

# Kantian Philosophy of Mathematics and Young Robots

**Could a baby 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/>

Last updated: [November 28, 2008](#) (liable to change).

---

These PDF slides will be in my 'talks' directory.

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

See also this longer presentation:

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

The conference paper is available here:

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

Kantian Philosophy of Mathematics and Young Robots

# Summary part 1

---

**We need a multi-pronged revolution in studies of intelligence,**

- replacing factional wars with healthy collaboration
- and with far more attention paid to the **problems** solved by evolution
- as opposed to focusing prematurely on supposed **solutions** to the problem,

e.g. hypothesised biological mechanisms.

Or even assuming that logic will suffice (McCarthy & Hayes 1969, Sloman 1971)

If we don't know what problems evolution actually solved, we (including biologists, neuroscientists, and psychologists) are likely to make serious mistakes about the mechanisms used to solve them.

This talk describes one facet of that revolution:

For other facets see my online papers and presentations.

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

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

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/>

# Summary part 2

---

## Some of the main points

- There are many designs for biological organisms – solving different problems.
- Most species (precocial species) are designed with genetically determined, **pre-configured**, competences tailored to a particular range of environments (all they can do is calibrate and adjust parameters).
- A small subset (altricial species) are designed to find out what sort of environment they are in and to go on extending their understanding of its properties and what can be done in it: they develop **meta-configured** competences.
- That can include acquiring new competences to acquire more competences, i.e. new meta-configured meta-competences, including substantively new ontologies.[\*]  
Humans in particular can go on learning new ways to learn, indefinitely.
- For some species learning **correlations** and conditional probabilities is not enough: they have to learn about **structures** and **structured processes** in the environment, and how they can be combined in novel ways. (E.g. building new tools, weapons, shelters, ...)

.....to be continued

---

[\*] In our IJCAI'05 paper Sloman&Chappell suggested that future robots will need different combinations of precocial competences and altricial competences. In our IJUC 2007 paper we started calling them preconfigured and meta-configured competences.

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0502> (2005)

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609> (2007)

# Summary part 3

---

- **The mechanisms required for such creativity are closely related to requirements for making not only empirical, but also mathematical, discoveries about the environment.**
- **As a result, some kinds of knowledge that start empirical in a child can be transformed, so as to be more like mathematical knowledge, i.e. non-empirical.**

Example: Learning that two ways of counting a set give the same result.

Note: The process whereby such discoveries are made is not bug-free. (Lakatos)

Examples of such bugs and how they can be fixed are given in

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

# Apologies and Acknowledgements

---

## I apologise

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.

## NOTE:

To remind me to skip some slides during the presentation they are 'greyed' with (X) at the top.

## Acknowledgements

The approach taken here is much influenced by reading G.Polya's *How to Solve It*.

The deepest influences were

- (a) experience of doing mathematics as a student and enjoying mathematical thinking.
- (b) Kant's philosophy of mathematics, expounded in *The Critique of Pure Reason* (1781) – his response to David Hume.

Although my interest in these problems goes back to my 1962 DPhil work on Kant's philosophy of mathematics, some of the new ideas reported here were stimulated by working on requirements for a robot that can manipulate 3-D objects in its environment, as part of the EU [CoSy](#) project:

“Cognitive Systems for Cognitive Assistants” (2004-2008).

<http://www.cognitivesystems.org>

<http://www.cs.bham.ac.uk/research/projects/cosy/>

This included many discussions with Jeremy Wyatt, Nick Hawes and other members of the School of Computer Science in Birmingham, and Jackie Chappell in the School of Biosciences, who works on animal cognition.

# Routes from DNA to behaviours: reflexes

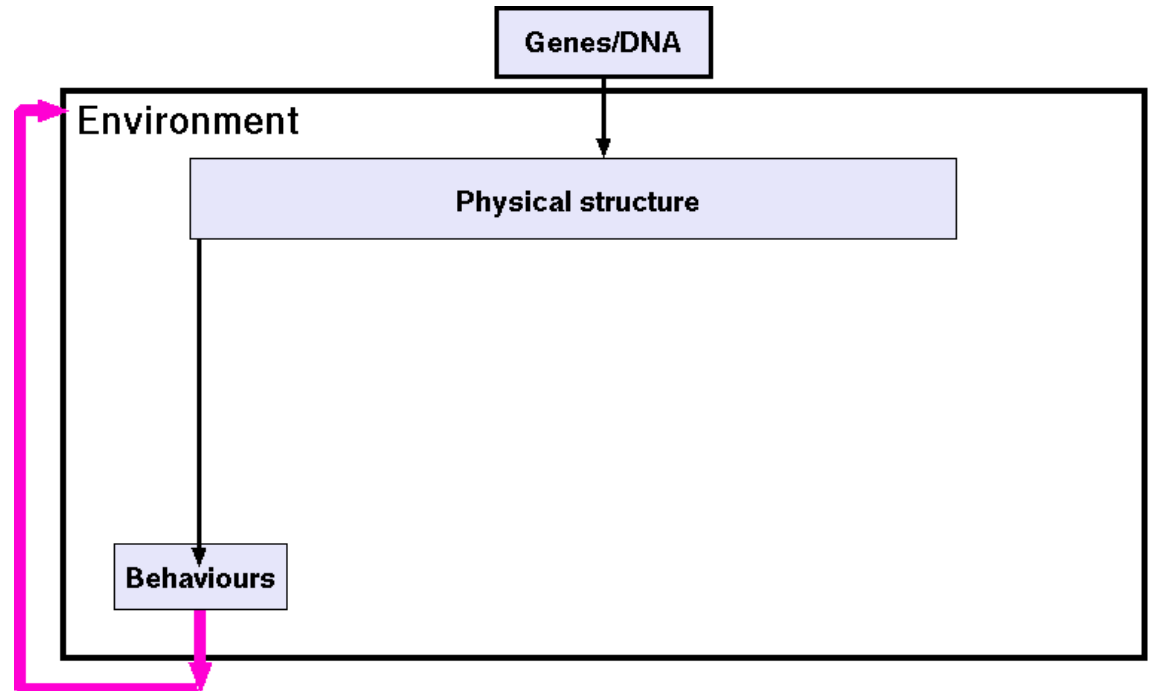
**Cognitive epigenesis:** Multiple routes from DNA to behaviour, some via the environment

## The simplest route from genome to behaviour:

Everything is hard-wired in a design encoded in the genome (subject to interpretation by epigenetic mechanisms).

The physical structures determine

- ongoing behaviours (e.g. breathing, or respiration, or pumping of a heart)
- specific reflex responses to particular stimuli  
e.g. the knee-jerk reflex.



Note: during development, the behaviours that produce effects on the environment may feed back into influencing further development, and learning.

**This work is based on collaboration with Jackie Chappell**

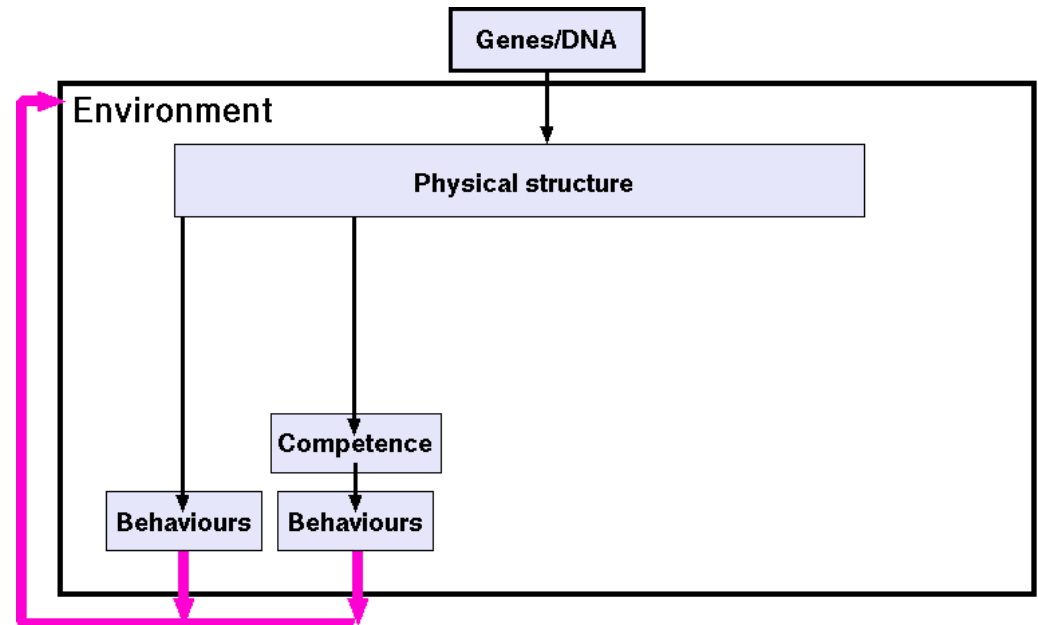
# Routes from DNA to behaviours: more flexible competences

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

A more complex route from genome to behaviour:

Everything is hard-wired in a design encoded in the genome (subject to interpretation by epigenetic mechanisms). But what is hard-wired is capable of modifying behaviour on the basis of what is sensed before and during the behaviour.

- The details of such behaviours are products of both the genome and the current state of the environment.



A “competence” is an ability to produce a family of behaviours capable of serving a particular goal or need in varied ways, e.g. picking something up, avoiding an obstacle, getting food from a tree by jumping, catching prey, avoiding predators, migration.

Some competences are “pre-configured” in the genome.

There is not necessarily a sharp division between reflexes and competences: the latter are more flexible and goal directed, but the degree of sophistication can vary a lot.

# Routes from DNA to behaviours: The role of meta-competences

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

A more complex route from genome to competences:

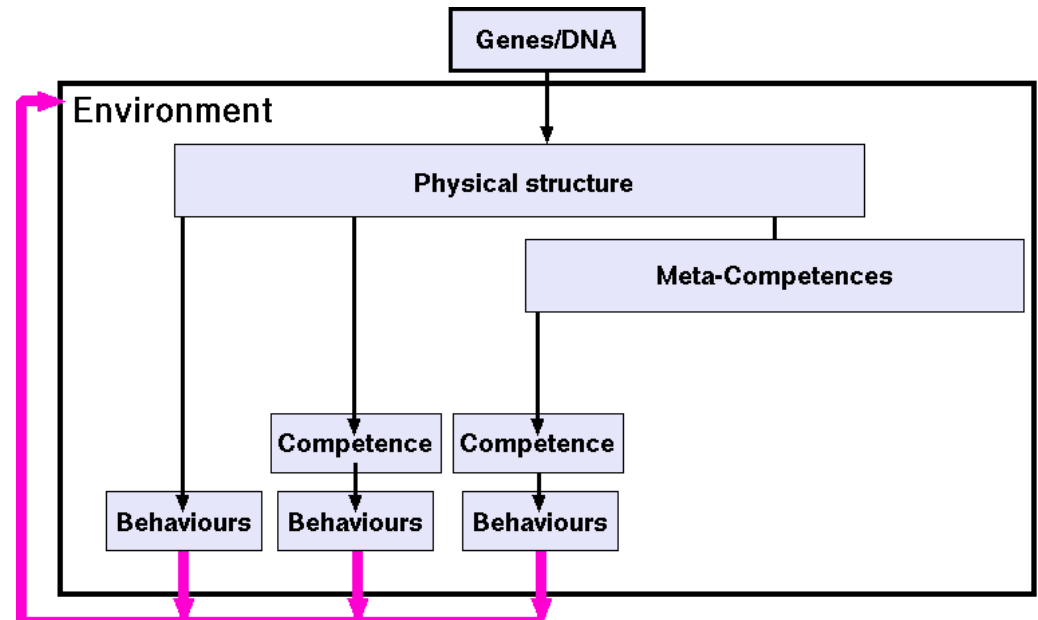
Instead of competences being hard-wired in a design encoded in the genome (subject to interpretation by epigenetic mechanisms), they may be developed to suit features of the environment, as a result of play, and exploration, leading to learning.

There may be hard-wired genetically preconfigured “meta-competences” that use information gained by experimenting in the environment to generate new

competences tailored to the features of the environment – e.g. becoming expert at climbing particular kinds of tree, or catching particular kinds of fish, while conspecifics in another location develop different competences produced by the same mechanisms.

The details of such behaviours are products of both the genome and the current state of the environment.

A “meta-competence” is an ability to produce a family of competences capable of serving varied goals in the environment of the animal. (Compare Alan Bundy’s talk: repair plans)





# Some virtual machines can extend themselves indefinitely

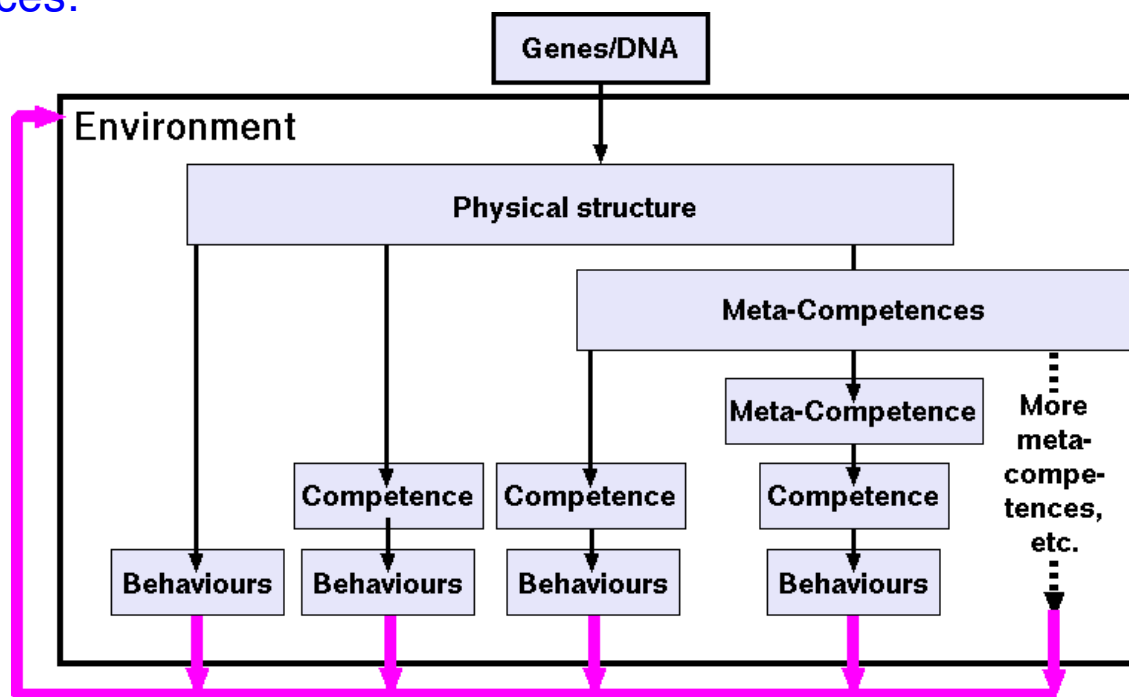
**Cognitive epigenesis:** Multiple routes from DNA to behaviour, some via the environment

## Meta-configured meta-competences:

Humans not only learn new kinds of things, they can learn to learn even more varied kinds of things. E.g. after completing a degree in physics you are enabled to learn things that a non-physicist could not learn, e.g. learning how to do more sophisticated experiments.

## Meta-configured meta-competences:

(towards the right of the diagram) are produced through **interaction** of pre-configured or previously produced meta-configured competences with the environment, including possibly the social environment



<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609>

Natural and artificial meta-configured altricial information-processing systems  
Chappell&Sloman, 2007, IJUC

Diagrams developed with Jackie Chappell and Chris Miall.

# Self-monitoring and virtual machines

---

Systems dealing with complex changing circumstances and needs may need to monitor themselves, and use the results of such monitoring in taking high level control decisions.

E.g. which high priority task to select for action.

Using a high level virtual machine as the control interface may make a very complex system much more controllable: only relatively few high level factors are involved in running the system, compared with monitoring and driving every little sub-process, even at the transistor level.

The history of computer science and software engineering since around 1950 shows how human engineers introduced more and more abstract and powerful virtual machines to help them design, implement, test debug, and run very complex systems.

When this happens the human designers of high level systems need to know less and less about the details of what happens when their programs run.

Making sure that high level designs produce appropriate low level processes is a separate task, e.g. for people writing compilers, device drivers, etc. Perhaps evolution produced a similar “division of labour”?

Similarly, biological virtual machines monitoring themselves would be aware of only a tiny subset of what is really going on and would have over-simplified information.

**THAT CAN LEAD TO DISASTERS, BUT MOSTLY DOES NOT.**

# Kant on Synthetic A priori Knowledge (e.g. mathematical knowledge).

---

Immanuel Kant wrote, over 200 years ago

**“There can be no doubt that all our knowledge begins with experience.**

**For how should our faculty of knowledge be awakened into action did not objects affecting our senses partly of themselves produce representations, partly arouse the activity of our understanding to compare these representations, and, by combining or separating them, work up the raw material of the sensible impressions into that knowledge of objects which is entitled experience?**

**In the order of time, therefore, we have no knowledge antecedent to experience, and with experience all our knowledge begins.”**

*Critique of Pure Reason*

# What about the genome?

---

Kant wrote before Darwin's theory of evolution, though he was aware of and wrote about the question whether humans had evolved from non-human animals.

However he seems not to have allowed for the possibility that many scientists now take seriously, that evolution produced genetically-determined cognitive competences and knowledge that are involved in learning and development from the earliest stages.

Think of precocial species like deer, horses, chickens, whose young are pretty competent after hatching – more competent than any existing robots. Some robots will need to be precocial also.

Compare

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

# Yet Kant implicitly supports such a theory in his question

---

For how should **our faculty of knowledge** be awakened into action did not objects affecting our senses ... work up the raw material of the sensible impressions into that knowledge of objects which is entitled experience?

Whatever this “faculty of knowledge” is, it could not be **awakened** without **already existing**.

**What is this faculty?**

**Perhaps it is a kind of information-processing architecture that drives exploration of the environment while it grows itself?**

# Genome (+) at work on a rug - after about 6 months



Video available online:

<http://www.cs.bham.ac.uk/research/projects/cosy/conferences/mofm-paris-07/slovan/vid/>

The infant seems to produce a new combination of previous competences: looking round, identifying an edge, setting up a goal, rolling over, stretching arm, opening fingers, moving down, closing fingers, pulling.

Human learning seems to progress in layers of competence, where new layers include components in old layers (including previously learnt actions and hypotheses) as objects to play with.

One feature of that is specially important for our topic:

Transforming empirical discoveries into non-empirical discoveries.

For the infant there may at first be just empirical discoveries about which movements produce which effects.

But members of the audience have got long past that:

you know a lot about what actions can and cannot succeed — and why.

# Motivation and Learning are not Reward-based but Architecture-based

---

**ASK ME LATER WHAT THAT MEANS**

# Genome plus much acquired competence at work with a broom (about 15 months)

---



Some snapshots from the video (available online):

<http://www.cs.bham.ac.uk/research/projects/cosy/conferences/mofm-paris-07/slovan/vid/>

---

## At least two different kinds of thing going on in parallel:

- Control of several interacting dynamical systems for maintaining stability and producing motion of the required sort.
- Various kinds of perceptual and cognitive processes involved in controlling the motion **by understanding how the broom (broom handle and brush) interacts with different things in the environment.**

The child feels that the broom handle is caught, looks round, then moves it away from between the railings. When the brush hits the skirting board he retracts and then redirects it.

Later, he sees that there's an opening off the corridor on the right and (unconsciously?) works out that steering into it will require the broom handle to be moved to the left shortly before the opening is reached.

There are also two fleeting bits of social interaction, indicating even more concurrent processing.



# Information initially gained empirically can be transformed, and recognised as non-empirical.

A key feature of the information-processing architecture produced by evolution, in humans and possibly a few other species is this:

Acquired knowledge that starts off empirical can be transformed, so that it has a new status.

E.g. is knowledge about a movement that will stop the railings impeding the handle empirical?

**This is because the environment has a great deal of discoverable structure.**

Spatial structure (geometrical, topological, continuous, discrete, orderings)

Process structure, combined, adjacent, synchronised processes, mappings, ....

Kinds of matter with distinct properties (rigid, flexible, elastic, inelastic, fluid, sticky, plastic, etc.).

The transformation from empirical to non-empirical (mathematical) is not modelled by any kind of statistical or probabilistic learning mechanism.

This needs an argument. E.g. changing probabilities of correlations does not allow  
creative prediction in new configurations,  
or new explanations  
or new plans  
or new designs

# This is not a change in probability – but something deeper:

---

## **It is a qualitative, not a quantitative change.**

Adding a new explanatory theory is not the same as modifying a probability

(Compare Alan Bundy's talk on ontology/theory changes)

Likewise developing a new form of representation is not a change in probabilities.

Explaining how this happens, and modelling the processes (e.g. in human-like robots), will require the development of new architectures, and perhaps new forms of representation.

Probably not using predicate calculus – maybe some formalism closely related to visual competences and human visual reasoning.

Compare the original presentations of Euclidean geometry.

Several examples, of different kinds, will be given below.

# An underrated, but powerful, learning process: Creatively playing, exploring, trying things out

---

This is very different from corpus-based, or data-mining-based learning, where the only criterion of success is an external judgement.

The learner who observes, plays, explores, and tests ideas in a complex but highly structured, and in many ways deterministic world, is in control.

Such a learner can tell:

- **when predictions are violated**
- **when goals are not achieved**
- **when hypotheses pass novel tests**
- **when unexpected things are observed**
- **when a new possibility turns up**
- **when an option has not yet been tried.**

Not always, and not infallibly, but often enough to drive learning by play and exploration.

# It can be, but need not be, a social process.

NOTE:

**You can't learn something by imitation or instruction  
if you don't have the cognitive resources in the first place.**

This is the obverse of McCarthy's dictum:

What you can discover for yourself you could learn by being told.

There are exceptions to McCarthy's dictum.

It's hard to learn by being told what the taste of pineapple is.

# Forms of representation

---

What form of representation does an infant or toddler use?

Or a chimpanzee or orangutan?

Or a nest building bird (e.g. Betty, the hook-maker)?

**Nobody knows.**

Could it be predicate calculus?

Could it be a language like human natural languages, e.g. English?

Could be more like a programming language, e.g. Lisp, Prolog, Java?

Are neural nets capable of expressing what's needed, and if so how?

**Nobody knows – none of those is plausible for a toddler.**

Some of us have been suggesting that some non-human animals and pre-verbal children have ways of manipulating structures specially tailored for spatial reasoning.

E.g. Sloman 1971, 1978 (Chapter 7)

See this presentation on the evolution of language

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

But the problems are deep and difficult and there are many problems to be solved.

NB: Beware of assuming the representing structures are **isomorphic** with what they represent.

(See the impossible object pictures below.)

# Empirical vs Non-empirical knowledge

---

## Example of the distinction, from Kant.

- **Empirical:**

All solid bodies are heavy

Many solid things require effort to raise them off a surface they are resting on. (But exceptions exist.)

- **Non-empirical**

All solid bodies occupy space

Kant: “All bodies are extended”

## Question:

Is all non-empirical knowledge essentially just a matter of definition?

Isn't it part of the concept of a solid object that it occupies space?

Arguably, it is impossible for a child to **discover** (by observation or experiment) that solid objects occupy space, for where else could it look for solid objects except in space?

Does that make all non-empirical knowledge trivial?

(Or “analytic”, in Kant’s terminology)

Or is some of it non-trivial, substantive, an addition to knowledge?

(Or “synthetic” in Kant’s terminology)

According to Kant, some non-trivial, substantive knowledge, may initially be acquired as a result of experience, but then come to be understood as non-empirical, and as **necessarily** true.

# Examples for babies and toddlers

---

The following are empirical discoveries at first:

- A certain sort of visual experience (of the edge of a rug) indicates the possibility of a certain sort of action (grasping the edge).

The action requires certain intermediate stages,

e.g. reaching, opening fingers, moving hand down, closing fingers, lifting.

- A rigid rod (e.g. broom-handle) poked between a pair of vertical railings cannot be moved sideways in the plane of the railings, but can be moved in the perpendicular direction.

A result of the second motion can be that the previously impossible direction of motion becomes possible.

Are these empirical facts for you?

Beware: That question has different interpretations.

Care is needed in the formulation/interpretation.

Another version:

- Could you imagine developing a mathematical theory of rug grasping?
- Could you imagine developing a mathematical theory of how rigid, impenetrable, rods or poles can move in the presence of other rigid, impenetrable things?

# Somatic and exosomatic ontologies

---

There are many “biologically inspired” researchers who think brains learn correlations between sensory and motor signals.

For such a researcher, committed to “the sensorimotor approach” the statement

“A certain sort of visual experience (of the edge of a rug) indicates the possibility of a certain sort of action (grasping the edge).”

would be a summary of a collection of correlations between visual input signals (or retinal patterns) and motor output signals.

It would summarise statistical facts about events on and inside the skin (taking sensors as part of the skin).

Such generalisations use a **purely somatic** ontology.

But some biological systems, e.g. humans, are capable of using an **exosomatic** ontology, referring to objects, properties, relationships, events, and processes existing outside the perceiver’s body, and capable of existing independently of any individuals sensory or motor patterns.

Without using an exosomatic ontology you would not be able to think about

- what’s going on out of sight,
- what happened in the past,
- what might happen in the future,
- what you should do in the future, and
- what’s going inside objects whose innards are not accessible to you.



# An exosomatic infant ontology

---

Whatever the state at birth, humans are eventually able to use an exosomatic ontology.

That can give a new meaning to

“A certain sort of visual experience (of the edge of a rug) indicates the possibility of a certain sort of action (grasping the edge).”

i.e. something of the form

**“When you see a certain 3-D structure in the environment related to you in certain ways, you can infer that certain 3-D processes are possible, involving your hands and the perceived structure.”**

The exosomatic ontology provides many advantages of economy, and power, including being usable for **reasoning** about processes in the environment, including actions and affordances of others.

Vicarious affordances can be important: a point ignored by J.J.Gibson, I think.

Warneken’s experiments show that even pre-verbal children can detect them

<http://email.eva.mpg.de/~warneken/video.htm>

## See also

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/nature-nurture-cube.html>  
(Learning to see a set of moving lines as a rotating cube.)

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

# Ontologies and theories develop together

---

(Compare Alan Bundy's talk.)

Whereas discoveries of correlations involving sensory and motor signals are empirical, within a **theory** of “external” 3-D structures and processes, discoveries can have a mathematical status, like theorems of geometry and topology.

What is possible and what is impossible within a configuration of structures can be derived from the specifications of the structures and their relationships.

# A theorem for infant science

---

E.g.

## Theorem 1:

If you stand with your soles next to a pair of slippers with their soles on the floor you cannot get your feet into the slippers just by sliding them horizontally.

(This might first be learnt empirically.)

(This is a special case of a much more general theorem about things with openings.)

## Theorem 2:

You can't shut an open door by grasping its edge and pulling it towards you until it is shut.

(Don't try that!)

Of course you can start the process that way if you can generate enough momentum in the door for it to complete the shutting process on its own.

For more on substantive ontology development and how meanings can be extended see this attack on "symbol grounding" theory:

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

# More theorems for infants and toddlers

---

In the first few years of life, as infant then toddler, a normal human learns many hundreds, possibly thousands, of facts about things that are and are not possible in the environment.

Because they become part of our common sense, and are “second-nature” to us, we do not notice that we and others learn them, nor do we find them remarkable. Things like

- How to adjust/improve grip on a held object.
- Results of pushing, prodding, squeezing, rotating objects of various shapes and sizes, made of different kinds of stuff.
- Results of moving around in various ways in relation to obstacles, openings, occluders, etc.
- Results of moving markers (pens, pencils, crayons) on markable surfaces.

E.g. What movements of a pencil point are needed to make it draw a square: A four year old tried, and failed and was annoyed – he could only draw curves, including more or less closed curves. He knew he had not drawn a square, but did not know how to do it. **Several months later he found it trivial.**

- How to get liquid out of a half-empty bottle, first when sucking and later when pouring.  
At first it's an empirical discovery that tilting your head back is needed for sucking. Later ???
- How to correct a mis-aimed teat on bottle (i.e. making a corrective movement depending on which part of the face it makes contact with).
- What changes in the mouth when one end of a rigid object is in the mouth and the far end is moved in various ways – e.g. a toothbrush.
- Various actions required at various stages of putting on and taking off various sorts of garments.  
**You can't put a sweater on just by holding it against you and pressing hard.** (Contrast a hat or cap.)

# 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.

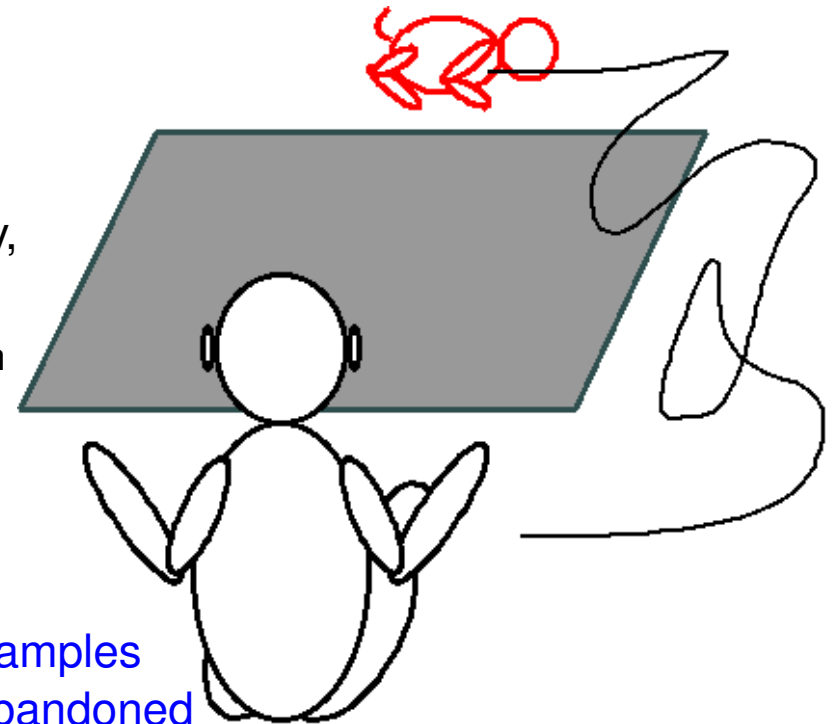
However, pulling a string that is not straight can transform it into a straight string, after which the object can be moved towards the puller.

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

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.

But some of what is learnt is about non-contingent truths that, once recognised as such, require no more empirical evidence.

**A fixed-length (inelastic) string gets straighter as its ends are moved apart.**



# From action affordances to epistemic affordances

In addition to discoveries about possible actions and processes

- what actions the individual can and cannot perform in what circumstances
- more generally what processes can and cannot occur in various circumstances,
- what the consequences of those actions and processes must be

a child also needs to learn things about available information

- What information about the world is available to the perceiver in various situations
- how the available information can be changed by various actions and other processes.

In other words, a child has to learn about both

- action affordances, and
- epistemic affordances, and
- how they are related.

# 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.

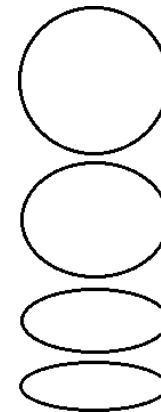
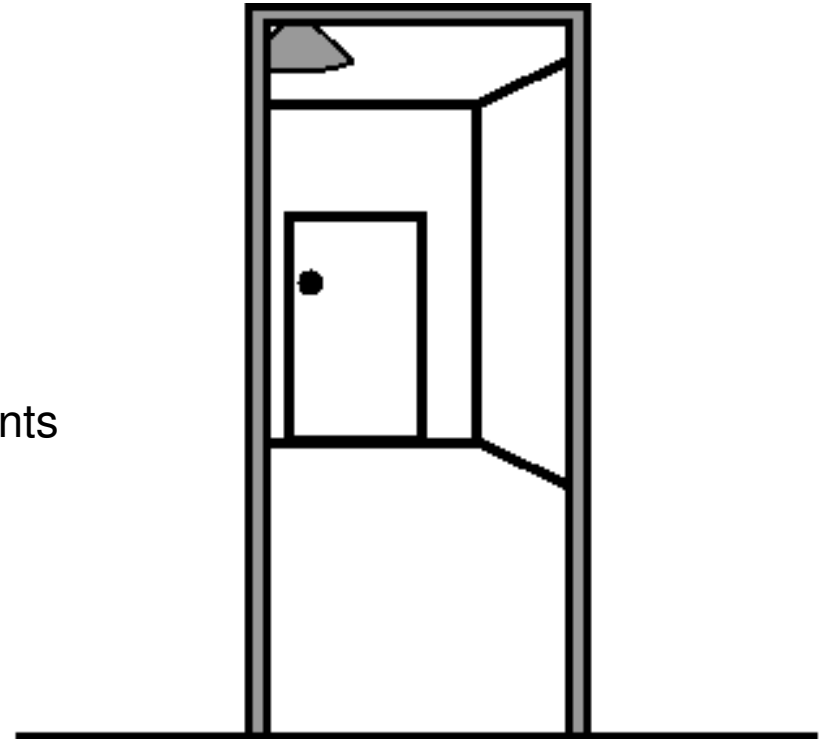
Why do those procedures work? How do they differ?

- 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

Why? How? Is it just a brute empirical fact?



# Towards a mathematics of door action affordances and other domestic action affordances?

---

- 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 avoid bumping into the **left** doorpost while going through a doorway by aiming further to the **right** – and what problem does that raise?
- How could you use the lid of one coffee tin to open the lid of another which you cannot prise out using your fingers?
- Why it is easier to carry a tray full of cups and saucers using a hand at each side than using only one hand on one side?

and so on and so on ....

Try to work out what theorems underlie such common sense knowledge.

How might the theoremhood become apparent to a child?



# Examples with discrete event sequences

---

Compare these discoveries a child might make by playing with things in the environment.

1. On one of the objects the child has access to there is a button, and whenever the button is pressed down the end of the object glows brightly, until the button is released – which makes the glow stop.
2. Another object is a circular disc, with a picture on each face, only one of which is visible when the disc is lying on the floor, but every time it is turned over the other picture becomes visible.

**Are both of those contingent truths?** Depends on how the processes are specified.

3. Another variant:

If the child has learnt to count, then she can count in synchrony with turning the disc over several times, noticing both the last number named and the state of the disc.

The following empirical discoveries could be made:

- If the last number is “four”, the picture on top is the same as the initial picture.
- If the last number is “six”, the picture on top is the same as the initial picture.
- If the last number is “three”, the picture on top is not the same as the initial picture.
- If the last number is “seven”, the picture on top is not the same as the initial picture.

What else could be discovered, from such experiments?

# Example: Counting invariances

---

A child may first learn to count.

Then learn empirically that counting objects from left to right gives the same result as counting from right to left.

Then learn that the two results **must** be the same: it is not empirical.

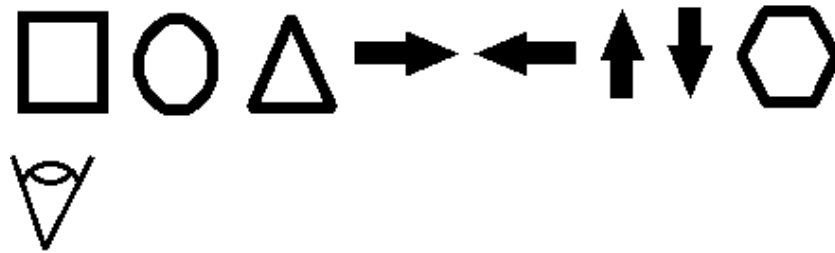
Then he may generalise to the notion of an invariant property of a set.

Then he may learn more about empirically discoverable non-empirical facts about numbers.

The process of identifying non-empirical truths is not bug free: faulty conjectures and proofs may have to be fixed.

## 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

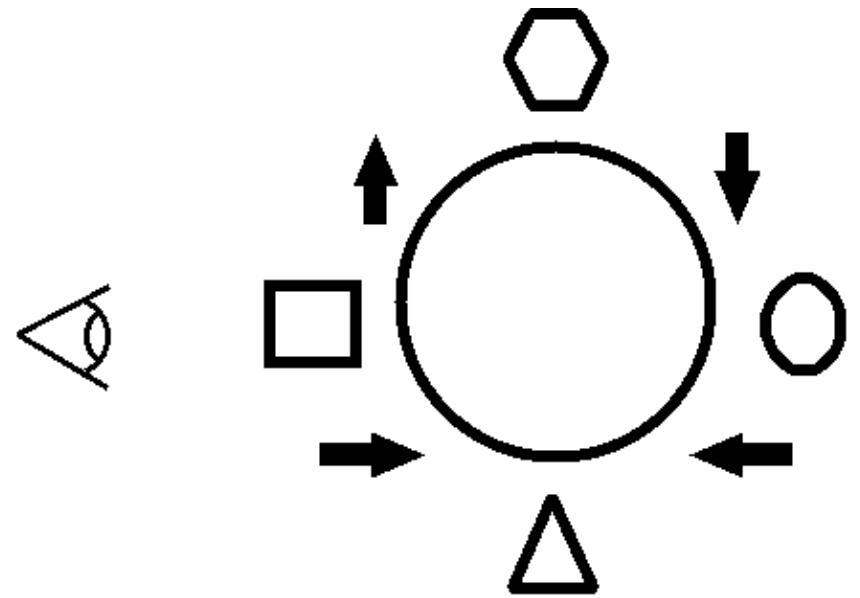
If you move round an **unchanging** configuration of objects arranged against the wall of a building, you will see them in a particular order, and as you continue moving round the building you will see the same sequence over and over again.

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

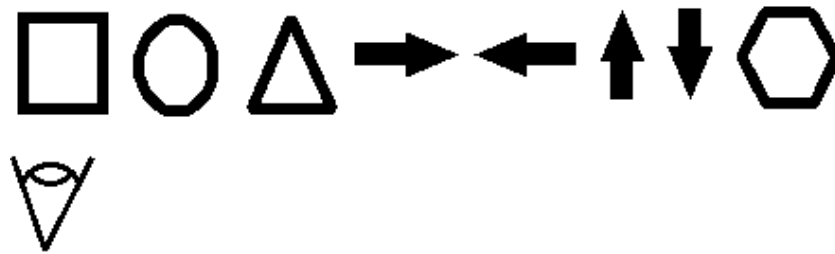
What information-processing mechanisms allow a child to make this discovery, or to use it?

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:
<b>A B C</b>	<b>A B</b>	<b>A</b>	
	<b>C</b>	<b>C B</b>	<b>C B A</b>

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”

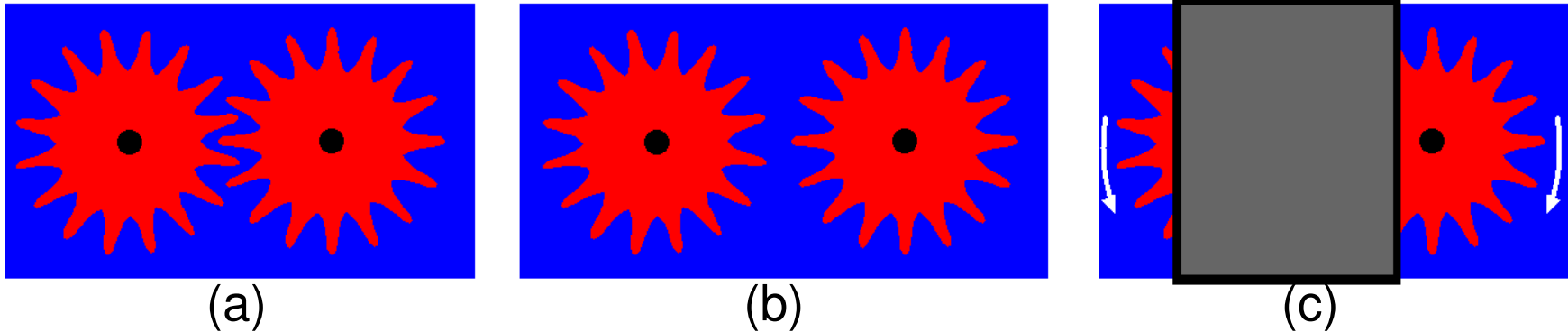
You can take the cards one by one from left to right and pile them up.

- 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.**

# 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. **Kinds of “stuff”**.



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**.

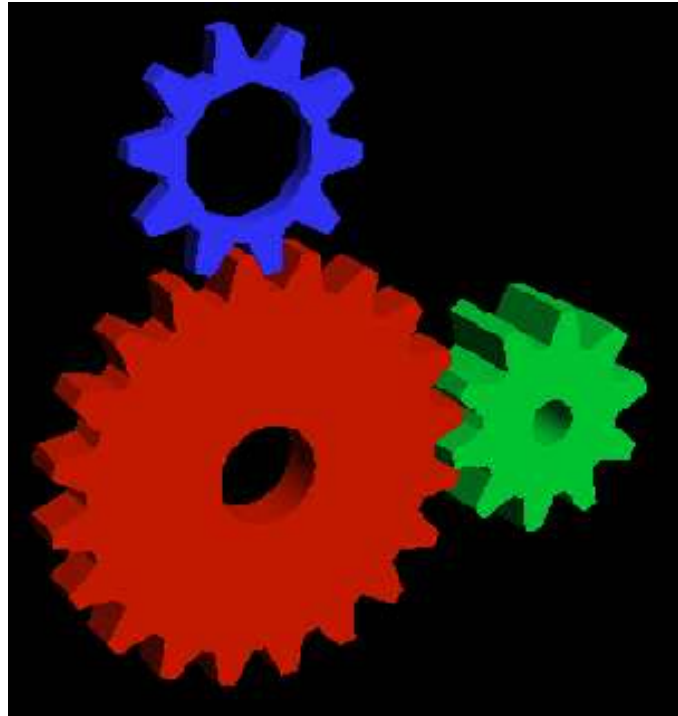
Kantian (structure-based, deterministic) causation is sometimes available, and is often more useful, but it requires a richer ontology and more general reasoning abilities.



# 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.

# Building A Collection Infant/Toddler Theorems

I would like to get help in building a much larger collection of examples of things that could be learnt empirically by human infants and toddlers, and perhaps some other animals, but which have an explanation that shows they are not merely contingent facts but can be understood as necessary truths within a more general theory of the nature of the environment, e.g. as containing 3 dimensional structures and processes, with a locally euclidean geometry, and containing materials with various properties, including rigidity, impenetrability, various kinds of elasticity, etc.

**Please send examples to me at [A.Sloman@cs.bham.ac.uk](mailto:A.Sloman@cs.bham.ac.uk)**

A lot more needs to be done to specify more precisely what should and should not be included, how it should be organised, what experimental and other procedures could be used to justify their inclusion, what new experiments could be generated on the basis of the examples, e.g.

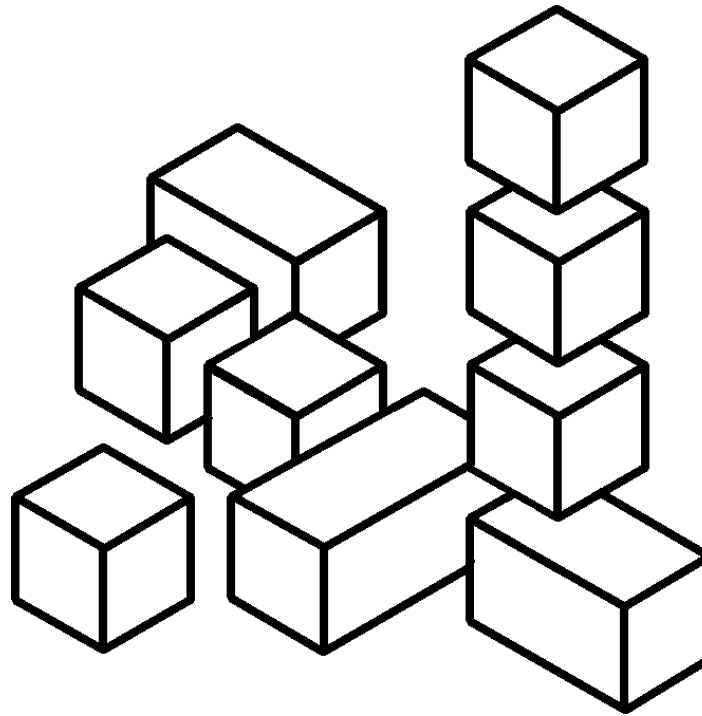
- (a) to find out which children (and other animals?) learn which examples, and under what conditions they make the transition from purely empirical understanding to grasping what **must** be the case;
- (b) to find out what brain mechanisms might be involved in the transition from empirical to “mathematical” understanding of particular cases;

and what sorts of robotic experiments could be constructed to demonstrate how such learning and the transitions between different sorts of understanding could work.

(Compare the work of Piaget, Karmiloff-Smith and others.)

# What can be built from a pile of cubes 1?

---

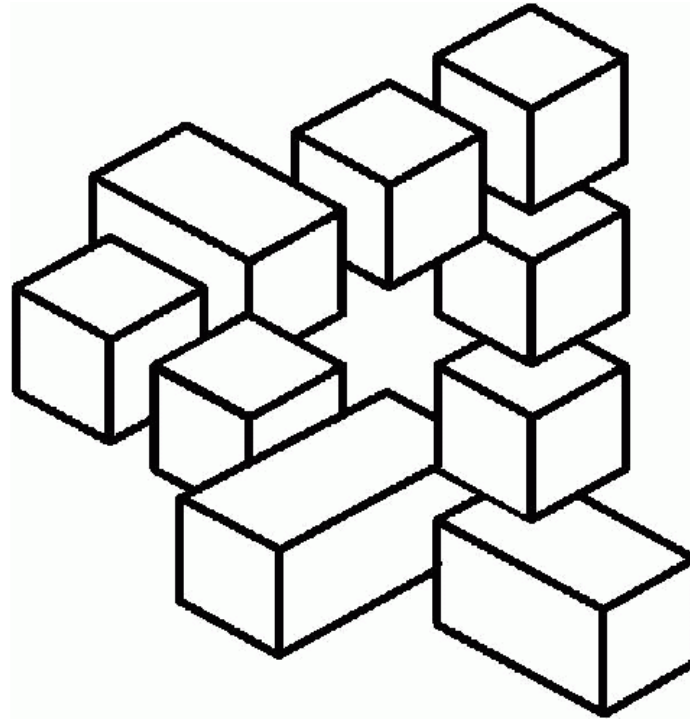
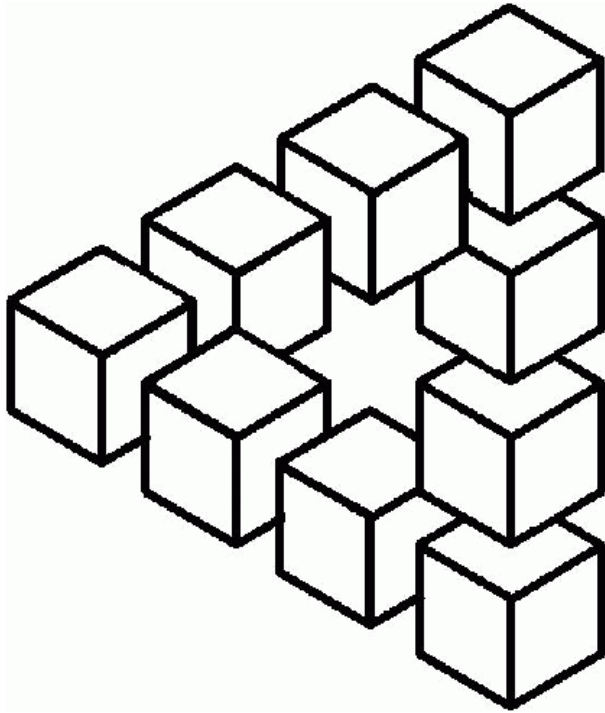


If you were given a box full of cubes and rectangular bricks could you assemble a configuration roughly like this?

(It's not a trick question. Compare the next slide.)

## What can be built from a pile of cubes 2?

---



Young children don't see anything wrong with these pictures?  
(Compare Escher's pictures.)

Why not?

What has not developed yet?

How does it develop later?

Did you see the similarity between the picture on the right and the one on the previous page? Only one block was moved.

# Should a robot be able to see impossible objects?

Impossible triangle by Reutersvard – Swedish artist 1934

At first this might simply look like a configuration of cubes in the form of a triangle.

However you should be able to convince yourself that the 3-D configuration cannot exist.

If you remove any two adjacent cubes at one of the corners, the remaining seven cubes form a consistent configuration.

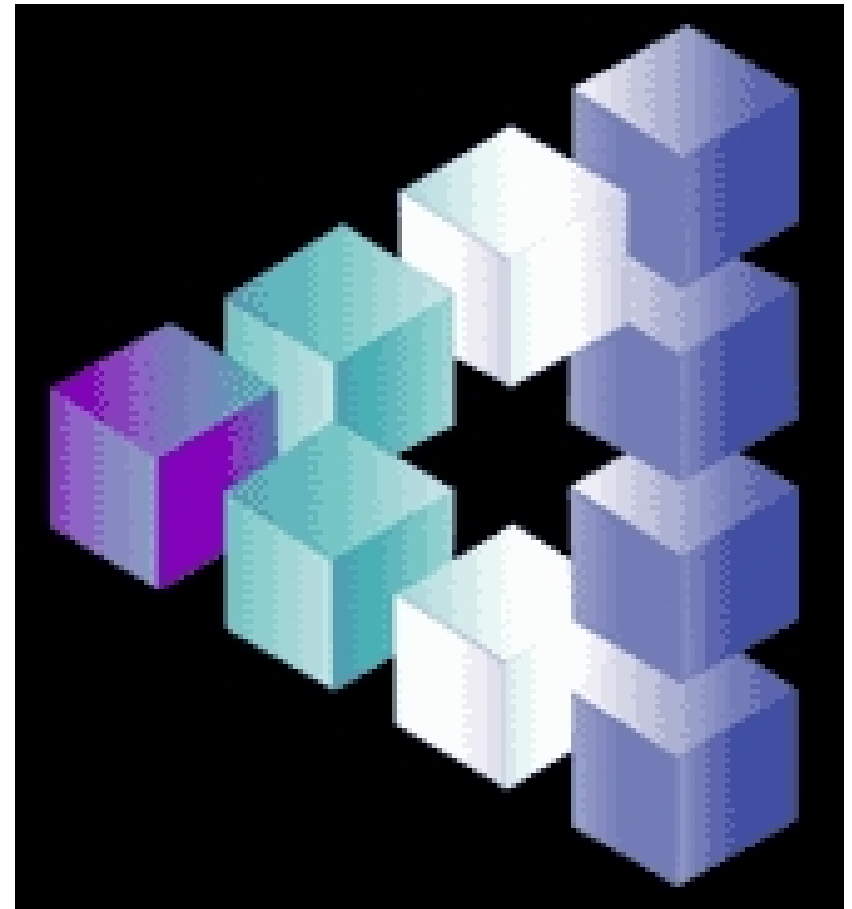
But if you put back all the cubes, the relative distances from you become impossible, like

$$A > B > C > D \dots > A$$

Should a robot looking at something like this notice the impossibility?

How can child, or you, see a lot of 3-D fragments that do not add up to a possible 3-D whole, without noticing the impossibility?

What does that tell us about how a robot should see a 3-D environment?



# 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
- Blocks moving from one location to another.
- Counting processes
- Coin-turning processes
- Processes of re-ordering items

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

What changes, and how? Don't assume it's an isomorphic model of the environment.

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.

If you have learnt a dance, do you have a constantly running simulation of your dancing in your brain?

# 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,

or perhaps a program,

or perhaps both:

    a grammar for recognition, and

    a program for production?

or something completely different, not yet thought of by any programmer or robot designer.

# The need for a self-monitoring architecture

---

The previous slides 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.



# My claims

---

- In the first five years of life, every child encounters hundreds, or perhaps thousands, of examples of knowledge each of which starts empirical but then acquires a new status, though the transitions mostly go unnoticed both by the child and by others (e.g. developmental psychologists).
- These capabilities depend both on the nature of human perception of spatial structures and processes, and on the ability to notice features of such perception when it occurs.
- These capabilities are also at the root of human mathematical competences
- Similar transitions will be needed for human-like robots.
- Very few (if any) developmental psychologists or AI researchers have noticed this transformation of empirical knowledge to apriori knowledge.
  - I expect any who have noticed it have wrongly characterised it as a change in subjective or objective probability produced by accumulating evidence.  
(Compare philosophers who think mathematical knowledge is empirical, e.g. J.S. Mill.)
- There are no AI models that I know of that are capable of explaining how this happens.
- It also seems to be beyond what neuroscientists are currently capable of studying.

# 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, non-mathematical subject matter.

Previous slides point towards an explanation of how mathematics can be applicable to the world in something like the way that Feynman suggested.

---

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>

# Conclusion

---

It's an old idea that a child is a young scientist.

We can now add more content to that idea.

Not just human children – maybe some other altricial species.

But humans have some special features (indefinitely many levels of meta-competences).

This may be connected with both special architectures and special forms of representation.

It may not be an all-or-nothing matter: perhaps some other intelligent animals, e.g. corvids, orangutans, and hunting mammals, have portions of the architecture and special cases of the forms of representation that give humans their extended capabilities.

There are forms of representation used by pre-verbal children and non-verbal animals that seem to be very powerful in discovering and creatively using facts about how the environment works.

This has deep implications for the nature of language learning and language evolution.

Maybe AI will look somewhat different when we know how to model all this.

In 100 years time – or in the next decade ???