plans in familiar situations, etc.

The analogy with compiling goes back at least to AI discussions in the 1970s.

I argued at the first Cognitive Science conference that translation to another form of representation was not the only mechanism: giving different mechanisms access to old information could speed things up, at the cost of loss of flexibility. I don’t know if that’s right.

See `http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#48`
Skills Learning and Parallelism (1981)

The need to move to a meta-level

Learning about possibilities by playing is a common biological capability, whereas finding out **why** those possibilities have certain features, understanding **why** things work as they do, is much less common: it requires a more complex architecture and new forms of representation.

It requires the ability to detect impossibilities.

To illustrate the ubiquity of the phenomenon some more examples are given.

If there is some important difference between humans and other animals it seems likely to be related this "meta-level" ability

- to monitor various processes, both internal and external,
- at the same time as they are being performed,
- to store information about them,
- to notice patterns and store them, (Selected how?)
- to detect and relationships between such patterns, as higher level patterns,
- to combine the information thus gained with future problem solving, planning or decision-making processes.
- to do some of this "off-line", i.e. representing possible events that are currently not happening.

More examples follow.

A possible experiment on young children

A child who has learnt to count, can play with a coin lying on a table and learn various things about patterns that can and cannot occur when turning a coin while counting.

- The coin has only two faces. (Call them H and T)
- There are two stable configurations of the coin on the table.
- You cannot pick up the coin using only one finger.
- It can be harder to pick up the coin using only two fingers than a cube.
- It is sometimes possible to turn a coin over with one finger.
- If a coin shows H and is turned over it will show T.
- If a coin shows T and is turned over it will show H.
- It is possible to keep turning a coin over, counting in synchrony as you do it:
  "one", "two", "three", "four", ...
- If a coin shows H and is turned over while you count "one" it will show T.
- If a coin shows T and is turned over while you count "one" it will show H.
- If a coin shows H and is turned over while you count "one" "two" it will show H.
- If a coin shows T and is turned over while you count "one" "two" it will show T.

What mechanisms enable a child to realise that it does not need to perform the turning operation 125 times, in order to answer the question:

If it starts showing H, and you turn it over 125 times, what will it show?

Can any non-human animals learn this? If not, what do they lack?

Exploring beyond easy discoveries

Discovering general facts about coin-turning and counting requires the ability to do several things in parallel (not necessarily consciously, if they are skilled operations):

- Turning the coin
- Counting (uttering the number names in the right sequence):
- Monitoring the coin-turning and counting to ensure that they remain in synchrony.
- Monitoring the combination of activities to look for patterns found.
- Remembering patterns discovered in different turning and counting process.
- Checking whether a pattern found in a particular process matches any previously stored patterns.
- Analysing discovered repetitions to see if they are accidental or necessary features of the counting process. HOW?

What sort of information-processing mechanism can do all that?

What “play-generating” cognitive mechanisms could inspire the child to look for a second coin, in order to explore relations between different ways of linking coin-turning to counting when there are two coins that can be turned, either separately or together.

Are there any interesting generalisations to be discovered with two coins?

Can you discover any, just by thinking about using two coins, without actually performing the actions?

It helps if you have already discovered the difference between even and odd numbers!

Mathematics and the world we inhabit

Points illustrated by the chocolate theorem and coin experiments:

- mathematics can be done in the abstract, as Russell says
- but the abstractions can be inspired by concrete examples
- and the results can be applied to new concrete examples.

  e.g. working out how long it will take you to divide a 2000x4000 slab of chocolate if you do one cut per second.

The mathematical knowledge once gained is not empirical.

Experiments in strong magnetic fields, or at high temperatures, or strong winds, or using materials on another planet, cannot refute the result. (Why not? – it's a complex story, not yet completely clear.)

That does not mean the knowledge is **innate**: it may be acquired through thought experiments, and reasoning processes, possibly triggered by interacting with the environment, or by thinking about previous mathematical discoveries.

We'll also see that mathematical reasoning is not **infallible.**
Mistakes can be discovered and corrected, as shown by Lakatos

Moreover it is a real discovery: when you learn the theorem you extend your competence, unlike trivial 'analytic' knowledge, e.g. All bachelors are unmarried

`http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#701`

So, as Kant claimed, the knowledge is

non-empirical (apriori, does not need experimental checking in distant galaxies)
synthetic (non-analytic, non-trivial, provides new insights, solves new problems).

Intelligence, mathematics and philosophy

A great deal of what goes on in the learning and development of a human or animal that perceives and manipulates objects in the environment and learns to predict and explain what happens, is deeply connected with our abilities to become mathematicians.

I.e. to do the mathematics that precedes logical formalisation.

The ability to do logical formalisation also builds on general perceptual and cognitive competences, including the ability to perceive and manipulate patterns in symbolic structures. (Sloman 1962, 1968).

It is also connected with the ability to become a philosopher.

There are links between mathematical reflection and philosophical reflection:

E.g. noticing how circular objects have more or less elliptical aspects from some viewing directions can lead to theories about qualia.

But we don't yet have working models of that process in a robot exploring its environment.

NOTE:

Some other animals may have only portions of the apparatus required.

That's a clue: the mechanisms are complex, and did not all evolve in one step.

Different species may have different subsets – also different robots.

Also humans at different stages of development.

Many examples need to be analysed

There are very many examples of interactions between a child, or animal, or robot and the environment, that need to be analysed in order to derive requirements for a working model in a robot.

More examples will be provided below, with indications of how perception, planning, moving, etc., require competences that are related to mathematical competence.

At present my collection of examples (and observations on them) is neither complete nor systematic.

AI still has a very long way to go before we can build working models replicating what children and other animals do.

Until we can do it, much of AI, including robotics, will be severely limited.

Also developmental psychology, work on animal cognition, etc.

A good working theory will also pose new problems for neuroscience.

And it may revolutionise philosophy of mathematics.

And philosophy of mind?

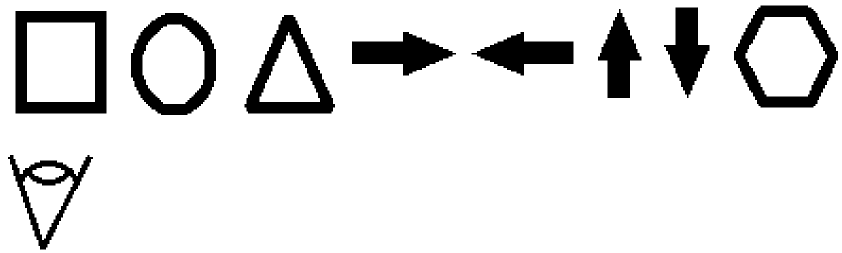
Compare: John McCarthy: “The well-designed child”
`http://www-formal.stanford.edu/jmc/child.html`

Theses

- Mathematics is not one thing: it has different facets with different features.
- There is a kind of mathematical knowledge that fits roughly with what Kant thought about mathematics, and with what I think Frege thought about geometry.
- To understand this view of mathematical knowledge we need to understand what sort of machine (what sort of information-processing machine) a mathematical thinker is.
- We can get some insights by investigating how mathematical capabilities grow – charting a partially ordered collection of stages, and explaining the ordering.
- That can provide a set of requirements, on the basis of which we need to try to design a working model of a child that develops mathematical competences and knowledge – and uses them.
- This will require us to produce new theories about human information-processing architectures and how they grow themselves.
- It will also generate new research questions.
- This can shed light on nature/nurture tradeoffs and the evolution of intelligence

  See papers by Chappell and Sloman 2005 onwards
- Developmental psychologists mostly investigate what children do when, leaving open how they do it.
- Neuroscience cannot tell us what information-processing problems are being solved, though it may give us some indication of where brain activity occurs and when.

Example: Linear order



If you move along a row of objects in a certain direction you will see them in a particular order (first the square, then the circle, then,. ...).

If nothing changes in the row of objects, and you reverse direction and move back to the starting point, you will see them in the opposite temporal order.

Is this an empirical discovery?

Could some experiment (in special conditions, on another planet perhaps) refute it?

What information-processing mechanisms would enable a child to notice that two sequences occur in the opposite order?

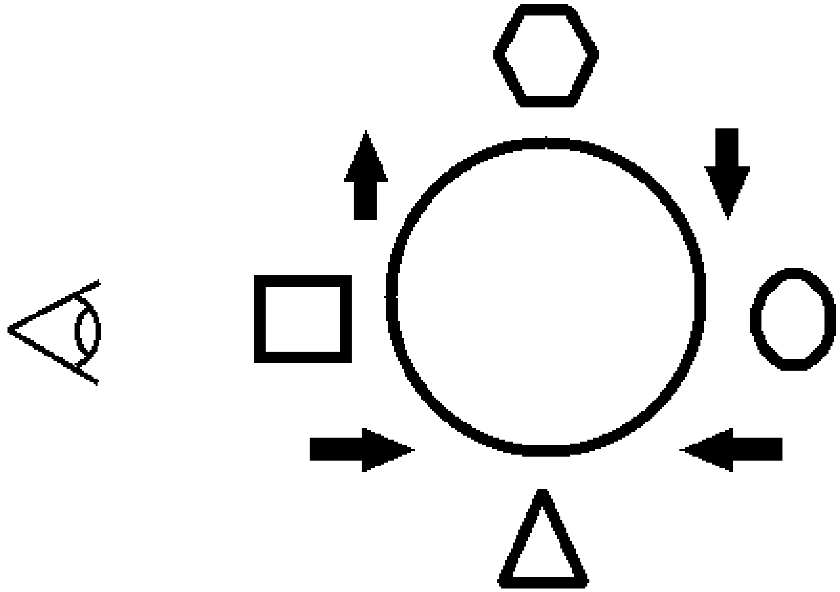
What mechanisms are required for a child to understand the necessary connection between the two temporal sequences?

Are YOU convinced? If so, why?

Example: Circular order

If you move round an unchanging configuration of objects arranged against the wall of a building, you will see them in a particular order, and as you continue moving round the building you will see the same sequence over and over again.

If you reverse direction of revolution you will see the unending sequence in the opposite order.

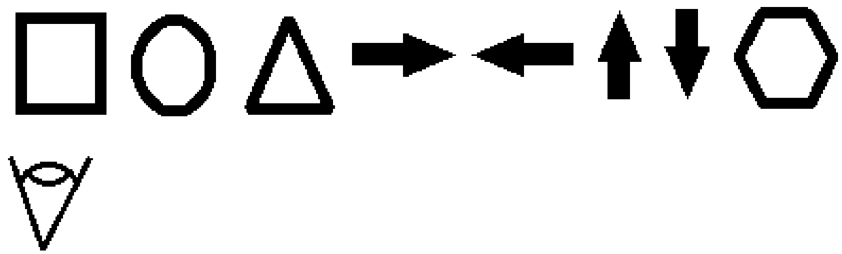


The geometric and topological relationships that hold between the objects in the environment make what is perceived predictable

What information-processing mechanisms allow a child to make this discovery, or to use it?

This is based on Kant's example of moving round a house in his discussion of causation in *Critique of Pure Reason* 1780

Example: More on linear order



If you

- move the rightmost object (hexagon) to a location below the leftmost object (square)
- then repeatedly move the rightmost object remaining in the original row to the right of the object last moved

then when all the objects have been moved they will be in the opposite order.

Become puzzled: What does being "in the opposite order" mean?

"Opposite" is a pattern linking two patterns – spatial or temporal or abstract.

A logician might give an inductive definition: does a child need that?

How many different ways do we have of specifying to ourselves what we mean by something?

Compare

- being in the opposite order

  and

- being seen in the opposite order (previous slide).

Doing a thought experiment

Suppose you don't move objects physically, as in the last slide, but merely think about doing it, using an abstract spatial representation, perhaps containing only a few objects, which you re-order in the manner required.

Can you tell by inspecting this representation of the process that all instances will have some common features?

```
Initial state     Next state:     Next state:     Final state:
   A  B  C     |     A  B      |     A        |
               |     C         |     C  B     |     C  B  A
```

Things that clearly can be varied without affecting the outcome include:

- the number of items in the row,
- their sizes,
- what they are made of, their colours,
- how heavy they are, ...
- which continent or planet they are on, ... and many more.

The only thing that is important is

- the position of each item in the initial row and
- the fact that each item is moved immediately after the item on its right.

  Another pattern:

  This procedure guarantees that taken two at a time neighbouring items will have their order reversed.

As noted in Sloman 1971 there is no real difference between doing this in your head or on paper. Why?

Experiments with orderings

It is possible to do many experiments with orderings.

- If you have a row of cards with pictures or labels on them, ordered in a certain way, from left to right, e.g.

  "A", "B", "C", "D", "E"

  You can take the cards one by one from left to right and pile them up.
- If you then unstack them, i.e. take the top one off, then the next one, and put each card to the right of the previous one, in what order will the results come out?
- Many computer programs make use of stacks in virtual machines, with the important property that the last thing stacked is the first one found on the top of the stack.
- Other things you can do is find out how many different ways there are of ordering a given set of objects.
- If you have a few classes of objects, e.g. three red, two black, four blue: in how many significantly different ways can you order them, i.e. ignoring which individual cards are where, taking account only of where the colours are in the order.
- Theorem: If there is only one colour, there is only one ordering
- What happens if you have two colours, and there are N cards of each colour?

  Anyone playing with this sort of question is taking an early step towards understanding Bose-Einstein statistics: important for Quantum Mechanics.

The need for patterns of motion, or change

Several of the examples have involved things changing in some experimental situation:

- A person moving nearer to a door and seeing more of a room
- A person moving past objects and seeing them in some order
- A person moving sideways and seeing different parts of a room
- A person moving past objects and seeing them in some order
- Cutting processes that increase the number of objects
- Counting processes
- Coin-turning processes
- Processes of re-ordering items

Perceiving a process clearly produces processes in the perceiver: things change in the perceiver.

If what is seen is remembered and re-usable, that implies that some information structure is stored which can be accessed later: a representation of the process (not necessarily representing every detail of the process).

Much is unknown about what forms of representations are good ones, what forms brains use, what forms should/could be used by robots – though there has been a lot of work on auditory memories suggesting that what is stored is itself a process, in some of those cases: rehearsal.

Re-usable information about processes

The ability to use a remembered process to produce or recognise a new process of the same type implies that there is some sort of pattern structure in the process representation: it can be re-instantiated to new instances of the pattern – perceived or created.

So our claim that patterns could be discovered in processes is not a very surprising claim – if the discovery of patterns is already a requirement for repetition or recognition.

However that leaves entirely unspecified what the form of that pattern is.

E.g. it could be an algorithm for generating instances.

If the pattern allows different forms, e.g. counting sequences of different lengths, the pattern may be stored in the form of a grammar of some kind.

Inspectable structures and processes

Some of what we have said about the difference between object-level knowledge or know-how, and meta-level knowledge, e.g. about discovering limitations of what is possible, or what must necessarily occur in certain conditions, depends on those process-representation patterns being inspectable.

We already have AI systems that can inspect some of their own data-structures and some of their own operations. (Cf. Sussman's HACKER 1975)

That's not all that different from what we need for logical information to be stored, and to be re-usable, and to be testable for validity or inconsistency.

Are spatial representations inferior to logical ones?

IF

it is clear that the representation of a process category would produce the same result if instantiated in various ways,

THEN

there is no less justification for using this as representing a whole class than for using a general logical pattern as a representation of a whole class.

Of course, the use of a system of logic can be specified as conforming to a relatively small set of rules, whereas the use of spatial structures and processes to represent new structure and processes is far less constrained, much more open-ended.

This has two consequences:

1. There is far more scope for creativity in devising a new use of spatial representations to serve a new purpose.
2. It is not possible to go through a comprehensive training process that covers all the types of structures – which is possible for a recursively defined class of structures.
3. There is more scope for error because a possibility has been overlooked.

But we have known for many years (helped by Lakatos) that mathematical reasoning is not infallible: so this is no surprise.

Mathematics and causation

Many of the examples already given and presented later are about necessary relationships between structures and processes.

But they can also be formulated in terms of causation:
doing X causes Y to happen

I.e. it is not just a coincidence or an unexplained regularity.

Understanding this requires going beyond a Humean (or Bayesian) notion of causation, and using a Kantian notion.

As discussed here:

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac`
Slides on Kantian and Humean causation in robots and animals.

Learning about (positive and negative) action affordances involves learning about causation: learning about what you can cause to happen, and how you can cause it, in various circumstances, and what can cause you to fail in certain attempts.

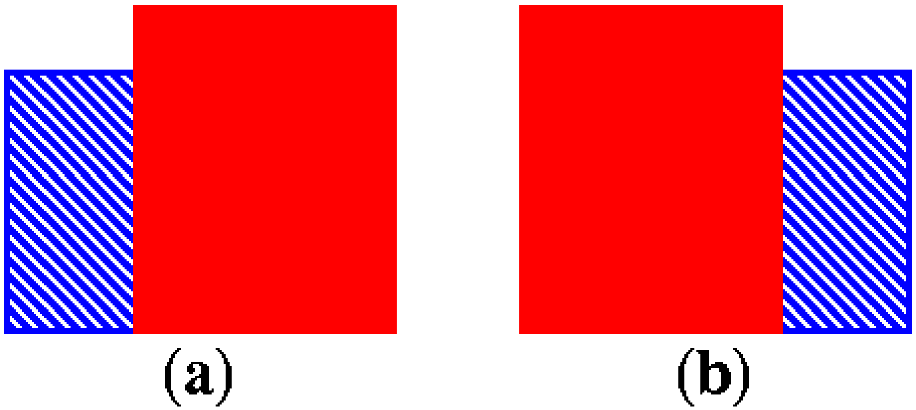
But there is also learning about epistemic affordances, i.e. learning about what **information** can be caused to become available or unavailable, by performing actions.

E.g. the example given earlier: you can see more of the contents of a room with an open door by moving closer to the doorway.

Learning about occlusion and epistemic affordances

A child can learn various things about the effects of moving in such a way as to change what it sees: at first empirically, and later understanding why

Moving from side to side can provide evidence, in the form of optical flow, that one object partially occludes another.



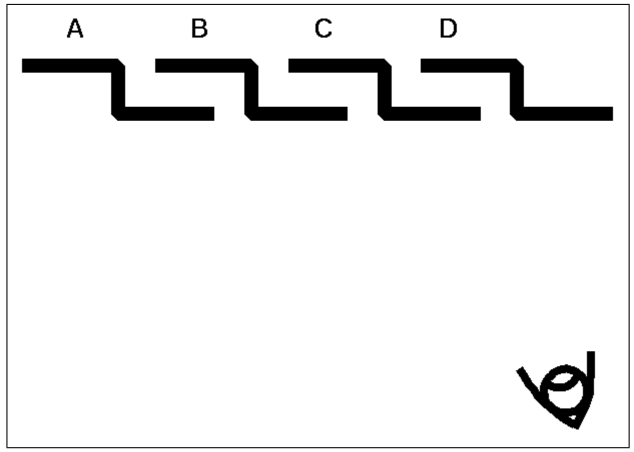
If you wish to see more of a partially occluded object you can do so by moving sideways in the direction in which the occluded object protrudes.

E.g. in situation (a) move left to see more of the blue object, in situation (b) move right.

A child could learn that the occluded object is further away than the occluding object, and "further" is transitive.

BUT....

Exceptions can be discovered



If objects are non-planar or are slanted obliquely, it can turn out that the generalisation that the occluded object is further away has to be modified:

> The portion of the occluded surface that is at the boundary been visible and invisible parts is further away than the edge of the occluding object.

So, it is possible for A to occlude B, B to occlude C, C to occlude D (all from a specific viewpoint) while D is nearer than A.

Discoveries about 3-D geometry, topology, etc. are potentially subject to indefinite refinements and corrections – but that does not make them empirical.

> Lakatos again.

How what is seen changes with motion

As you move from side to side, or rotate your head (or eyeballs) to look in different directions (including downwards or upwards) or move backwards or forwards, what you see changes systematically in many detailed ways, providing information both about what is in the environment and how you are moving, but also about what is and is not possible in the current situation, and about what information is and is not available.

The importance of this was emphasised by J.J.Gibson in *The Ecological Approach to Visual Perception*, and many examples relevant to child development are presented by E.J.Gibson and A.D.Pick in *An Ecological Approach to Perceptual Learning and Development*.

However the variety of types of information available in the environment is even richer than they suggested, and the ways in which the information can be represented, manipulated and used more diverse than they thought: there is a lot more than sensorimotor invariants.

E.g. J.J.Gibson focuses much on the use of perception for controlling action, but ignores the use of perception for finding explanations, devising multi-step plans and designing new things.

However, E. Gibson and Pick do draw attention to a child's need for representation of future possibilities, e.g. alternative routes round an obstacle.

J.J. Gibson also did not address the difference between
(a) being able to acquire and use information, and
(b) understanding why things are as they are, including predicting and explaining novel effects.

Information can be used with, and without, understanding

Many animals, e.g. insects, use many of the sorts of mathematical facts discussed here, but they do not know that they use them, or why they are usable with confidence.

Robots can also be built that learn and use associations without understanding what they are doing or why it works.

That description fits all current robotics, as far as I know.

- The systematicity in the relationships between perceptual contents and changes and what is happening in the environment can be used through purely reactive highly trained associative mechanisms (e.g. neural nets).
- However it can also be used in a different way for reasoning, predicting, explaining, and solving novel problems creatively.

  This occurs in robots that can use planning mechanisms to find new routes.
- Robots can do such things without knowing what they are doing or why it works – which would require something closer to mathematical competence.
- The relationship between perceptual competence and mathematical competence is, I believe, closely related to Kant's philosophy of mathematics.

Seeing uses exquisite, and changing, structural correspondences between what is in the environment, where the viewer is, and how the viewer moves.

We have already given some examples, including the multiple changing projections of fixed 3-D shapes into the 2-D optic array.

As the Gibsons noted: passively observed changes are rich in information, and actively produced changes can provide even richer information.

They also generate philosophical problems, e.g. about qualia, the changing contents of perceptual mechanisms!

TO BE EXPANDED

We need to talk more about similarities and differences between vision and other forms of perception: are there auditory, or haptic, or olfactory inference patterns?

You can see something cause something else to happen: can you hear causation happening or smell it happening, with the same kind of necessity, or inevitability in the events concerned?

Different sources, representations and uses of information

placeholder

Learning how to bring something to you

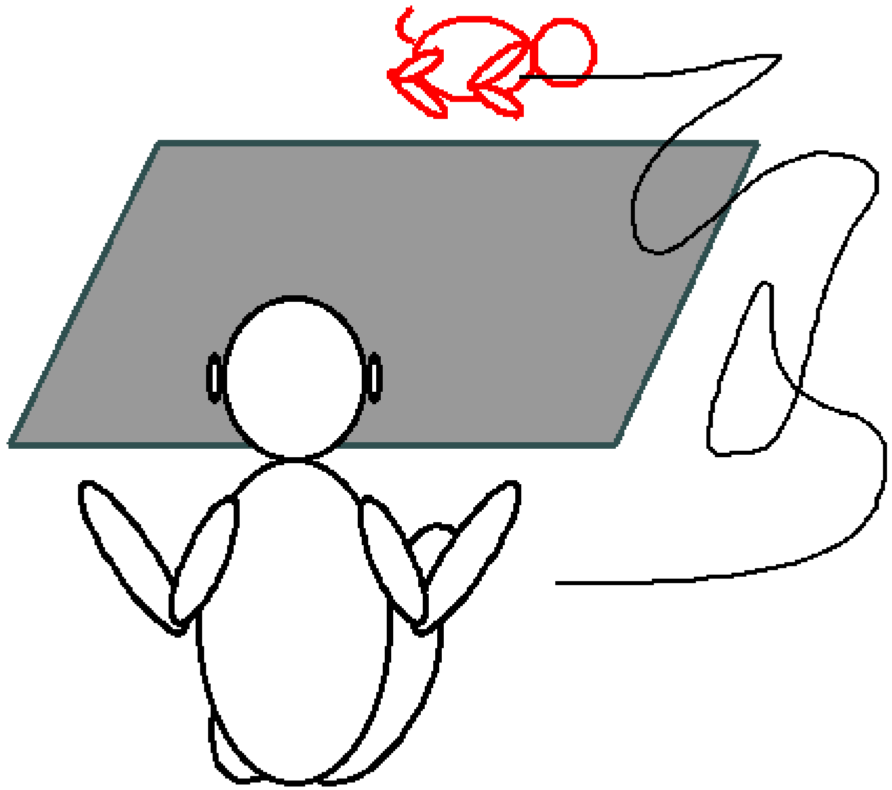
A child can learn various ways of getting hold of a visible object, depending on whether it is within reach or not, whether it is resting on something that is within reach, whether something that is within reach is attached to it, etc.

If the toy is not on the mat pulling the mat will not move the toy, but pulling the string may move the toy.

If the string is straight, pulling it will move the toy, but not if the string is curved.

However, pulling a string that is not straight can transform it into a straight string, after which the object can be moved towards the puller.

But not if the string goes round a fixed remote obstacle, e.g. a remote chair leg.



Once again, generalisations are learnt, and then counter-examples discovered, causing the generalisations to be abandoned or de-bugged, and in the process ontologies are extended, e.g. length of a curved string, knots, loops, forces being transmitted indirectly, elasticity, etc.

Recap: Mathematics and Meta-Management

- Unlike the vast majority of animal species, humans (and perhaps a few other species) have an information-processing architecture which (as a result of developmental processes) includes a "meta-management layer"

  (Terminology due to Luc Beaudoin: PhD, 1994.)

- This allows some of what is being perceived and some processes of reasoning, planning and deciding, to be simultaneously monitored and modulated.

  (Self-monitoring cannot be complete. Minsky ''Matter Mind and Models" (1868))

- Recurring patterns can be discerned and used for planning, predicting, debugging, etc.
- The ability to notice
  - which features of particular cases are immaterial and
  - which are crucial to derivations and predictions

  allows generalisations to be formed.

- This yields new knowledge about the subject matter that is not empirical.
- However such competences are not infallible either, and in some cases learning about mistakes and how to fix them can go on indefinitely generating a vast array of cases: a growing collection of concepts, including new levels of abstraction.

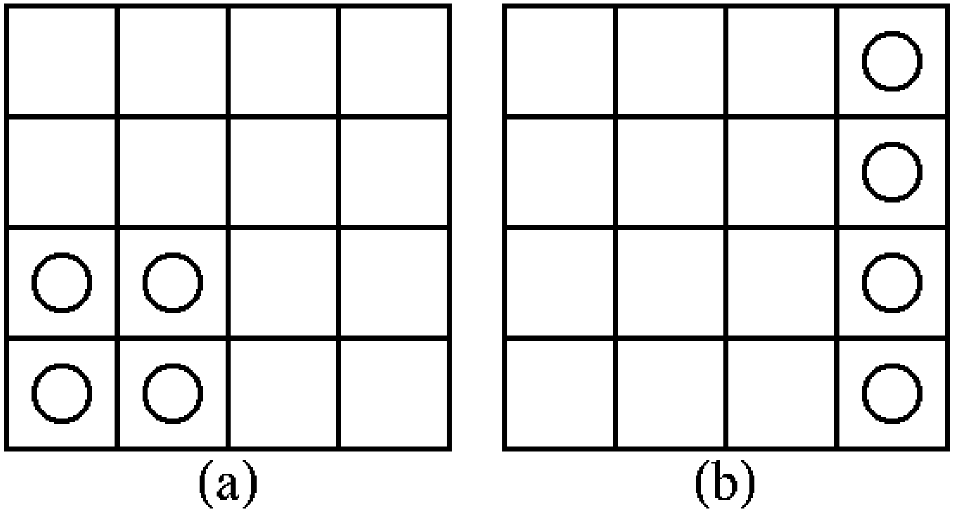
Seeing motion possibilities

For any configuration that happens to occupy some part of space, there are always variants that are possible: learning to see which changes are and which are not possible is a crucially important aspect of learning to see – for an animal or robot that can act in the world.

Some possibilities are obvious, others not so obvious.

For example if you have coins placed on a board divided into squares, it is obvious that you can slide them around into different places.

Suppose you consider only moves that are diagonal: no coin can go straight up or down or horizontally.



Question:

Using only diagonal moves, can you transform configuration (a) to configuration (b)?
What is the minimum number of diagonal moves?

Some people will find the answer obvious, whereas others will have to experiment.

You may find a some general pattern in the combinations of diagonal moves that convinces you that from any 2x2 starting configuration, the coins can always be converted to a row of four, or a column of four, using only diagonal moves.

What about converting them to a diagonal of four?

Seeing impossibilities

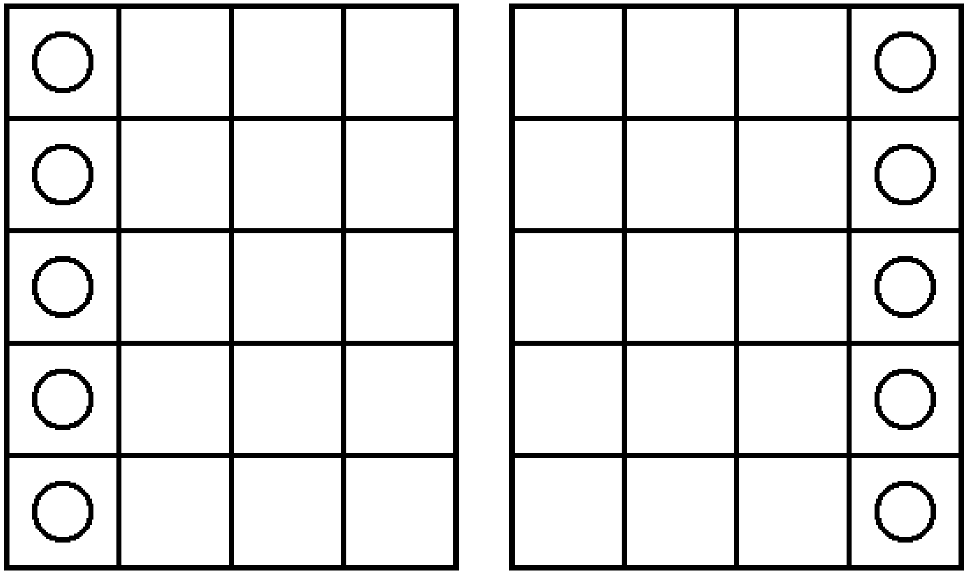
What happens if we try a different pair of configurations?

Question:

Using diagonal moves, can you transform configuration (a) to configuration (b)?

This task may look easier than the previous one, because the starting and ending configurations look very similar: in both cases it is just a vertical column of coins.

But if you try it you will eventually find that it is impossible.


(a) (b)

Question:

How can you understand why it is impossible?

You could try all possible collections of diagonal moves: the board is finite and there is only a finite number of configurations using that number of coins.

But an exhaustive analysis is very tedious: is there a better way than exhaustive search?

Mathematical intelligence essentially involves laziness, i.e. productive laziness.

In this case we want a way to see why the transformation is impossible, in a much simpler and cleaner way than by trying all possible moves.

One way is to see some possibilities that are quite different from the possibility of moving coins around.

Discovering parity

I watched someone working on the five coin problem: at first she thought it was going to be easy. Then she tried, and found a problem. After trying a few different ways, she began to suspect it was impossible. After tracing routes she saw a pattern relating locations in the left column to relations in the right column. That pattern made it easy for her co conclude that the task was impossible. She had seen the link with chess boards.

Someone once discovered that a grid of squares has an interesting property: they can be divided into two colours in such a way that squares of the same colour are never adjacent: they meet only at corners.

That fact is used in chess boards.

On the basis of that clue you may be able to work out why it is impossible to perform the transformation from (a) on the previous slide to (b) on the previous slide, by moving coins only diagonally.

That requires you to notice a pattern that is never changed by such moves.

I don't know at what age young children are capable of discovering and using the information about diagonal colouring, but more importantly I don't know what has to change in their information-processing architectures to enable them to make the discoveries discussed here.

Perhaps we can come up with good hypotheses by trying to design robots that are capable of such discoveries.

Another route to solving the coin-sliding problem

The coin-sliding puzzle involves several different kinds of structures and processes superimposed, including spatial and numerical structures and processes: Consequently there are different routes to a solution, using different forms of representation, and different cognitive competences.

A mathematical expert saw this as a problem of translating a vertical column of coins horizontally to a new location, which he expected to be easy, so he initially tried moves that he expected would demonstrate a solution, and was surprised when they failed.

He suspected that parity was involved, and realised that diagonal moves preserve the parity of dx+dy, concluding (after a few more steps) that it was possible to shift the column of coins 2k steps horizontally, but not 2k+1 steps horizontally (e.g. column 1 to column 4).

Another person experimented with diagonal moves and soon discovered that three coins on the left were competing for two slots on the right, making the task impossible.

He then noticed the connection with a bishop's moves in chess, and suggested that the task would be much easier if the squares were coloured.

Some people might notice, after playing with the puzzle, that there are many superficially different variants of the same puzzle.

E.g. on an infinite board can you get a 2x2 square of coins to shift one column right using only diagonal moves for each coin? What about a 3x3 square?

Playing with the arithmetisation of geometry

Descarte's arithmetisation of geometry was one of the greatest and most important intellectual achievements in human history; that discovery required geometrical insight as well as arithmetical expertise.

Without it, Newton's mechanics would have been impossible
(Stephen Muggleton drew my attention to this.)

A child might find it convenient for some game (e.g. "battleships") to label rows and columns of a rectangular grid with numbers: then each square in the grid can be identified using two numbers.

After that the step to arithmetisation of continuous space is smaller but still non-trivial.

That invites various kinds of playing: e.g. what happens if you write into each box the sum of the two numbers that identify it?

Suddenly a deep link between colouring possibilities, diagonal moves and the difference between even and odd numbers becomes evident.

What enables a learner to realise that it does not matter how many squares there are, and that it works even if the grid has holes, like the odd slab of chocolate illustrated earlier.

You might suspect that if the colouring process went round a hole in the grid it might come round and be inconsistent with the starting layout. Why is this impossible?

What happens if you write the **difference** of the two numbers, into each box?

Try the **product** of the two numbers: is there anything interesting to be found in the resulting pattern?

Moving to the third dimension

Embedding the coin-sliding problem in a 3-D space allows a new version of the problem to become solvable.

A person who starts thinking about variants of the coin-sliding problem may notice partial analogies with other structures and processes.

If the problem is posed using a grid on a sheet of paper or other flexible material, then you might wonder what happens if you make a tube by joining two edges, e.g. top and bottom.

Two cases can occur, depending on the number of rows on the grid.

In some cases, diagonal moves across the join no longer preserve square-colour (parity).

As a result any configuration of coins could then be transformed into any other configuration using diagonal moves, but only because the 2-D grid is embedded in a 3-D space.

A logician could avoid using spatial intuition and map the tubular grid of squares into a purely logical structure, about which appropriate logical theorems could be proved, in the manner suggested by Bertrand Russell.

However, seeing the equivalence of the logical and the geometrical structures and processes would require use of geometrical insight.

Another option is to mark the grid on both sides of the sheet, and make a möbius strip.
See `http://en.wikipedia.org/wiki/M%C3%B6bius_strip`
`http://www.metacafe.com/watch/331665/no_magic_at_all_mobius_strip/`

A pattern in the preceding examples

We can see some common patterns in the preceding examples, which may help us design more human-like machines, help us understand better how humans and some other animals work, and perhaps even help us design far better educational strategies.

The pattern seems to be something like this:

- Competences are acquired that allow actions to be performed on objects in the environment.
- Mechanisms involving those competences require use of representations of the structures and processes.
- Those representations of possible occurrences can to some extent be created and manipulated independently of what is actually going on in the environment.
- Consequences of the manipulations can be used for predicting or explaining actual occurrences, or for planning new ones to achieve goals.
- The forms of representation can themselves become objects of play and exploration, sometimes with the aid of externalisations (e.g. diagrams).
- This can allow the representations acquired for different competences to be combined playfully and the consequences explored (with or without external aids).
- A meta-management architectural layer observing things that happen in play and in use can notice and store patterns that have some interesting feature.
- Often those patterns allow new problems to be solved.
- Sometimes trying to solve a specific problem also leads to discovery of a new and powerful pattern.

Logical patterns and spatial patterns

I have tried to bring out a deep similarity between spatial and logical forms of reasoning: in both cases relationships between patterns can be discovered that allow a single abstract representation to cover a wide range of phenomena in a way that allows predictions to be made in individual cases

But there is also a deep difference:

- Insofar as logical formulae are all Fregean (in the sense defined in Sloman 1971) the only way to alter a logical formula LF is to replace a function representation with another symbol, or an function-argument representation with another symbol, where the thing substituted may be more or less complex than the original, or to embed LF as an argument or function representation in a larger logical formula.
- In contrast, the ways of modifying and combining spatial representations (or spatio-temporal representations, when simulations are used) are far less constrained: many kinds of superposition, entanglement, adjoining, partial removal, extrusion, etc. can transform one structure to another, with many relationships, including causal relationships altered simultaneously.
- Moreover, the changes that are possible when a Fregean complex formula is modified are all discrete, whereas alterations of properties and relationships in a spatial or spatio-temporal representation can be continuous.

  There are techniques for accommodating an approximation to continuous change into a Fregean formula, namely by associating a numerical value with a property or relationship, and allowing the numerical value to undergo small discrete changes: this is how many graphical simulation programs work: but getting effects of these changes to propagate requires more work than in an analogical spatial representation (the “frame problem”).

Patterns versus Analogies and Metaphors

I have emphasised the discovery of reusable patterns: new abstractions that have wide applicability.
This is subtly and importantly different from discovery of analogies/metaphors.

A metaphor (or analogy) uses concepts from two domains and the notion of a mapping.

For example: the sequence of numbers 1 to 7 can be used as a metaphor for the days of a week, using the mapping

```
Sunday Monday Tuesday Wednesday Thursday Friday Saturday
One    Two    Three   Four      Five     Six    Seven
```

However instead of learning about and remembering all the many mappings between similar structures, you can create and store a new representation for the pattern that is common to all the "ends" of those mappings, something like this:

```
[ ( )    ( )    ( )    ( )    ( )    ( )    ( ) ]
```

This abstract specification (of an ordered collection of seven items), is applicable to very many ordered sets, and is also an entity that can be manipulated in its own right, along with other such entities, and relationships between them can be explored.

For example, you can discover different ways in which this structure can be constructed from smaller structures, and what you learn will be applicable to all the concrete ordered sets that the patterns can be applied to.

In short: reification is part of the power of mathematical thinking, and is more powerful than metaphor/analogy.

Tasks for the future

The ideas sketched above have been explored in restricted ways by some AI researchers, e.g. Doug Lenat (AM, Eurisko), Simon Colton (HR), Alison Pease (Modelling Lakatos).

This work has, as far as I know, been completely ignored by roboticists.

Likewise those researchers used forms of representation and reasoning that were not suitable for use by robots: e.g. they did not use mechanisms that could be part of a 2-D or 3-D visual perception system or action control system.

It remains an open question whether a geometric intuition mechanism (of the sort indicated in Sloman 1971) could be implemented in a virtual machine running on a von Neumann machine: I see no reason why not, but Penrose seems to think that is impossible.

See *The Emperor's New Mind: Concerning Computers Minds and the Laws of Physics*, OUP, 1989.

One of the many tasks remaining is to specify a kind of information-processing system that could discover new abstractions and represent them independently of the instances from which they are derived, as suggested in the previous slide; this well lead to a much better theory of concepts than the partly similar "prototype" theory of Rosch.

I am suggesting that a new synthesis, guided by investigation of tasks that arise naturally when a child or animal learns about the environment, may lead to major advances relevant to several disciplines.

However we also need to look at ways in which things can go wrong: they may be clues to the types of mechanisms and forms of representation that play a key role, and indicate some unavoidable flawed intermediate stages in the development of new competences, including new mathematical competences.

Some examples of buggy reasoning have been presented previously: another one follows, which could be part of a natural stage in the process of learning about space, time, motion, and measurements.

Reasoning as simulation: Simulating potentially colliding cars

The two vehicles start moving towards each other at the same time.

The racing car on the left moves much faster than the truck on the right.

Whereabouts will they meet – more to the left or to the right, or in the middle?

Where do you think a five year old will say they meet?

Five year old spatial reasoning

The two vehicles start moving towards each other at the same time.

The racing car on the left moves much faster than the truck on the right.

Whereabouts will they meet – more to the left or to the right, or in the middle?

One five year old answered by pointing to a location near 'b'

Me: Why?

Child: It's going faster so it will get there sooner.

What produces this answer:

- Missing knowledge?
- Inappropriate representations?
- Missing information-processing procedures?
- An inadequate information-processing architecture?
- Inappropriate control mechanisms in the architecture?
- A buggy mechanism for simulating objects moving at different speeds?

Partly integrated competences in a five year old

The strange answer to the racing car question can perhaps be explained on the assumption that the child had acquired some competences but had not yet learnt the constraints on their combination.

- If two objects in a race start moving at the same time to the same target, the faster one will get there first
- Arriving earlier implies travelling for a shorter time.
- The shorter the time of travel, the shorter the distance traversed
- So the racing car will travel a shorter distance!

The first premiss is a buggy generalisation: it does not allow for different kinds of 'race'.

The others have conditions of applicability that need to be checked.

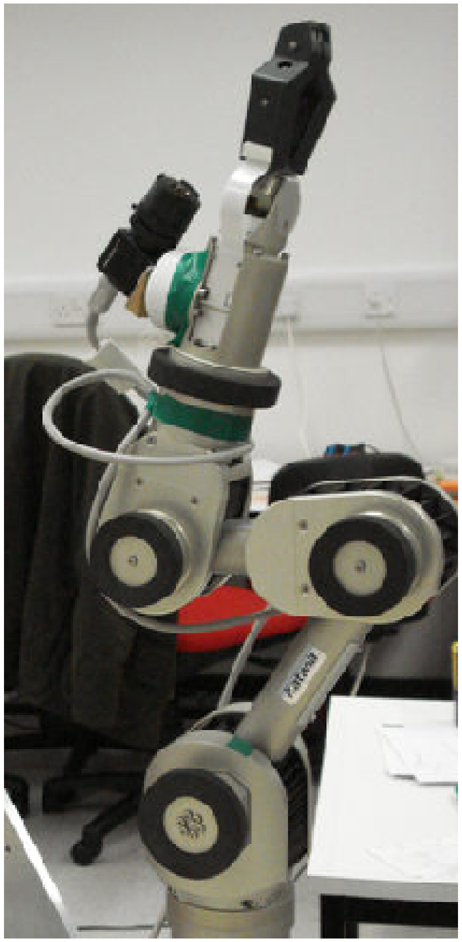
Perhaps the child had not taken in the fact that the problem required the racing car and the truck to be travelling for the same length of time, or had not remembered to make use of that information.

Perhaps the child had the information (as could be tested by probing), but lacked the information-processing architecture required to make full and consistent use of it, and to control the derivation of consequences properly.

Is Vygotsky's work relevant? Some parts of Piaget's theory of "formal operations?"

The CoSy PlayMate robot

Our robot has a camera on its wrist, looking past the gripper



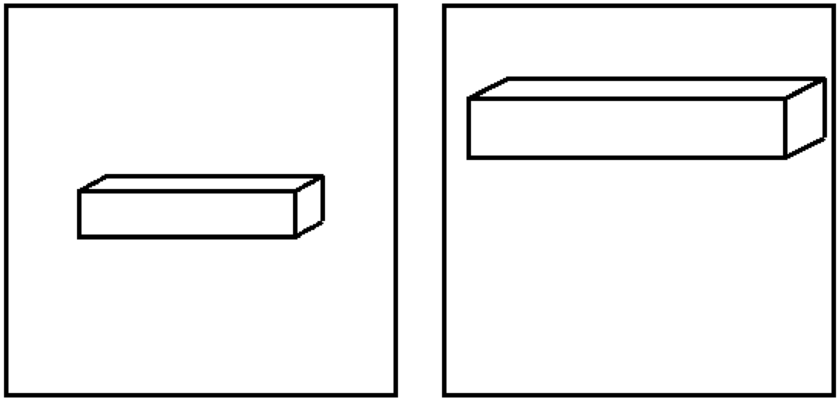
[SHOW TYPICAL PICTURES]

http://www.cs.bham.ac.uk/research/projects/cosy/photos/fleaCamPics.pdf

Epistemic affordances in grasping

The rigid relationship between eyes and mouth can be used to control motion towards an object to be grasped by biting.

The images represent two views as the eyes move down towards the object to be grasped by biting.
One of the images is taken when the gripper (i.e. mouth or beak) is still some way from the block to be grasped and the other is taken when the gripper is lower down, closer to the block.
Now, if the eye (or camera) is directly above the gripper is the gripper moving in the right direction?



An agent can use the epistemic affordance here by reasoning about the effects of its movements on what it sees and how the effects depend on whether it is moving as intended or not (e.g. in this situation aim higher).

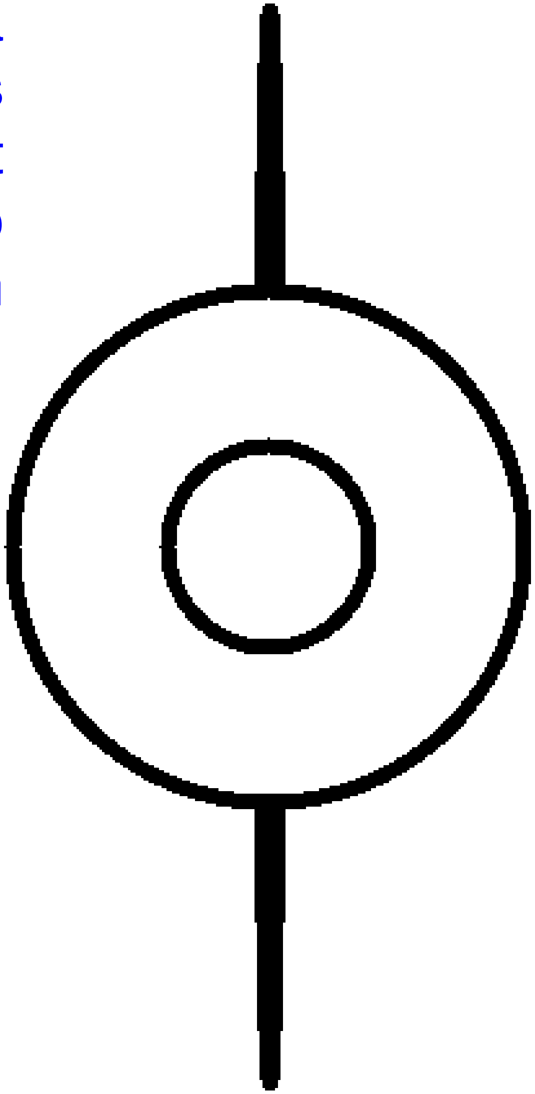
Instead of explicit reasoning using general knowledge about space and motion, an individual could simply be trained to predict how views should change if the target is being approached, and to constantly adjust its movements on the basis of failed predictions.

I.e. it can either use explicit knowledge, and the ability to reason about changing 3-D relationships applicable in varied situations, or implicit pattern-based reactive knowledge, produced by training, applicable only to situations that are closely related to the training situations.

Reactive pattern-based competence may work fast, but not generalise well.

Droodles can probe geometric competences

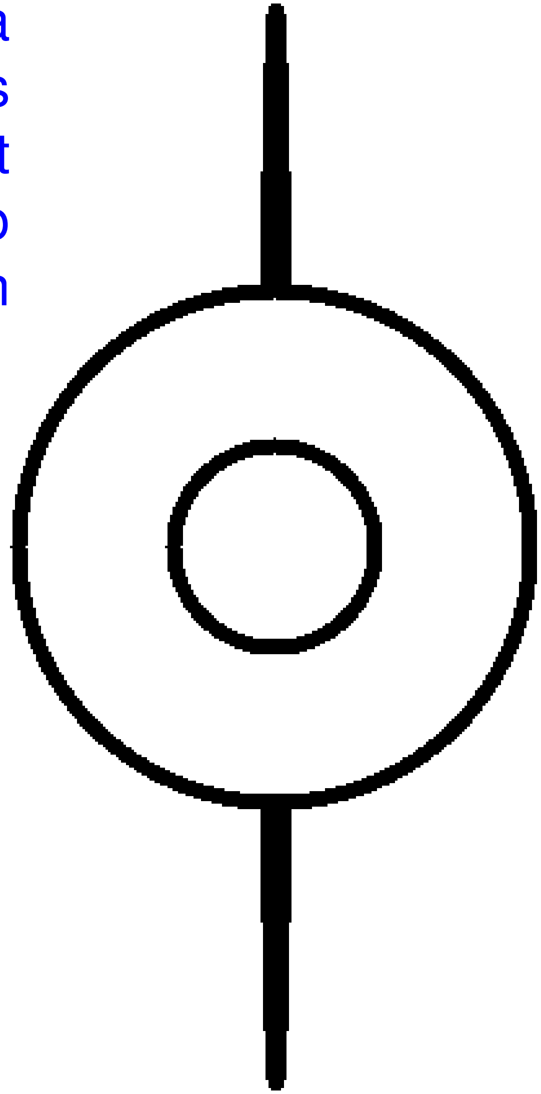
Many droodles work only because we have learnt a great deal about shapes of various familiar objects and we have also acquired the competence to project views of those shapes from different directions, or to infer viewing direction from projected shapes, as in this well known droodle.



What is it???

Droodles can probe geometric competences (2)

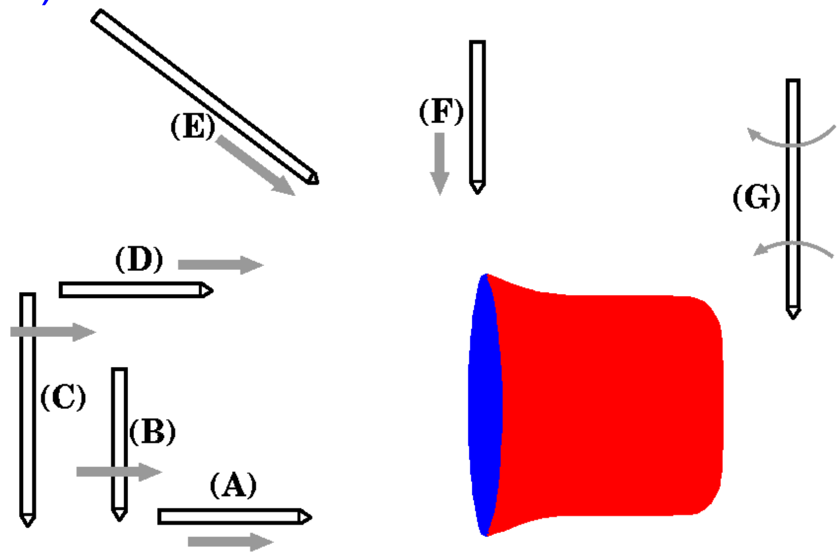
Many droodles work only because we have learnt a great deal about shapes of various familiar objects and we have also acquired the competence to project views of those shapes from different directions, or to infer viewing direction from projected shapes, as in this well known droodle.



A mexican riding a bicycle.

Uncertainty-reducing affordances

A very common problem in robotics is the uncertainty that comes from low resolution or noisy sensory input, or inadequate algorithms for interpreting sensor input (as in current machine vision).



The diagram shows various possible configurations involving a pencil and a mug on its side, along with possible translations or rotations of the pencil indicated by arrows.

Assume all the pencils lie in the vertical plane through the axis of the mug.

For each starting point and possible translation or rotation of the pencil, consider questions like:

- o Will it enter the mug?
- o Will it hit the side of the mug?
- o Will it touch the rim of the mug?

In some cases the answer is clear.

In cases where the answer is uncertain, because the configuration is in the “phase boundary” between two classes of configurations that would have clear answers we can ask how the pencil could be moved or rotated into a new initial configuration, to make the answer clear.

If pencil A moves horizontally to the right, will it enter the mug? If the answer is not clear, what vertical change of location of the pencil will make the answer clear?

If pencil G is rotated in the vertical plane about its top end will it hit the mug? If the answer is not clear what translations will make the answer clear?

Perceiving a scene can include perceiving possible ways of changing the epistemic affordances related to actions under consideration.

Two ways of dealing with uncertainty

Because uncertainty is so common in robotics, a vast amount of effort has gone in to ways of coping with uncertainty, including:

- Improving sensor quality
- Adding multiple sensors (e.g. multiple video cameras)
- Using different types of sensors (e.g. combining video cameras with laser range finders).
- Using sophisticated mathematics to compute probability distributions, and combining that with sophisticated decision-making algorithms to control actions.

A child or animal who is confronted with something uncertain, because of poor lighting, bad eyesight, dirty windows, occluding objects, distance of objects may not be able to adopt any of those engineering solutions.
(Except when it is possible to open the curtains or turn on a light.)

However a child can learn other ways of coping with uncertainty, by using the epistemic affordances in the environment to remove or reduce uncertainty.

That typically involves changing what you are doing rather than changing the way you process information.

So alter your heading to remove uncertainty about a collision, look from a different viewpoint, or move an object, or rotate an object to remove uncertainty about things that are occluded or self-occluded.

Selecting an appropriate strategy can often use geometric or topological reasoning, rather than manipulating probability distributions and expected utilities.

Multimodal sensorimotor ontologies are not general enough

Full human competence in a 3-D environment requires more than a somatic ontology based on patterns in input and output signals.

For some purposes an exosomatic ontology (of 3-D surfaces, objects, substances, motions, causal interactions, etc.) is required.

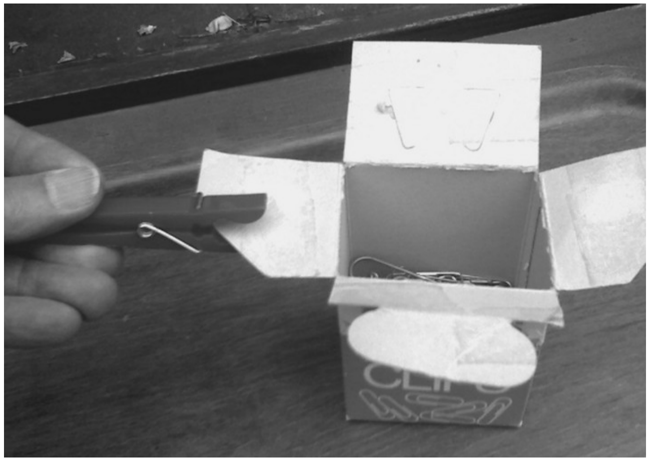
For more on this see

`http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0603`
Sensorimotor vs objective contingencies

`http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0601`
Orthogonal Recombinable Competences Acquired by Altricial Species

`http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0606`
Requirements for going beyond sensorimotor contingencies to representing what's out there (Learning to see a set of moving lines as a rotating cube.)

Abstraction to an amodal exosomatic ontology



Five different cases of grasping occur here, with very different realisations (projections) in the image plane: what is common to the different cases can be abstracted to a 3-D relationship between two facing surfaces and an object between them.

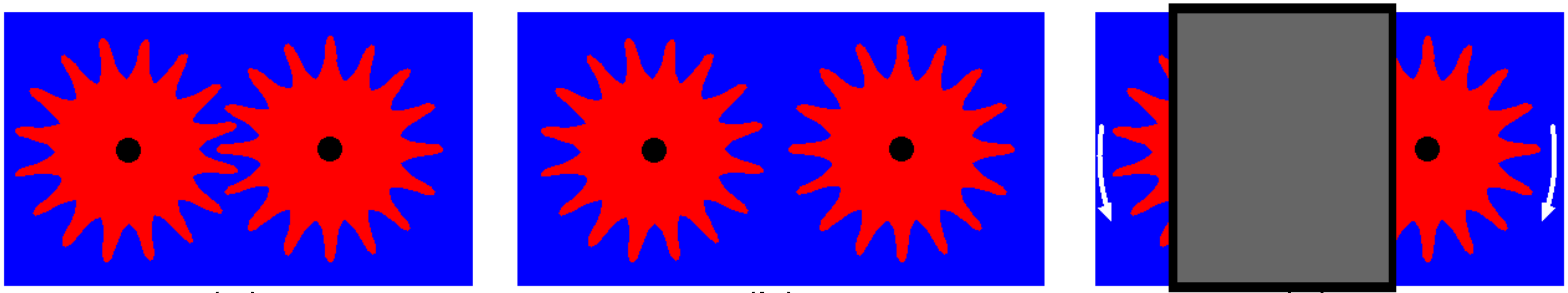
Instead of a somatic sensori-motor ontology referring to contents of sensory and motor signals, this uses an exosomatic amodal ontology referring only to things in the environment, i.e. outside the body.

Using the exosomatic ontology makes it possible to predict, control, and understand motions in far more varied situations, e.g. using the fact that if two surfaces come together firmly with an object between them, then when the two surfaces move the other object will move with them.

Using exosomatic ontologies requires an architecture that supports hypothetical reasoning about 3-D geometric relationships, and causal consequences of changes.

Adding non-geometric features

The ability to understand and predict what is going on can be enhanced if the exosomatic ontology goes beyond geometric and topological relations and processes, to include properties of materials, e.g. rigidity, impenetrability, weight, thermal properties, etc.



(a) (b) (c)

Assuming impenetrability and rigidity makes it possible in (a) to predict that whichever way the left wheel is turned the other will turn the opposite way. This uses geometric reasoning about movement of gears.

In case (b) the only way to predict is to use statistics/probability information from many observations.

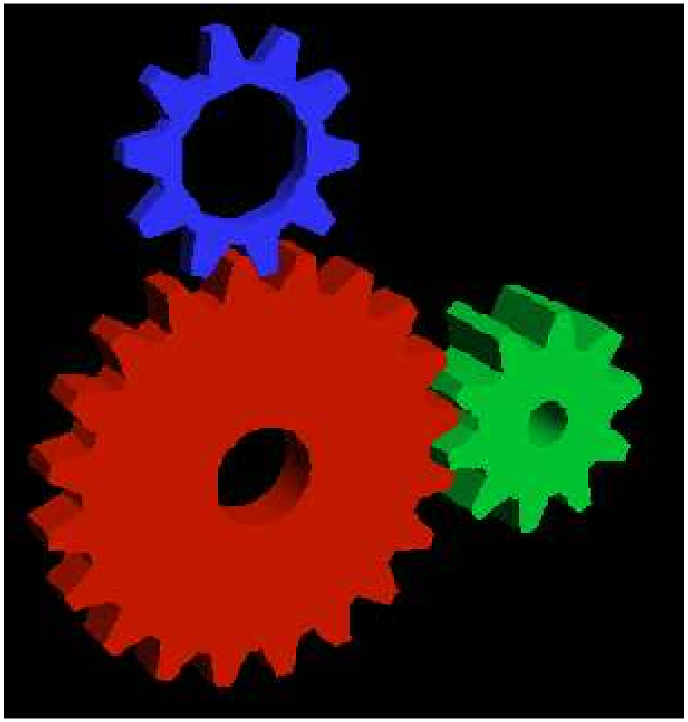
Case (c) shows that sometimes (if relative distances are used carefully) hidden mechanisms can be guessed on the basis of geometrical/causal reasoning, and used to explain what is observed.

Humean (nowadays Bayesian) causation based on collection of evidence of correlations, or statistics, is the only kind of causation available **when the mechanism is not understood.**

Kantian (structure-based, deterministic) causation is sometimes available, and is often more useful, but it requires a richer ontology and more general reasoning abilities.

Chains of causation

You can probably imagine various chains of causation by doing "What if reasoning" about this 3-D structure, with the initial causation being located in different places, including rotating, sliding, in various directions, etc.



Snapshot of the 'glxgears' program running on linux

E.g. if the small blue wheel moves towards you then starts rotating it may leave the other two unaffected, but not if it rotates where it is.

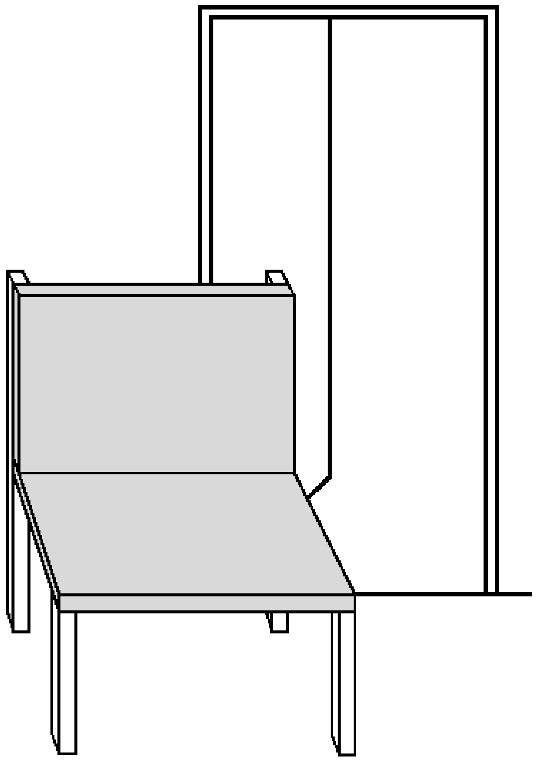
Our ability to represent different combinations of processes that are possible in any situation is rich and varied – but limited, and sometimes partly dependent on previous practice.

Carrying a chair through a door

Process fragments (proto-affordances) can be combined, in sequence or in parallel, in action or in hypothetical reasoning, to form new complex processes (actual or possible).

Affordances can interact in complex ways when combined, because of changing spatial relationships of the objects involved during the processes of performing the actions.

A large chair may afford lifting and carrying from one place to another, and a doorway may afford passage from one room to another, but attempts to combine the two affordances by lifting and carrying the chair to the next room may fail when the plan is tried.



A very young child may not be able to do anything about that, but an older child who has learnt to perceive the possibility of rotation of a 3-D object, may realise that a combination of small rotations about different axes combined with small translations some done in parallel, some in sequence. can form a compound process that results in the chair getting through the doorway.

Is any other type of animal is capable of understanding that?

Even the very familiar process of grasping an object is a complex combination made of various successive sub-processes, and some concurrent processes, with concurrently changing relationships between different parts of the surface of the object and different parts of the grasping hand.

Problem: What needs to be added to traditional AI planners to enable them to construct plans involving such continuous, concurrent, interacting processes?

Interacting processes

Processes that occur close in space and time can interact causally in a wide variety of ways, depending on the precise spatial and temporal relationships and constraints.

It is possible to learn about the consequences of such interactions by observing them happen, but humans and some other animals sometimes need to be able to consider and work out consequences of possible combinations that they have never previously observed.

The ability to think about and reason about novel combinations of familiar types of processes is often required for solving new problems.

One source of fallibility of mathematical generalisations about interacting spatial structures is the fact that whatever space encloses those processes could, in principle also contain something else that interferes with their normal consequences.

Thus the necessity in such causation, and the validity of spatial mathematical reasoning is always conditional, but often we don't understand the conditions well enough to formulate them apart from the nearly vacuous *ceteris paribus* (other things being equal).

Perhaps slightly better: *provided nothing else intervenes*.

Don't forget Lakatos.

An example from recent history of science

The development of the theory of plate tectonics was extremely important in unifying a collection of theories and observations in physics, chemistry, materials science, geology, geography, meteorology, and evolution: the total picture arising as a result enables us to view the earth as something like a computer, involving:

- different levels of explanation, with causal interactions between them;
- some levels close to the physical sciences – using geometry, topology, mechanics;
- other levels involve information processing
  (at the levels of individual organisms, species, and ecosystems);
- coexistence of hardware and software virtual machines, as in a computer.
  (one way to interpret the Gaia hypothesis).

The ability of scientists to understand and think about plate tectonic theory used information processing mechanisms partly similar to our ability to think about a wide armchair moving through a doorway: there are concurrent interacting 3-D processes.

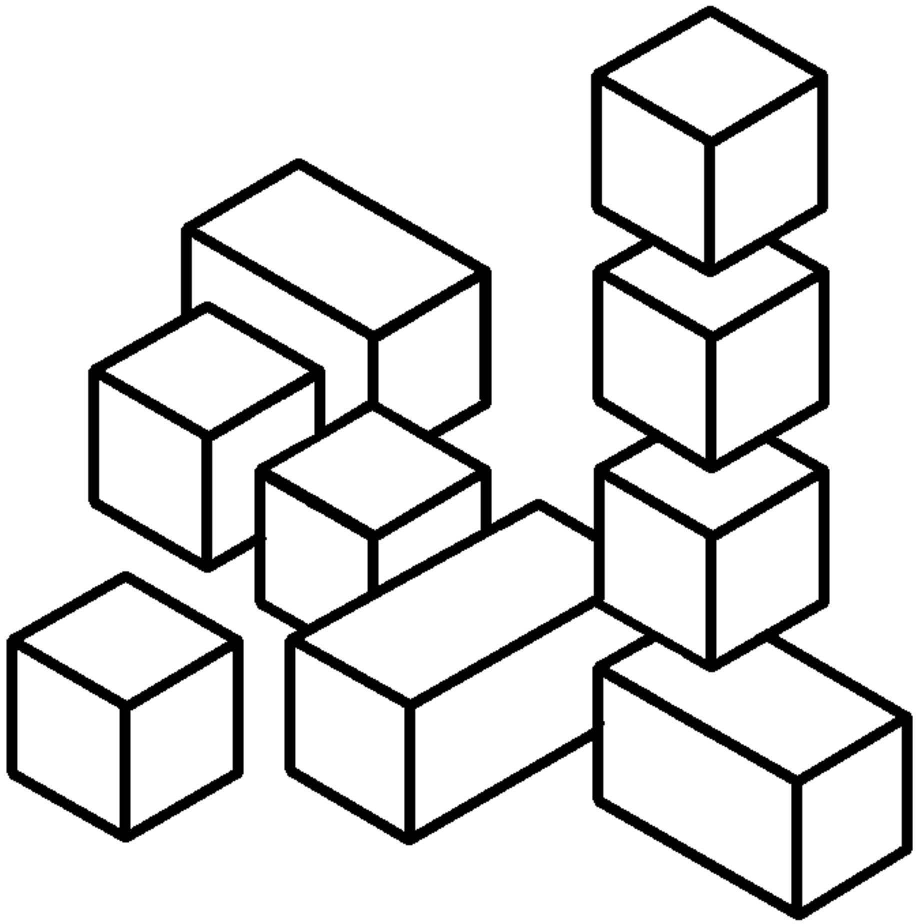
This geometrical/topological form of representation of and reasoning about continuous interacting processes would be hard to express in an effective way using only logical and algebraic formalisms.

See `http://www.ucmp.berkeley.edu/geology/tectonics.html`
Also `http://www.bbc.co.uk/radio/aod/radio4_aod.shtml?radio4/inourtime`

How we represent 3-D structures and processes is problematic

We can see a 3-D configuration of cubes

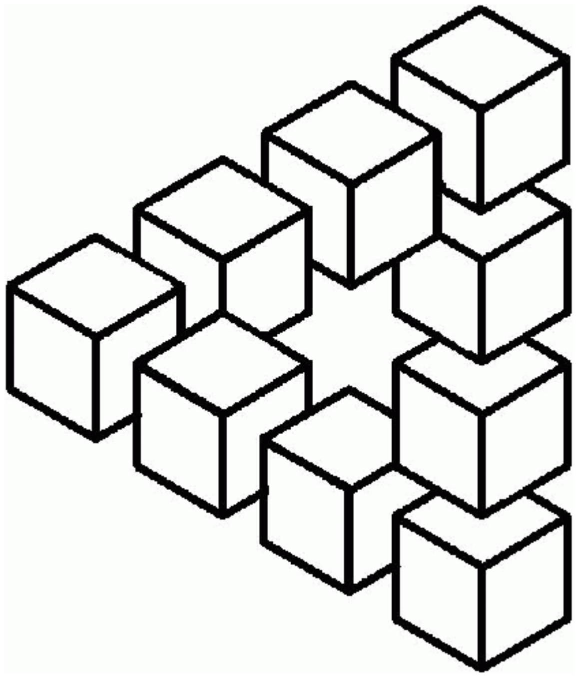
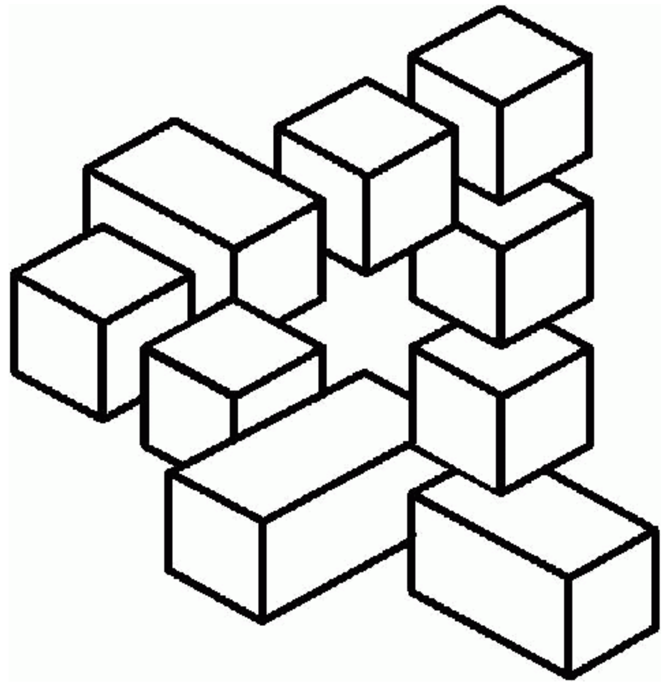
LIKE THIS:



You could build a model of what you see.

But what happens if we rearrange the cubes, ... ?

Like this? or Like this?

Given a pile of cubes, could you build configurations like these?

These examples – inspired by Oscar Reutersvärd (1934)[*] – show that 3-D perception does not involve building internal objects that are isomorphic with the things seen.

That's not at all surprising from the standpoint of Sloman 1971, which argued that analogical representations need not be isomorphic with what they represent. (See next slide)
`http://www.cs.bham.ac.uk/research/projects/cogaff/04.html#200407`

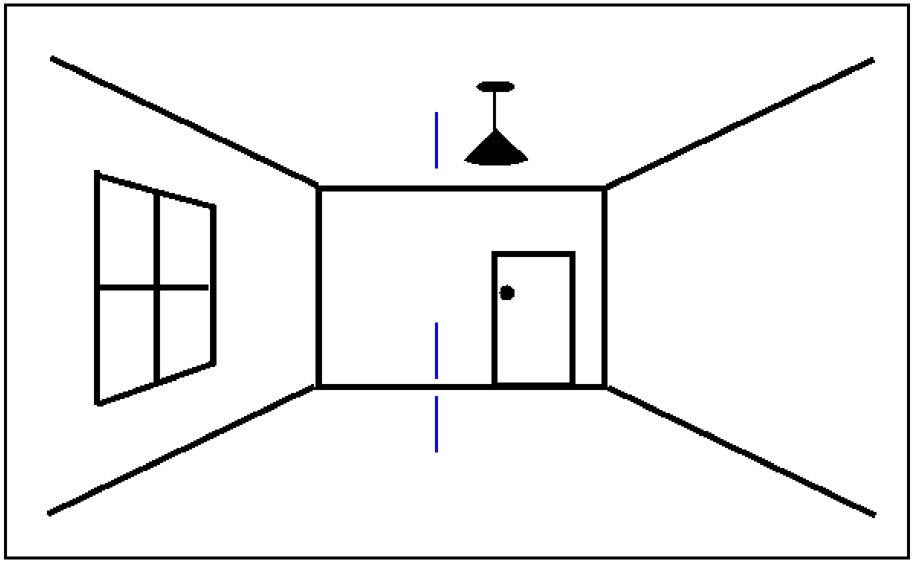
[*] `http://www.sandlotscience.com/EyeonIllusions/Reutersvard.htm`

Isomorphism is not needed

Here's a modified version of a picture from chapter 7 of *The Computer Revolution in Philosophy*, also in the 1971 IJCAI paper.

Objects and relations within a picture or optic array need not correspond 1 to 1 with objects and relations within the scene, as is obvious from 2-D pictures of 3-D scenes.

For example: pairs of points in the image that are the same distance apart in the image can represent pairs of points that are different distances apart in 3-D space – e.g. vertically separated points on the walls, and horizontally separated points on the floor and ceiling. (And *vice versa*.)



Some pairs of parallel edges in the scene are represented by parallel picture lines, others by converging picture lines.

The small blue lines can be interpreted in different ways, with different spatial locations, orientations and relationships. On each interpretation the structure of the image remains unchanged, but the structure of the 3-D scene changes. Compare the Ames room and the Ames Chair

Isomorphism is not needed for pictures of objects to be useful.

In the case of pictures of 3-D impossible objects, whatever the interpretation is cannot be isomorphic with what it refers to, since nothing can be isomorphic with an impossible object. (Why not?)

Potentially inconsistent fragments

The crucial point is that the result of making sense of perceptual input is neither some sort of sensory copy of the stimulation pattern (e.g. a bitmap), nor an isomorphic model of what is taken to exist in the environment, but a collection of re-usable separate items of information about things, surfaces, processes, relationships, and possibilities in the environment derived from the sensory input (often using prior knowledge).

- In the 1960s it was thought by some that a major result of perception would be some sort of "parse tree" or "parse graph" for images and scenes, based on a grammar for spatial structures. E.g. S. Kaneff (ed) *Picture language machines* 1970.
- But the proposed grammars were often very arbitrary, and the approach did not lead to good ways of representing information about 3-D structures and processes for a robot to use.
- An aspect graph, linking hypothesised views to actions that will lead to those views being seen has more 3-D information, implicitly, and allows partial ignorance to be represented.
- These ideas need to be generalised to allow more kinds of actions and more kinds of consequences to be represented, in addition to structural information.
- The form of representation needs to be capable of being driven by sensory input, and also by hypothesised future actions, e.g. in planning or predicting.
- Results of such manipulations need to be accessible for reasoning.

Mr Bean's underpants

This paper (from a conference on thinking with diagrams in 1998)

`http://www.cs.bham.ac.uk/research/cogaff/00-02.html#58`

discusses how we can reason about whether Mr Bean (alias Rowan Atkinson) can remove his underpants without removing his trousers.

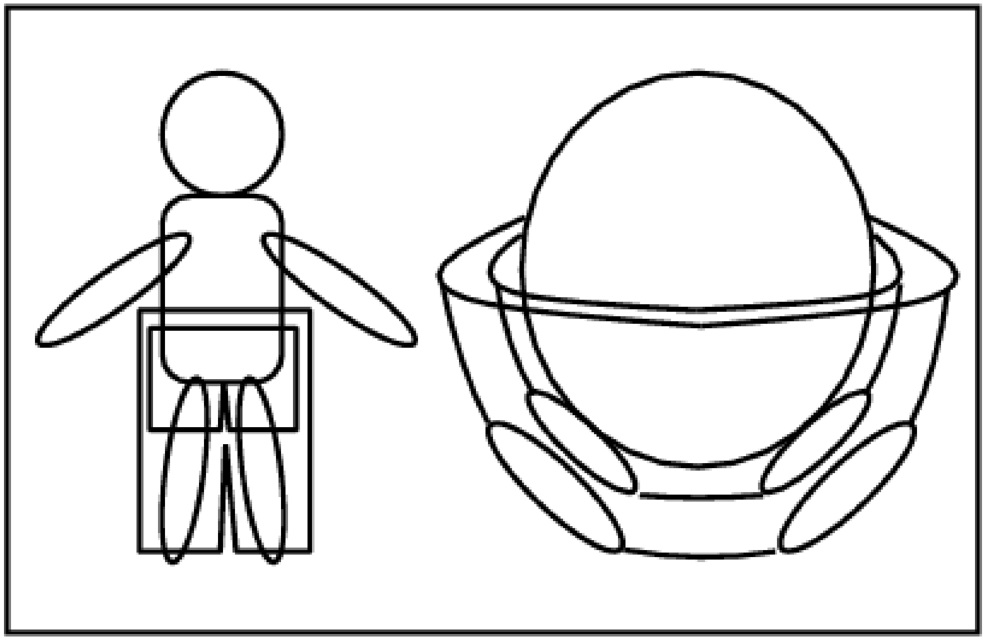
People often don't see all the possibilities at first.

The paper discusses how changing the simulation to a topologically 'equivalent' one can help us count the possible ways to perform the task.

Children can learn to perform such actions (as party tricks) physically long before they can reason with the mental simulations.

What changes as the simulation ability develops?

In part it seems to require an introspective ability to understand the nature of the simulations we use.



For more on learning about topology, see:

Jean Sauvy & Simonne Sauvy *The Child's Discovery of Space, From Hopscotch to Mazes: an Introduction to Intuitive Topology* (Translated P.Wells 1974).

Seeing structure, motion, and invariants in mathematics

- Hume is often interpreted as claiming that all knowledge is either analytic (i.e. true by definition and essentially empty), or empirical, requiring experiment and observation for its confirmation, and therefore capable of turning out false in new situations.
- Kant thought there were counterexamples, especially in mathematical knowledge, which he claimed was synthetic, i.e. amplified our knowledge, and non-empirical (or *a priori*), i.e. immune from empirical refutation.

  (Section C of Franklin's chapter on 18th Century philosophy of mathematics, mentioned later, suggests that Hume's view of mathematics is closer to Kant's than normally supposed.)
- My Oxford D.Phil thesis (completed in 1962, never published, but now online) was an attempt to defend Kant against Hume, but, like Kant, I did not have adequate conceptual tools for the job. We are a little closer now, insofar as we may in future be able to design working models of how a mathematician uses mechanisms that are needed for perception of and thinking about complex structures can be deployed in making mathematical discoveries, including seeing why 7 + 5 must always be 12 (Kant's example – discussed below).
- I now try to show that this is connected with our understanding invariant properties of one to one mappings, which most people can visualise in terms of spatial connection, even though the mathematical notion is far more general and not restricted to spatial objects.
- A child learning to count eventually has to understand all this, in order to understand what numbers (at least the positive integers treated as cardinal numbers) are, and what mathematical truths are. Unfortunately their teachers may be too confused to help children who do not discover these things spontaneously. (Understanding numbers is NOT a matter of being able to "subitize".)
- When we go beyond the positive integers things get far more complex in ways that very few people understand, alas, so they just learn rules of thumb that work – their minds remain partly underdeveloped for life. (This is true of all of us, in some respects.)

Discovering facts about counting

Not all mathematical discoveries are based on visual reasoning.

For example, a very different collection of discoveries, some of them documented in Chapter 8 of (Sloman 1978), occurs as child learns to count, and then discovers different uses for the counting process and different features of the counting process, such as the fact that the result of counting a collection of objects is not altered by rearranging the objects but can be altered by breaking one of the objects into two objects.

This kind of mathematical discovery depends on perceiving structures and relationships in procedures as they are followed, like seeing the pattern in chocolate slicing.

For example, a child who has learnt to count may discover that in order to work out the size of the set formed by combining a collection of M objects and a collection of N objects all it has to do is recite N numerals after M.

(E.g. reciting three numerals after “five” gives “six, seven, eight”.)

At first the child may discover empirically that this produces the same result as counting five numerals after three, but at a later stage this can lead to the non-empirical understanding why it must be so.

Part of this is discovering that any one-to-one mapping between elements of two finite sets can be converted into any other by successive changes of the correspondences.

KANT'S EXAMPLE: 7 + 5 = 12

Kant claimed that learning that 7 + 5 = 12 involved acquiring *synthetic* (i.e. not just definitionally true) information that was also not *empirical*.

I think that's because he saw its connection with what I've been saying about discovering properties of structures and processes.

You may find it obvious that the equivalence below is preserved if you spatially rearrange the twelve blobs within their groups:

```
ooo        o      oooo
ooo  +     o  =   oooo
o        ooo      oooo
```

Or is it?
How can it be obvious?
Can you *see* such a general fact?
How?

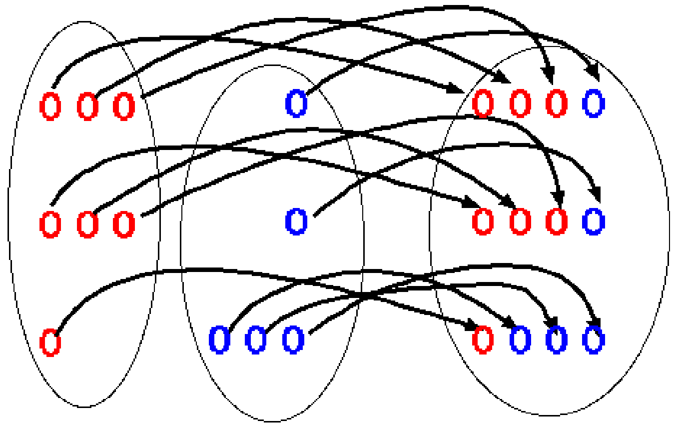
What sort of equivalence are we talking about?

I.e. what does "=" mean here?

Obviously we have to grasp the notion of a "one to one mapping".

That can be defined logically, but the idea can also be understood by people who do not yet grasp the logical apparatus required to define the notion of a bijection — if they have a way of thinking about the consequences of motion of the blobs.

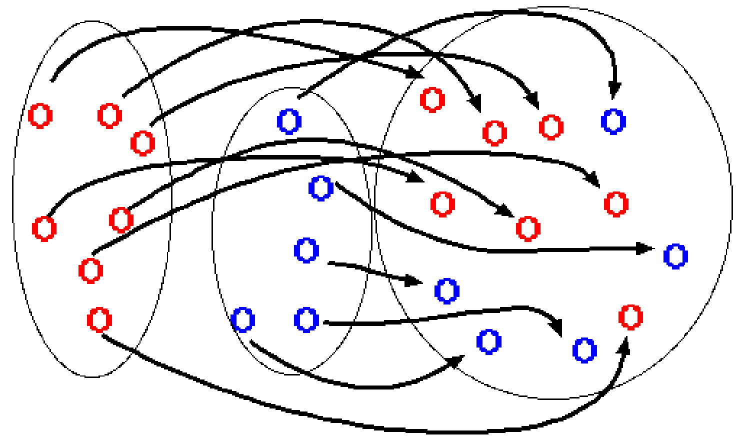
SEEING that 7 + 5 = 12



Join up corresponding items with imaginary strings.

Then rearrange the items, leaving the strings attached.

Is it 'obvious' that the correspondence defined by the strings will be preserved even if the strings get tangled by the rearrangement?



Is it ‘obvious’ that the same mode of reasoning will also work for other additions, e.g.

777 + 555 = 1332

Humans seem to have a ‘meta-level’ capability that enables us to understand why the answer is ‘yes’. This depends on having a model of how our model works – e.g. what changes and what does not change if you add another pair of objects joined by a string, or if you swap locations of one end of two pieces of string.

Compare the reversed ordering example given previously.

Another way to think about 7 + 5 = 12

OBJECTION: the emphasis on those diagrams with groups of dots misses the point, because it does not address the link between numbers and counting operations – i.e. associating elements of a set not with elements of some other arbitrary set, but with an initial sequence of numerals.

Here are three initial sequences of number names:

```
A: one two three four five
B: one two three four five six seven
C: one two three four five six seven eight nine ten eleven twelve
```

Then "5 + 7 = 12" could be taken to state that if there are two sets of objects Sa and Sb, and elements of Sa can be mapped one to one onto elements of A, and elements of Sb can be mapped one to one onto elements of B:, then combining Sa and Sb will give a new set Sc, whose elements can be mapped one to one onto elements of C:, and no other initial segment of the number sequence.

One way to demonstrate that is combine A: and B: and note that the combination maps one to one to C:

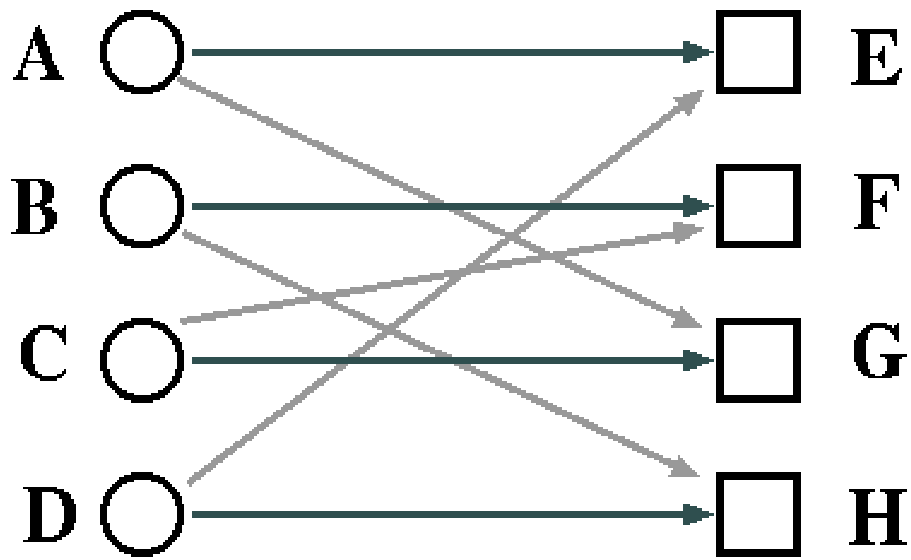
```
one two three four five one two   three four five six    seven
 |   |    |    |    |    |   |      |     |    |    |      |
one two three four five six seven eight nine ten  eleven twelve
```

That then leaves only the problem of checking that being in a one to one correspondence is a transitive relation between sets – demonstrated by showing that two short links can always be combined to form one long link, joining the first element of the first link to the second element of the second link. (Demonstrable graphically)

Converting mappings

A child may discover empirically a strategy for converting one mapping to another and implicitly understand that it will always work, without necessarily being able to articulate the strategy nor explain why it works.

This depends on the architecture allowing one process to observe that another process has some consequences that do not depend on the particularities of the example.



A one-to-one mapping from one set of objects to another (e.g. the grey arrows) can be converted to any other such one-to-one mapping (e.g. the black arrows) by swapping ends on one side, two at a time.

E.g. the right hand ends of the grey arrow from A to G and the grey arrow from D to E can be swapped, then the right hand ends of arrows from B to H and from C to F, etc. gradually eliminating differences between grey and black mappings.

Discovering that any one-to-one mapping between elements of two finite sets can be converted into any other by successive changes can make use of simultaneous perception of spatial and temporal relationships.

Formulating the general algorithm is left as an exercise. Could a robot do this?

I don’t think this is how children normally come to understand the invariance: What alternatives are there?

Extending mappings

Suppose that you have checked, by counting or by using strings to link pairs of objects, that there is a one-to-one mapping between apples in box A and bananas in box B.

I.e. apple a1 is paired with banana b1, a2 with b2, etc. and apple aN with banana bN.

Now suppose that someone else chooses a subset of the apples and a subset of the bananas and sets up a one to one mapping between the subsets.

E.g. apple a1 is now paired with banana b9, a2 is paired with b3, and 5 is paired with b7. with banana bN.

No apple or banana has been added to or removed from either box, and they are all intact: can you be sure that that partial mapping of elements of A with elements of B can be extended to a **full** one-to-one mapping between the two sets?

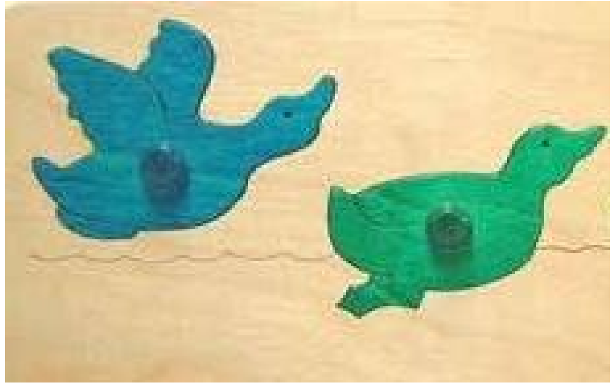
I suspect very few people ever explicitly articulate the fact that if there exists a one-to-one mapping between two sets, then any one-to-one mapping between subsets of the sets, whether consistent with the original mapping or not, can be extended to a new full mapping between the whole sets: nevertheless they often presuppose special cases of that fact.

E.g. if you have counted out the right number of forks for your dinner guests it does not matter in which order you hand them out.

Does that also apply to **infinite** discrete sets?

Why not?

Pre-mathematical discoveries



There are developments that would not normally be described as mathematical, yet are closely related to mathematical competences.

For example a very young child who can easily insert one plastic cup into another (of the sort shown in the figure may be able to lift a cut-out picture from its recess, and know which recess it belongs to, but be *unable* to get it back into the recess: the picture is placed in roughly the right location and pressed hard, but that is not enough.

The child apparently has not yet extended his or her ontology to include boundaries of objects and alignment of boundaries.

Some time later the child copes easily.

How such extension of competences happens is not at all clear, but what has to be learnt, namely facts about boundaries and how they constrain possible movements, is something that can be studied mathematically, and might be so studied later.

Specialised mathematical education builds on general abilities to see structures and processes and see how some structures can constrain or facilitate certain processes, including processes of information acquisition.

High level perceptual processes can ignore low-level details

- I am suggesting that when we watch or imagine things moving we simulate the motion (i.e. we create and run representations) at different levels of abstraction.
- Some of them we probably never become conscious of as they are used only in relatively automatic control of common processes, for instance as optical flow patterns are used in posture control.
- What we say we are conscious of is often closely related to what we can report, to ourselves or to others, and that will typically be things happening at a high level of abstraction, that are relevant to our current goals and needs, though we can direct our attention to details just for the sake of examining details, and we can become aware of details that are too rich and complex to be reported, even to ourselves, e.g. watching swirling rapids in a fast flowing river or hundreds of leaves stirring in the wind.
- What we are conscious of seeing may depend on what the current task is, and sometimes we do not notice details even if a low level system processes them – e.g. because what we attend to when answering a question includes only the contents of the more abstract simulations.
- But that does not mean that the details have not been processed, as I have shown elsewhere: one of your subsystems concerned with posture-control may be conscious of optical flow even when you are not.

Development of perceptual sub-systems

The ability to run simulations while seeing is not static, and may not even exist at birth:

- Visual capabilities described here develop in part on the basis of developing architectures for concurrent simulations and in part on the basis of learning new types of simulation, with appropriate new ontologies and new forms of representation.
- The initial mechanisms that make all of this possible must be genetically determined (and there may be limitations caused by genetic defects).
- But the *contents* of the abilities acquired through various kinds of learning are heavily dependent on the environment – physical and social, and on the individual's history.
  Some innate content is needed for bootstrapping.
- For instance someone expert at chess or Go will see (slow-moving!) processes in those games that novices do not see.
- Expert judges of gymnastic or ice-skating performance will see details that others do not see.
- An expert bird-watcher will recognize a type of bird flying in the distance from the pattern of its motion without being able to see colouring and shape details normally used for identification.

A deeper theory would explain the variety of types of changes involved in such developments: including changes in ontologies used, in forms of representation, and perhaps also in processing architectures.
These will be changes in virtual machines implemented in physical brains.

Unusual developmental trajectories

I have not discussed, though I am very aware of the problem of whether and how this discussion generalises to children born blind, or limbless, or with other deficiencies, some there from birth, some resulting from early trauma (e.g. Helen Keller).

How can a congenitally blind child grow up to be a mathematician?

How can a child born without limbs grow up to have a normal understanding of spatial structures and processes?
(See Alison Lapper's web site.)

I suspect that a key part of the answer is that even children born with these deficiencies still have parts of the brain that evolved to work with a full complement of human sensors and effectors.

So even though those brain mechanisms normally develop through a certain trajectory making use of ordinary sensing and ordinary acting, the flexibility that allows them to work in a very wide variety of environments (like children who now learn to use a mouse to play a computer game – something none of their ancestors ever did), also allows them to grow and develop without a full set of normal interfaces to the environment.

HOW this happens is no doubt a complex and difficult story.

If this is right a great deal of what has recently been written about the role of embodiment in cognition is probably badly misguided.

It's your ancestors who need to be embodied, not you.

Links with becoming a philosopher

Many examples of the ability to do mathematics seem to depend on an information-processing architecture that supports concurrent processes including processes that monitor other processes.

Much of philosophy depends on our ability to notice things we do and do not do, and to become puzzled by certain features of those processes.

However, the self-observation provided by an information-processing architecture that includes metamanagement cannot provide full detail, and may be limited in various ways, including not having a good ontology for describing complex information-processing of the sort that goes on in itself. (Neither do most humans.)

Likewise no robots that have any self observation capability at present, have nearly as much knowledge about what is going on in their virtual machines as the designers (when the program is working properly).

It is also the case the knowledge acquired about the environment often has features that can support arguments that they have suspect credentials.

This can also lead to (sometimes misplaced) philosophical puzzlement and investigation of alternative explanations.

Things to do

There is still much to do, and many topics to discuss, including:

- The variety of extrapolations to limiting cases, e.g. infinite discrete sequences, infinitely long lines, infinitely large areas, infinitely thin lines, infinitely small points, infinitely dense textures,...
- Discuss relevance of Gödel's incompleteness theorem, and how one can talk about the intended model for the natural number series (if a human can have one, why not a machine?)
- Many issues to do with continuity
- Extending the notion of number from discrete, countable, sets to amount of something that can vary continuously, eg. length.
- How can a child come to understand the notion of half the area or volume of an asymmetric spatial region or volume.
- How to extend the idea of number to a measure of an arbitrarily shaped area: the importance of rectangular grids and the limiting case as grid size shrinks.
- Using finite spatial structures to represent infinite sets and infinite ordinals.
- Using what you can imagine to help you imagine what you can't imagine.
- What to think about Euclid's parallel axiom: is there some way of constructing a pair of straight lines that forces them to go on indefinitely exactly the same distance apart?
- Does the construction come unstuck before grids of lines with different orientations are considered?
- Need to go back to the elastic sheet proof of Euler's theorem: what mechanisms would enable a robot to imagine the process of stretching a polyhedron's surface flat?

Unanswered questions

The form of representation, the mechanisms for manipulation, and the architecture for combining the various information-processing components of an intelligent individual are still barely understood.

A brave attempt at theory construction can be found in

Arnold Trehub, 1991, *The Cognitive Brain,*

`http://www.people.umass.edu/trehub/`.

The retinoid theory seems to be only a partial model, though richer than many others.

The work of Eric Baum may also be relevant, and his approach (looking closely at how humans solve particular problems) overlaps with what I have been doing.

Eric Baum's web site `http://www.whatisthought.com/eric.html`

"A Working Hypothesis for General Intelligence" 16 pages Draft October, 2006.

`http://www.whatisthought.com/working.pdf`

There is probably a lot of other relevant work that I don't know about or have forgotten (and may be unwittingly reproducing!).

We may be able to move towards a design specification if we study and analyse more and more examples in order to work out detailed information-processing requirements, which may lead us to features that may suffice to explain the desired behaviours.

A presentation on the exploration of "kinds of stuff" in infants and toddlers is here (still under development)

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#babystuff`

Assembling bits of stuff and bits of process, in a baby robot's world: A Kantian approach to robotics.

In conclusion

- I have tried to identify an array of features of normal perception, action, learning, reasoning, control, planning, and explaining which seem to be products of complex developmental processes influenced by both evolution and the current environment (including other humans), and which also are able to play a role in generating mathematical explorations and supporting mathematical reasoning.
- Making this more precise and detailed will require considerably extending the state of the art in robotics and AI, and giving robots new ways of representing and reasoning about spatial structures and processes, as well as giving them architectures that support self-observation of a kind that drive new learning and developments.
- I suspect that as this research programme develops it will provide new, testable, insights into what goes on when children (and adults) learn many basic aspects of mathematics: understanding the mechanisms and the processes could at last replace educational policies based on deep explanatory science rather than shallow empirical results.
- These mechanisms required for intelligently coping with the environment, including other intelligent individuals, can, as some science fiction writers have pointed out, produce both philosophical activities, and when they become really buggy, even theological activities.

  (Isaac Asimov: “Reason” in *I Robot*)

WARNING

This presentation should not be treated as an authoritative account of the ideas of any of the thinkers mentioned herein.

My summaries of the views of others are often interpretations of things I have read recently or things I remember reading in the past. I am interested in them as theories that could be held by serious thinkers, and don't really care who did or did not hold them.

Anyone seriously interested in the views of Kant, Hume, Frege, Russell, Feynman, etc. should study their writings instead of relying on my summaries.

A useful overview of ideas about the nature of mathematical discovery and mathematical reasoning around the time of Kant can be found in

James Franklin, Artifice and the Natural World: Mathematics, Logic, Technology,
in *Cambridge History of Eighteenth Century Philosophy*, Ed. Knud Haakonssen
Cambridge University Press, 2 volumes 1423 pp. 2006

Franklin's chapter is available online without footnotes here:

`http://web.maths.unsw.edu.au/~jim/18c.html`

It is interesting that he too notes the possibility of testing ideas in philosophy of mathematics by implementing working intelligent systems:

> Hume's views on inference are seen to better advantage if they are thought of not in terms of formal logic, or even introspection, but as a research proposal to be implemented in, say, silicon chips. Modern Artificial Intelligence, like most eighteenth century writing, is concerned with the implementation of a system of inference, not just the formal structure of the system itself. From that point of view, it is necessary to answer questions that do not arise in formal logic, such as how the symbols become attached to the things they mean.

On the last point see also this critique of symbol grounding theory:
`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models`

See his survey of philosophies of mathematics: `http://web.maths.unsw.edu.au/~jim/philmathschools.html`

Related online papers and presentations

Functions of vision, with speculations about the role of multiple multistable dynamical systems:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#compmod07

Discussion paper on Predicting Affordance Changes (including epistemic affordances)
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0702

For a critique of the notion that reduction to logic can explain mathematical knowledge (similar to Frege's argument) see my DPhil thesis (Oxford, 1962), e.g. Appendix II
http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#706
Knowing and Understanding: Relations between meaning and truth,
meaning and necessary truth, meaning and synthetic necessary truth

The arguments were partly replicated here:
http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#712
Explaining Logical Necessity *Proc. Aristotelian Society*, 1968/9, Vol, 69, pp 33–50.

Then linked to AI theories of representation in 1971
http://www.cs.bham.ac.uk/research/projects/cogaff/04.html#200407

Computational Cognitive Epigenetics (With J. Chappell in BBS 2007):
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0703
Commentary on Jablonka and Lamb (2005)

See other papers and presentations with J. Chappell here:
http://www.cs.bham.ac.uk/research/projects/cosy/papers/

Our presentations on causation at WONAC, Oxford June 2007:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac

An argument that internal generalised languages (GLs) preceded use of external languages for communication, both in evolution and in development:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang
What evolved first: Languages for communicating, or languages for thinking
(Generalised Languages: GLs) ?

More

Diversity of Developmental Trajectories in Natural and Artificial Intelligence (AAAI Fall symposium 2007):

`http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0704`

The importance of virtual machines and supervenience:

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#bielefeld`

Ideas about how machines or animals can use symbols to refer to unobservable entities (why symbol grounding theory is wrong)

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models`

Introduction to key ideas of semantic models, implicit definitions and symbol tethering

Additional papers and presentations:

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/`
`http://www.cs.bham.ac.uk/research/projects/cosy/papers/`
`http://www.cs.bham.ac.uk/research/projects/cogaff/`

Other useful references

James Franklin, Adrian Heathcote, Anne Newstead: 'MANIFESTO: Mathematics, the science of real structure: An Aristotelian realist philosophy of mathematics,'
http://web.maths.unsw.edu.au/~jim/manifesto.html

James Franklin, Last bastion of reason New Criterion 18 (9) (May 2000), 74-8. (Suggests Lakatos is dishonest about mathematics.)
http://newcriterion.com:81/archive/18/may00/lakatos.htm

Teri Merrick, 'What Frege Meant When He Said: Kant is Right about Geometry',
Philosophia Mathematica Vol 14, 1, 2006. doi:10.1093/philmat/nkj013

Lisa A. Shabel, "Kant's Philosophy of Mathematics" in The Cambridge Companion to Kant,CUP, 2006.
http://people.cohums.ohio-state.edu/shabel1/cv.html

Jeanette Wing on 'Computational thinking' *CACM* 49,3, March 2006, pp. 33-35.
http://www.cs.cmu.edu/afs/cs/usr/wing/www/publications/Wing06.pdf

Anette Karmiloff-Smith *Beyond Modularity*. See the BBS precis:
http://www.bbsonline.org/documents/a/00/00/05/33/index.html

Douglas Hofstadter (Thanks to Dan Ghica for the link.)
http://www.stanford.edu/group/SHR/4-2/text/hofstadter.html
"on seeing A's and seeing As", *Stanford Humanities Review* Volume 4, issue 2

Mateja Jamnik *Mathematical Reasoning with Diagrams: From Intuition to Automation* (CSLI Press 2001)
http://www.cl.cam.ac.uk/~mj201/research/book/index.html

Daniel Winterstein *Using Diagrammatic Reasoning for Theorem Proving in a Continuous Domain*
PhD thesis, Edinburgh, 2004, http://winterstein.me.uk/academic/

Alison Pease, *A Computational Model of Lakatos-style Reasoning* PhD thesis, Edinburgh, 2007
http://hdl.handle.net/1842/2113

Samson Abramsky & Bob Coecke, Physics from Computer Science IJUC, vol. 3:3, 179–197, 2007 (QM in diagrams)
http://web.comlab.ox.ac.uk/oucl/work/samson.abramsky/YORKIJUC.pdf