

Evolution of two ways of understanding causation: Humean and Kantian

Aaron Sloman

School of Computer Science, University of Birmingham

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

Jackie Chappell

School of Biosciences, University of Birmingham

<http://www.biosciences.bham.ac.uk/staff/staff.htm?ID=90>

Based on work done with the Birmingham CoSy team on requirements for human-like robots, and discussions with Dean Petters about babies.

Location of these slides

These slides are available here:

<http://www.cs.bham.ac.uk/research/cogaff/talks/wonac/#sloman>

Jackie Chappell's slides are available in two formats at

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

Slides extracted from the above, and then expanded, on

[Causal competences of many kinds](#)

are at

<http://www.cs.bham.ac.uk/research/cogaff/talks/wonac/#causal>

A closely related talk presents a (possibly) new theory of vision

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0505>

See also

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

(Includes several joint papers by us on these topics)

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

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

(The Birmingham Cognition and Affect Project)

Note on these slides

This is one of two linked presentations on causation and the altricial-precocial contrast. The second was Jackie Chappell's presentation (see links on previous slide):

[Understanding causation: the practicalities](#)

A third set of slides, on why we should not try to find behavioural tests for whether an animal, or robot, understands causation, was extracted from this set and is now available separately here:

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

[Causal Competences](#)

NOTE:

I (AS) try to make my slides readable by anyone interested – without having to hear me present them.

This means

- (a) that they contain too much clutter for presentations
- (b) that there are usually far more slides than can be included in a single presentation.

These slides are no exception.

Only a small subset of the slides could be presented during the workshop.

NOTE: these slides are produced using LaTeX and developed and presented on linux.

Summary of both talks

- There are two concepts of causation:
 1. **Humean causation**: evidence-based, probabilistic, statistical
 2. **Kantian causation**: structure-based, deterministic
- Most current theorising about causation in philosophy, psychology and AI is **Humean** in a modern form: e.g. using Bayesian causal nets.
- This ignores deep features in **Kantian** causation connected with reasoning about spatial and temporal structures and the role of properties of different kinds of stuff.
- Kantian understanding of causation, when available, also allows more creativity, and recombination of different kinds of knowledge to deal with new problems
 - because of the way different structures and processes are embedded in the same spatial region.
- The growth of understanding of Kantian causation is linked to forms of learning and development found only in animals usually classified as altricial, for reasons that are only beginning to become clear.
- We need to revise and update some biological and computational ways of thinking about animals and machines and their evolution and development.
- Close observation of play and exploration in children and animals, including their failures as well as their successes, provide clues as to what is going on: including development of ontologies and forms of representation, requiring **abduction**.
- Systematic biological and psychological research, along with design and implementation of working models can add more clues and help to test the theories.

Abstract

The current emphasis on causation as correlational/statistical, i.e. Humean, as in Bayesian nets, ignores a deeper notion of causation as structure-based and deterministic, i.e. Kantian. The history of science frequently involved moves from Humean (merely observed) to Kantian (intelligible) causation, and that also seems to happen in young humans and a subset of other animals.

A Kantian grasp of causation in our environment typically requires understanding of spatial structures and relationships and being able to reason about what happens when they change (often with multiple relationships changing concurrently). This uses geometrical and topological reasoning, while taking account of properties of different kinds of matter (e.g. rigid, flexible, brittle, elastic, tearable, etc.). A result is the ability to cope creatively with some novel problems. This is closely related to the development of some human mathematical competences. (The ability to do logical and statistical reasoning requires other forms of representation.)

The differences between ontologies and forms of representations required for Humean and Kantian understanding of causation will be discussed in relation to cognitive development and requirements for manipulation of 3-D structures (e.g. adding twigs to partially built nests, getting at flesh in animal prey, or playing with meccano and other construction kits). There are deep implications for various disciplines, including neuroscience, linguistics, education, philosophy.

Evolutionary considerations explain why varieties of causal competence are related to the altricial-precocial distinction: Kantian causal understanding seems to require characteristics of altricial species – which acquire successive layers of orthogonal recombinable competences, using evolved mechanisms for learning by using playful, creative, exploration of the environment, together with what has previously been learnt, whereas precocial species have only pre-compiled causal competences. This extends some old ideas of Oliver Selfridge.

Some of the theoretical ideas, empirical data, and challenges for biological researchers, will be discussed in more detail in Jackie Chappell's talk.

Prologue: Hume and Kant

Some background to this presentation: an abbreviated approximate history

David Hume was one of the great empiricist philosophers who thought that every concept has to be derived from experience of instances — a mistaken but often re-invented theory recently revived in connection with the label ‘symbol grounding’ criticised in <http://www.cs.bham.ac.uk/research/cogaff/talks/#grounding>.

He also argued that we do not experience any kind of causal connection apart from co-occurrence or sequential occurrence. So he concluded that ‘causes’ means approximately ‘is regularly correlated with’, since nothing stronger than co-occurrence is ever experienced, though learnt correlations often lead to strong ‘feelings’ of expectation. But we do not experience one thing ‘making’ another happen.

See Hume’s **A Treatise of Human Nature** and his **Enquiry Concerning Human Understanding**.

This is now a very popular interpretation of causation, usually elaborated using ideas of Bayesian probability.

Immanuel Kant responded that (a) concept empiricism is false, because in order to have any experiences you need concepts, so they cannot all come from experience, (b) to make sense of our experiences as referring to an external world with properties that we do not directly experience we need a notion of causation as involving **necessity** not merely correlation, and (c) he thought we had ways of understanding necessity that Hume had not recognised e.g. the ability to discover synthetic necessary truths in mathematics (arithmetic and geometry).

The majority of thinkers seem to side with Hume rather than Kant, but I think they are missing some deep facts about human and animal knowledge about the environment, and requirements for intelligent robots. The use of Humean causation may be widespread in many animals but human, primate, and avian intelligence requires something deeper, closer to Kant, involving understanding of structures and structural changes.

This presentation introduces some of the ideas and some of the evidence, though there is still much work to be done showing how the required mechanisms can actually work in brains or in other machines.

The ideas are developed further in a theory of perception presented in various papers and discussion notes, and this slide presentation: <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0505>

Conjecture, and Videos of human infants

- Broom video
- Yogurt video
- Hook video (child, not crow)

The main difference:

- Humean causation is concerned with statistical relations between 'atomic' facts
- Kantian causation is concerned with deterministic relations between complex structures and complex processes involving those structures (e.g. old clocks).

In particular when multi-strand relationships change multi-strand processes occur.

Conjecture:

The evolution of mechanisms for perception of 3-D processes, such as grasping, twisting, levering, sliding, breaking, fighting (and many more) led to the ability to understand Kantian causation and the ability to do Kantian causal reasoning.

Two quotations

Annette Karmiloff-Smith:

“Decades of developmental research were wasted, in my view, because the focus was entirely on lowering the age at which children could perform a task successfully, without concern for how they processed the information.”

Karmiloff-Smith, A. (1994).

Precis of *Beyond modularity: A developmental perspective on cognitive science.*

Behavioral and Brain Sciences 17 (4): 693-745.

(Preprint: <http://www.bbsonline.org/documents/a/00/00/05/33/index.html>)

This prompts the question:

What about the research on finding out under what conditions various animals do interesting things?

Ulric Neisser:

“we may have been lavishing too much effort on hypothetical models of the mind and not enough on analyzing the environment that the mind has been shaped to meet.”

Neisser, U. (1976)

Cognition and Reality San Francisco: W. H. Freeman.

Our position:

Understanding the phenomena requires both a deep understanding of important properties of the environment the animal or child interacts with and the information processing mechanisms that make it possible to engage fruitfully with **that** environment.

John McCarthy on The Well Designed Child

Quotes from his unpublished online paper: 'The well designed child',

<http://www-formal.stanford.edu/jmc/child1.html>

“Evolution solved a different problem than that of starting a baby with no a priori assumptions.”

...

“Animal behavior, including human intelligence, evolved to survive and succeed in this complex, partially observable and very slightly controllable world.

The main features of this world have existed for several billion years and should not have to be learned anew by each person or animal.”

Biological facts support McCarthy:

Most animals start life with most of the competences they need – e.g. deer that run with the herd soon after birth. For them, there's no **blooming, buzzing confusion** (William James)

So why not humans and other primates, hunting mammals, nest building birds, ...?

Perhaps we have not been asking the right questions about learning.

We need to understand the nature/nurture tradeoffs, much better than we currently do, and that includes understanding what resources, opportunities and selection pressures existed during the evolution of our precursors, and how evolution responded to them.

See the papers by Chappell and Sloman, starting with:

The Altricial-Precocial Spectrum for Robots, in *Proceedings IJCAI'05*

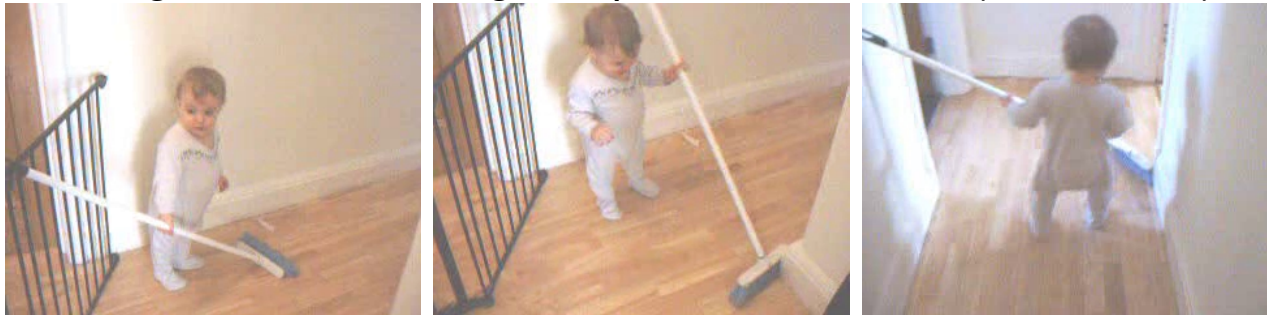
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0502>

Causal competences: Yogurt, broom and train

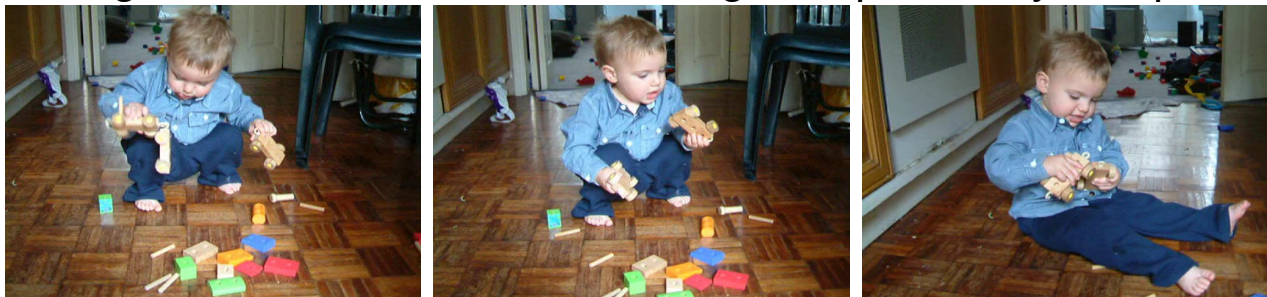
Playing with yogurt and with ideas (11 months)



Pushing a broom – using competence and luck (15 months)



Failing to understand hooks and rings despite many competences (18 months)



The videos will be shown, on request, during discussions.

The Child as Scientist: 1



Yogurt can be food for both mind and body in an 11 month baby.

Video available at

http://www.jonathans.me.uk/josh/movies/josh23_0040.mpg

Hypothesis

Alongside the innate physical sucking reflex for obtaining milk to be digested, decomposed and used all over the body for growth, repair, and energy, there is a genetically determined information-sucking reflex, which seeks out, sucks in, and decomposes information, which is later recombined in many ways, growing the information-processing architecture and many diverse recombinable competences.

HOW ???

There is more discussion of this video in

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0603>

Child as scientist 2: Failing to deal with hooks at 19 months

1: Lifting two trucks makes the third disengage.
2-3: He picks it up with his left hand & shakes off the hanging truck with his right.
4: He notices the blank end & puts the truck down, rotating it.
5: He makes a complex backward move from crouching to sitting – while leaning forward to pick up the rotated truck.
6: He sees two rings.
7-9: He tries to join the rings, ignoring the hook, fails and gets frustrated, bashing trucks together and making an angry sound.



See the video http://www.jonathans.me.uk/josh/movies/josh34_0096.mpg

Within a few weeks, he had learnt to see and use the hook-affordances. How? (Nobody saw how.)

The Child as Scientist: 3

- The idea that an infant, or possibly an older child, is like a tiny scientist investigating the world is often reinvented.
- It is obviously false if taken literally, for instance, because there are many conceptual, representational and mathematical tools used by scientists that are not available to a child, not even highly talkative and competent four-year-olds.
- A currently popular view, exemplified by work of Alison Gopnik and colleagues online here <http://ihd.berkeley.edu/gopnik.htm> is that young children (or at least their brains!) have the prerequisites for making causal inferences consistent with causal Bayes-net learning algorithms, which deal with conditional probabilities.
- On that view the concept of cause is viewed as concerned with correlations – Humean causation with probabilities replacing universal correlations.
- Another view, implicit in Kant's critique of Hume, points to a deterministic, notion of causation concerned with structures and their interactions. On this view understanding causation is, at least in some cases, akin to proving, or at least understanding, mathematical theorems as in geometry.
- We suggest that the **probabilistic/correlational** (Humean) kind of causality is what most animals have, but humans and maybe a few others also have something deeper: a Kantian, **deterministic, structure-based** understanding of causality – the sort that drives deep science. To be a Kantian scientist is to be, in part, a mathematician.

Humean causation today: Bayesian nets

In recent years, developments of Humean ideas about causation have led many researchers to assume that causal information can be represented in the form of Bayesian nets, which sum up a collection of conditional probabilities – with or without numbers indicating the precise probabilities.

The diagrams can be interpreted as expressing Humean associations:

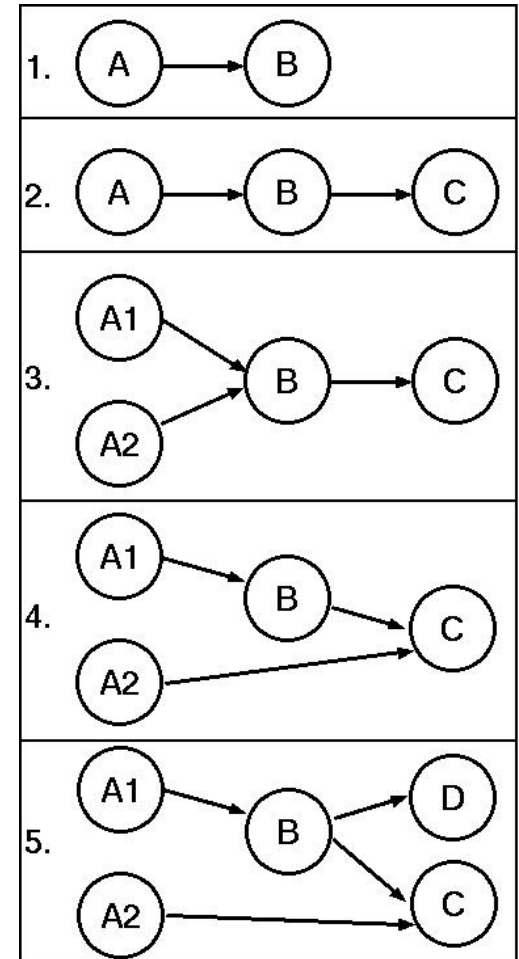
1. could represent: **As cause Bs** (e.g. failed brakes cause crashes)
2. could represent: **As cause Bs and Bs cause Cs**
(failed brakes cause crashes and crashes cause injuries)
3. alters that to: **Bs can be caused either by A1s or by A2s or both**
4. changes the above to represent **A2s being direct co-causes of Cs, along with Bs**
5. Adds **Bs also cause Ds**

Numbers can be added to the diagrams to express prior probabilities and conditional probabilities.

Experimental tests e.g. **clamping** some thing to prevent it from changing, or **directly** altering something to discover what does and does not change, can be used to evaluate a particular Bayesian net as a theory about a set of causal relationships.

Each link can be regarded as representing a true counterfactual conditional statement, or a conditional probability, if numbers are added.

(For a gentle introduction see: Steven Sloman, *Causal Models*, OUP, 2005)



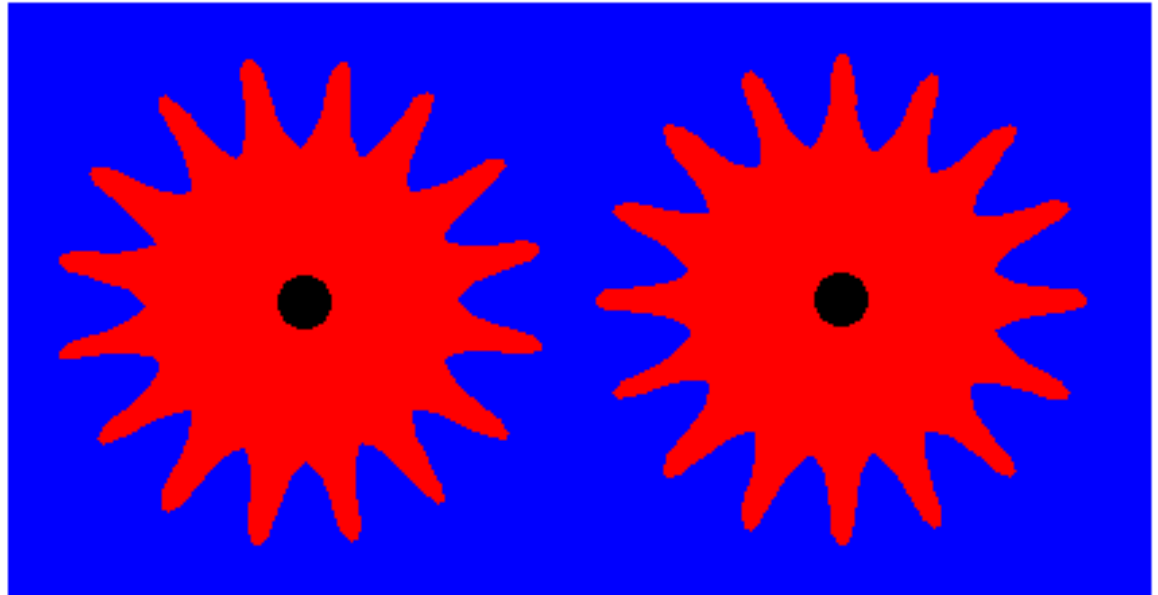
Causation in gear wheels (Humean)

Two gear wheels attached to a box with hidden contents.

Can you tell what will happen to one wheel if you rotate the other about its central axis?

Not on the basis of what you see.

You (or a child) can build a Bayesian net by experimenting.



- By experimenting you may or may not discover a correlation between the rotations – depending on what is inside the box.
- In more complex cases there might be a knob or lever on the box, and you might discover that which way the second wheel rotates depends on the position of the knob or lever. (Compare learning about gears in driving a car.)
- In still more complex cases there may be various knobs and levers, modifying one another's effects through hidden mechanisms. There could also be motors turning things in different directions, competing through friction devices, so that the fastest one wins, etc. etc.

See the work of Alison Gopnik

Causation in gear wheels (Kantian)

Gear wheels closer together and meshed:

You (and some children) can tell, by looking, how rotation of one wheel will affect the other.

HOW? You can simulate rotations and observe the consequences.

What you can see includes the following:

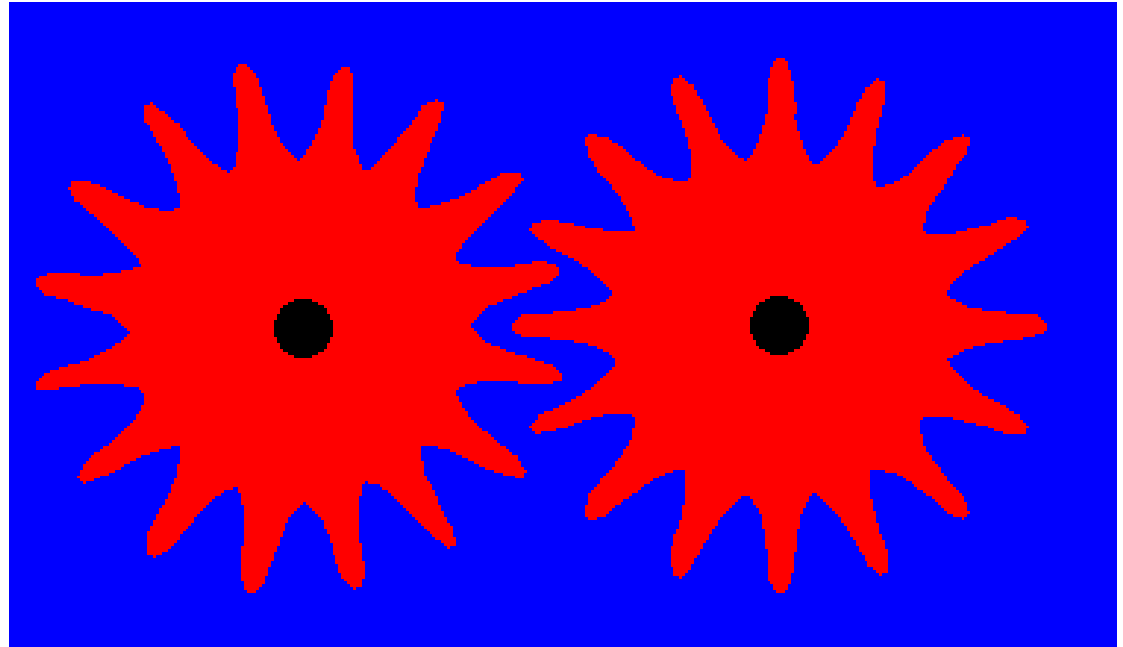
As a tooth near the centre of the picture moves up or down it will come into contact with a tooth from the other wheel.

If both are rigid and impenetrable, then if the first tooth continues moving, it will push the other in the same direction, causing its wheel to rotate in the opposite direction.

AN EXAMPLE OF A MULTI-STRAND PROCESS

Compare understanding causation in a lever.

(We are not claiming that children need to reason **verbally** like this, consciously or unconsciously: explaining the forms of representation used is work to be done.)



Understanding, simulating, proving

Kantian reasoning about effects is partly like running a simulation, except that it does not require precise details of the process to be simulated.

The simulation happens at a level of abstraction that is comparable with reasoning about a theorem in Euclidean geometry.

E.g. such and such a construction will produce a line that divides the triangle into two parts with the same area.

We don't need perfect precision in our diagrams or our simulations – because they are representations of processes, not replicas.

See Sloman, (IJCAI 1971), Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence

<http://www.cs.bham.ac.uk/research/cogaff/04.html#200407>

Not just geometry

Besides perceived shape, and interactions between shapes, the simulations that you run can also make use of **unperceived constraints**:

e.g. **rigidity** and **impenetrability** and the fixed location of the axle.

These concepts need to be **part of the perceiver's ontology** and integrated into the simulations, for the simulation to be deterministic.

Constraints and processes using them need not be conscious, or expressed in linguistic or logical form.

There is no suggestion that a child can do all this from birth.

How this competence develops, and what information processing is involved, remains to be explained.

A Doodle: What do you see?

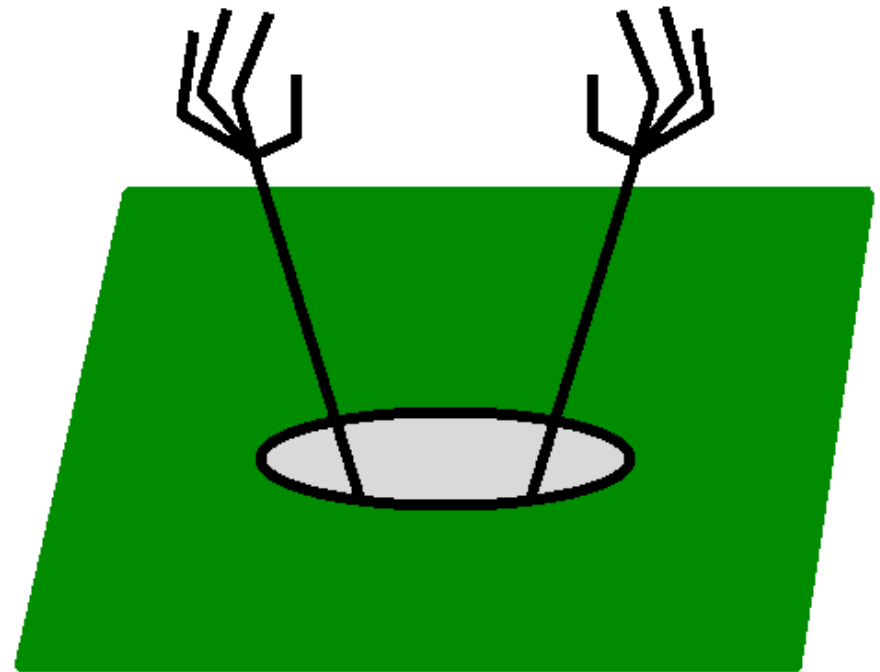
In many cases what you see is driven by the sensory data interacting with vast amounts of information about sorts of things that can exist in the world.

But doodles demonstrate that in some cases where sensory data do not suffice, a verbal hint can cause almost instantaneous reconstruction of the percept, using contents from an appropriate ontology.

Often understanding a doodle requires the ability to do Kantian causal reasoning about unseen mechanisms and processes: you can visualise a process producing the depicted result.

See also

<http://www.doodles.com/archive.html>



Verbal hint for the figure: 'Early worm catches the bird' or 'Early bird catches very strong worm'

Limitations of Humean causation

The Humean conception of causation in general, and the Bayesian net model in particular, treats causation as a relation between events or states of affairs without representing the kind of detail required for Kantian causal understanding.

This does not use the ability of humans and some other animals to understand causation in terms of **the operation of structured mechanisms**, like gears, levers, strings, pulleys, where causation involves concurrent interaction between multiple parts.

Something other than Humean causation is needed to explain causal understanding based on our ability to see **structured 3-D processes** in which

- there are multiple relationships between objects and parts of objects
- different relationships change concurrently
 - **some continuously**
(e.g. direction, distance, angular or linear velocity, acceleration, pressure, location of contact, etc.)
 - **some discretely**
(e.g. contact beginning or ending, containment changing, collision impending or not, something moving into or out of view, occlusion, containment, obstruction starting or ending, etc.)
- some changes are **necessarily** connected with others
(e.g. because of the geometry of the situation: if X continues in the same direction it **must** hit Y)
- There are **additional, non-geometric constraints** that come from physical properties of the materials involved, e.g. rigidity, impenetrability, elasticity, density, stickiness, etc.

The over-rated “DO” operator: for surgery on nets

There is a long philosophical tradition emphasising the role of experimentation in discovering what causes what – and this can lead to a spurious distinction between learning from observation and learning from doing.

This idea is sometimes enshrined in the notion of a **DO** operator on a Bayesian net.

E.g. If A-events and B-events often occur together, and never alone, and a C-event follows immediately, you may conjecture that A and B together cause C-events.

If you can “**DO**” A while you “**DO**” preventing B from occurring, and you then observe that C occurs, you have learnt that your causal net can have a link from A to C without B being involved: the net needs “surgery” (Pearl).

But there is no logical difference between noticing this when you have done the experiment and noticing it when someone else does the experiment or noticing it when the experiment is produced without any human agency by the wind blowing in such a way as to produce A without producing B.

- The main difference between passive observation and active intervention is **speed of learning**: you may have to wait a long time to observe something passively.
- Active intervention can speed up discovery of new relations and testing of hypotheses but it does not provide stronger evidence than passive observation: it merely provides more, faster.
- Of course people may **treat** the results of their own interventions as somehow providing better evidence, but that is just one of many ways in which human reasoning can be fallacious.
- **The study of how people actually reason (psychology) should not be confused with the study of valid reasoning (logic and mathematics).**

The “DO” operator was invented by Evolution

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

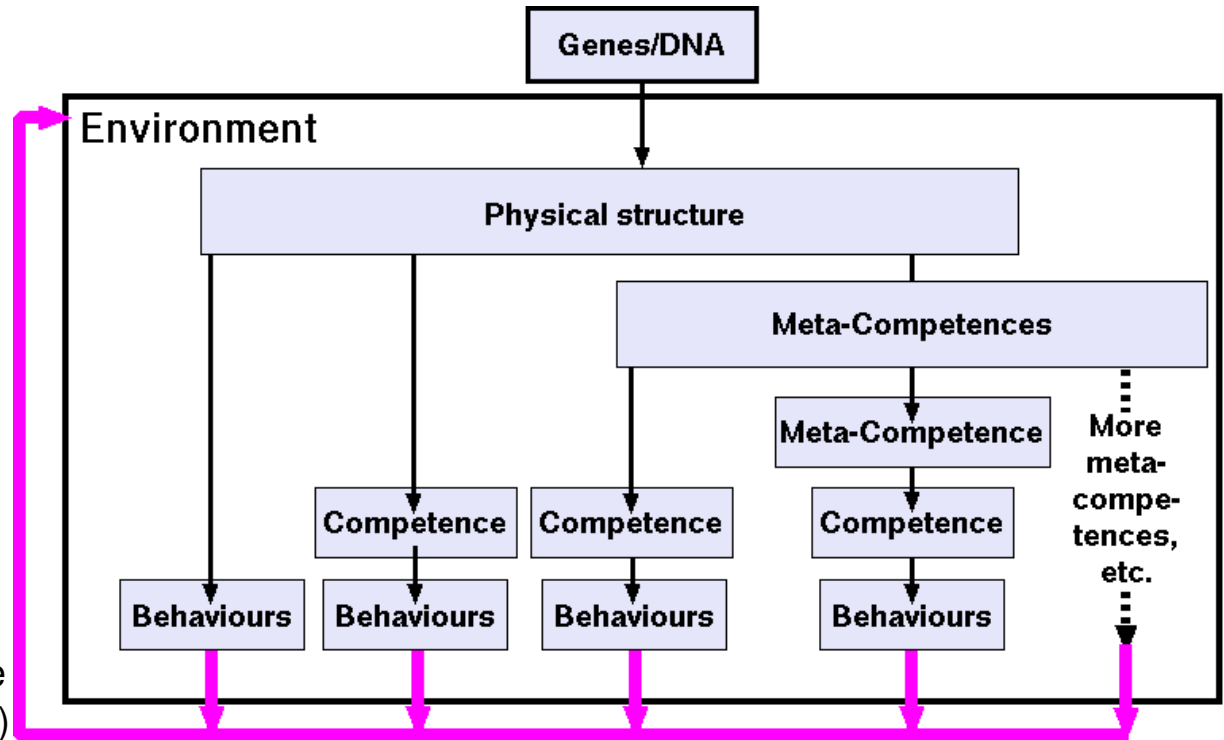
Pre-configured competences:

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 DNA.

They occur towards the left.

Meta-configured competences:

are produced through the **interaction** of pre-configured or previously produced meta-configured competences with the environment (internal or external)



(Chris Miall helped with the diagram.)

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**.

PERHAPS THE INFANTS THAT STARE LONGER ARE TRYING TO DEBUG A THEORY?

A “DO” operator is more powerful when combined with Kantian causal models.

Don't ask which animals or machines understand causation

There is no fixed, sharp, distinction between understanding and not understanding causation.

Instead we can distinguish a fairly long list of types of competence related to causation, and then ask:

- which animals have which subset,
- at what age they typically develop,
- how those competences are acquired or extended,
- what forms of representation, mechanisms and architectures support them
- what the trade-offs are between alternative sets of competences and alternative implementations,

See the slides on this here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/causal-competences.pdf>

That way we can replace futile debates about how to label phenomena with productive research, on what sorts of competences different animals have, what the implications of those competences are, and what mechanisms can explain them.

We can do that for many debates about cognition and development in humans, animals and machines.

I.e. replace ill-defined and poorly motivated dichotomies with analysis of spaces of possible designs and corresponding niches to be analysed, compared, explained, modelled, ...

Some limitations of current AI

Alas, current AI systems, including robots, have only a small set of the previous types of causal competence:

e.g. they may be able to do things,
but, when doing something, don't know
what they do, why they do it, how they do it,
what they did not do but could have done
why they did not do it differently,
what would have happened
if they had done it differently,
etc.

They also cannot tell whether someone else
is doing it the same way or a different way,
and if something goes wrong explain why it went wrong.

All this is related to lacking Kantian causal competences.

Advantages of Kantian Causation

- If the only way you can find out what the consequence of an action will be is by trying it out to see what happens, you can acquire knowledge of causation based only on observed correlations.
 - This is 'Humean causation' – David Hume said there was nothing more to causation than constant conjunction, though modern theories have enriched Hume's ideas in the form of Bayesian nets.
- However if you don't need to find out by **trying** because you can **work out** consequences of the structural relations (e.g. by running a simulation that has appropriate constraints built into it) then you are using a different notion of causation:
 - This is Kantian causation, which is **deterministic, structure-based and generative**:
it supports understanding of novel situations, and designing new actions and machines.
- As children learn to understand more and more of the world well enough to run deterministic simulations they learn more and more of the Kantian causal structure of the environment, including how changing or moving spatial structures interact.
- Typically in science causation starts off being Humean until we acquire a deep (often mathematical) theory of what is going on: then we use a Kantian concept of causation.
 - At first sight Quantum Mechanics refutes this. But
 - (a) the QM wave function is deterministic, and
 - (b) a mathematically specified statistical mechanism can also have determinate statistical properties that follow from its structure.

Kant's most famous example: viewing a house

What will you see when you move round a house, in one direction or the other; or when you move from top to bottom of a house, or from bottom to top? See *The Critique of Pure Reason* Second Analogy

What Kant intended by this example is open to interpretation, but, whatever he meant, the example nicely makes the point that if you understand the geometric and topological structure of the environment you can deterministically predict the consequences of various movements.

- Actions that alter your location cause different parts of the house to change their appearance or become visible or invisible in **systematic** ways, which depend on the structure of the house.
- Some changes are **continuous** (e.g. continuous foreshortening) and some **discrete** (e.g. as a previously invisible face of the house becomes visible, or vice versa): and both sorts can occur at the same time.
- As with causation in levers, pulleys, or engaged gears, this is a collection of **effects that are inferrable** on the basis of an understanding of the structure of the situation: you can **work out** the effects of different combinations of movements.

As illustrated in Sloman (1971) 2nd IJCAI (Reference above)

- This is closely related to our ability to do mathematical reasoning, especially geometrical and topological reasoning.

Compare the use of 'aspect graphs' in computer vision, and Minsky's 'Frame systems' (1973?).

For more detailed discussion of Kant's views see, for example,

<http://www.trinity.edu/cbrown/modern/litrev/Kant-causation.html>

Kant's Second Analogy: Literature Review by Melanie Butler (?)

on Curtis Brown's web site: <http://www.trinity.edu/cbrown/>

Sensorimotor knowledge and Kantian causation

The 'sensorimotor' approach in an extreme form assumes that everything an organism needs to know is expressed in **sensorimotor contingencies**: conditional probabilities linking patterns in sensory and motor signals, at various levels of abstraction. It can also be presented as a theory of consciousness.

(O'Regan, Noë, Hurley and others. <http://lpp.psychol.univ-paris5.fr/tikiwiki/tiki-index.php>)

It has strong connections with 'symbol grounding' theory. It may provide an adequate characterisation of how most simple organisms represent the environment, e.g. microbes and perhaps insects most of the time. But as soon as a representation is of 3-D spatial structures, processes, relationships, and allows unobservable but inferrable properties of kinds of stuff, as we have suggested supports Kantian conceptions of causation, it goes beyond sensorimotor representations and uses an a-modal exosomatic ontology.

From that viewpoint "mirror neurones" should have been called "abstraction neurones".

It is argued elsewhere that one reason such ontologies and forms of representation evolved was to drastically reduce the combinatorics of representing grasping and other actions performed by hands that move independently of each other and of the eyes (unlike mouth and beak).

See <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/nature-nurture-cube.html>
How to see a rotating 3-D wireframe cube (Necker cube).

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0601>
Learning orthogonal recombinable competences.

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0603>
Critique of sensorimotor contingency theory.

Could the portia spider perform her amazing feats if she had only representations of her sensory and motor signals and relations between them? See <http://www.freerepublic.com/focus/f-chat/1640513/posts>

No robots have Kantian causal understanding yet

Current work on robots and AI systems has not addressed the ability to acquire and use Kantian understanding of spatial structures and processes and their interactions, or to learn about different kinds of “stuff” and their properties
e.g. rigidity, impenetrability, elasticity, hardness, softness, etc.

That’s mostly because

- (a) those working on vision have focused largely on **recognition**, rather than **understanding structure**,
- (b) researchers have not noticed the need to give machines Kantian causal understanding,
- (c) researchers have not given machines the ability to think and reason about their causal reasoning
- (d) giving machines an understanding of spatial interactions is very difficult.

Work on fault diagnosis in circuits could be regarded as an exception to (b), and some early work on “naive physics” (Pat Hayes) and “qualitative” physics (by Kuipers and others) long ago, now mostly abandoned(?). (But see the ongoing work by Ken Forbus and colleagues and work on qualitative reasoning by Tony Cohn.)

The parts need to be assembled in an architecture

How can we put everything together? We need to adopt the **design stance** (Dennett 1978) and make significant use of present and future concepts from information engineering and science. That will reveal a **logical topography** underlying the **logical geography** of concepts currently in use, pointing at the possibility of new deeper conceptualisations, as happened to ordinary concepts of kinds of stuff, following discoveries about the architecture of matter.

See <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/logical-geography.html>

The Birmingham Cognition and Affect project has produced a draft high level classification of types of mechanisms, requiring many further subdivisions (the CogAff schema):

Requirements for subsystems can refer to

- **Types of information** used (**ontology** used: processes, events, objects, relations, causes, functions, affordances, meta-semantic....)
- **Forms of representation** (continuous, discrete, Fregean, diagrammatic, distributed, dynamical, compiled, interpreted...)
- **Uses of information** (controlling, modulating, describing, planning, executing, teaching, questioning, instructing, communicating...)
- **Types of mechanism** (many examples have already been explored – there may be lots more ...).
- **Ways of putting things together** in an architecture or sub-architecture, dynamically, statically.

Perception	Central Processing	Action
	Meta-management (reflective processes) (newest)	
	Deliberative reasoning ("what if" mechanisms) (older)	
	Reactive mechanisms (oldest)	

In different organisms or machines, boxes contain different mechanisms, with different functions and connectivity, with or without various forms of learning. In some the architecture grows itself after birth.

A special case of CogAff: H-CogAff

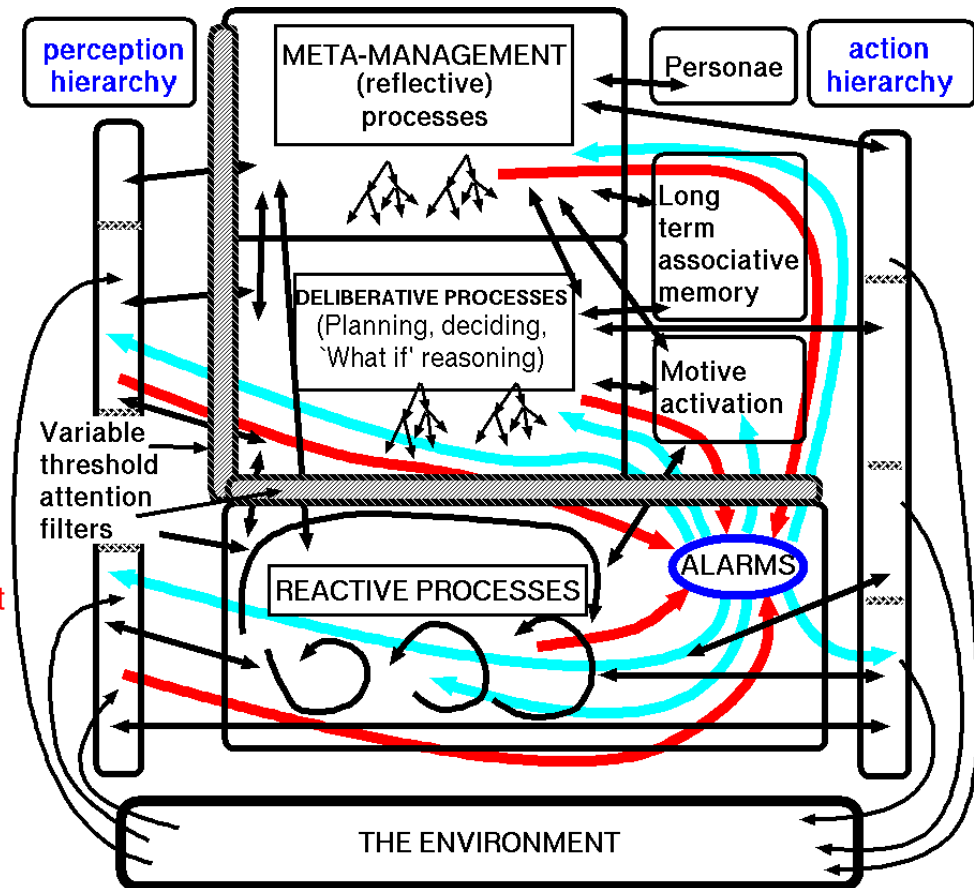
H-CogAff specifies a sub-class of human-like architectures within the generic “CogAff” schema. (“H” stands for “Human”)

This is a sketchy indication of some of the required subsystems and how they are connected. Note the implication that both vision and action subsystems have several different (concurrently active) layers of functionality related to the different central layers/mechanisms.

Where could this come from?

Different trajectories for different layers:

- evolutionary,
 preocial competences from the genome
- developmental,
 altricial competences and architectures built while interacting with the environment
- adaptive changes, (small adjustments)
- skills compiled through repetition
- social learning, including changing personae...



Much work remains to be done.

Kantian causal understanding and reasoning probably cannot occur in the reactive layers. Why not? Different variants may occur in deliberative and metamanagement layers.

For more details, see the presentations on architectures here: <http://www.cs.bham.ac.uk/research/cogaff/talks/>

Can non-human animals use Kantian causation?

Jackie's talk will discuss this.

See her slides – two versions: for screen viewing (with hyperlinks) and for printing:

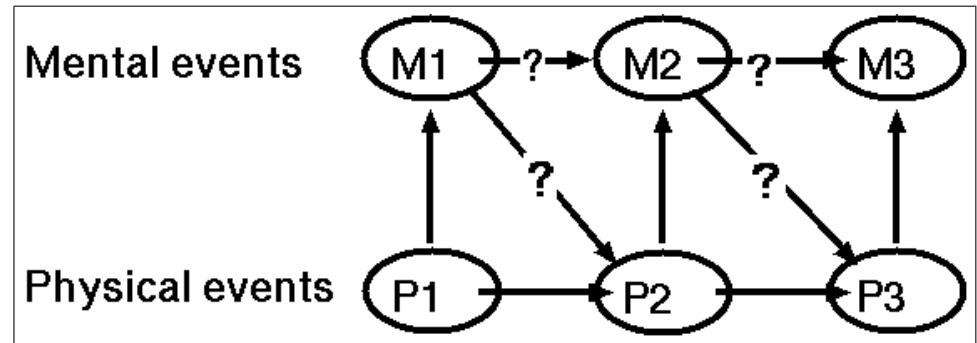
<http://www.cs.bham.ac.uk/~jmc/chappell-wonac-slides-screen.pdf>

<http://www.cs.bham.ac.uk/~jmc/chappell-wonac-slides-print.pdf>

Virtual machine events can also be causes

We now know that something like this can exist in computers, though there are people who dispute the causal links marked with '?'s.

Events in virtual machines, like inserting a character in a line, or a spelling checker correcting a typing mistake, or a chess program deciding to attack your queen are like mental events in many ways: e.g. such events have no physical location, cannot be observed using physical sensors, and are defined by concepts not definable in terms



of the concepts of the physical sciences. Compare mental events.

Understanding all that is important for scientists (and philosophers) trying to understand minds, brains, and what they do and how they do it.

People who challenge the links marked “?” (especially downward links) regard events in minds and computer virtual machines as **epiphenomenal** – i.e. as effects but not causes.

This denial is based on a false theory of causation (e.g. if it is thought of as a kind of “conserved fluid” flowing through the universe).

If instead we regard causal talk as merely a way of expressing which of a (rather complex) set of counterfactual conditionals are true and which false, then virtual machine events can be causes.

(But that is not the main topic of this talk.)

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

Who knows about virtual machine causes and effects?

Not everything that can learn about and make use of causal facts is capable of thinking about or learning about causation in virtual machines.

However that is part of what we do as scientists, philosophers, gossips, novelists, novel readers, etc. and as users or designers of many software packages (computer games, spelling checkers, operating systems,).

In principle it is possible that some non-human animals can acquire and use this sort of knowledge about causation in virtual machines, but that is not something we shall discuss.

A special case of this is being able to think about mental states and processes in others, e.g. the contents of their perceptions, their desires, their values, their beliefs, their decisions, their intentions, etc. This is not an all-or-nothing capability: many special cases may be known about without others being known.

Although there is much (albeit narrowly focused) research on when children acquire these abilities, and research on which other animals can do it, the topic is too complex to be discussed further here.

A research agenda: Learning, development and Kantian Causal Competences

This is a large research topic requiring deep interdisciplinary collaboration between philosophers, robot designers, mathematicians, developmental and clinical psychologists, biologists, neuroscientists, educationalists,

Maybe an outcome of this workshop could be a research agenda, with some well defined early steps?

—

The ability to do causal reasoning in different domains has to be learnt.

The ability to **work out consequences** requires learning to build simulations with appropriate structures, appropriate permitted changes, and appropriate constraints.

What is appropriate depends on what is being simulated: simulating the rotation of a rigid gear wheel (e.g. one made of steel) is not the same as simulating the rotation of something soft and malleable, e.g. putty or plasticine.

Appropriate constraints ensure the right counterfactual conditionals are true as the simulation runs.

The detailed representational, algorithmic, mechanistic and architectural requirements to support such learning, and the growth of the ontology involved, require much deeper analysis than we can give at present.

Part of the point of the CoSy Robot project is to investigate these issues, especially the **requirements** for human-like competence, which we need to understand before we can build **designs** or **implementations**, though the process of designing and implementing can help the process of understanding requirements.

But the problems are very hard and progress is very slow.

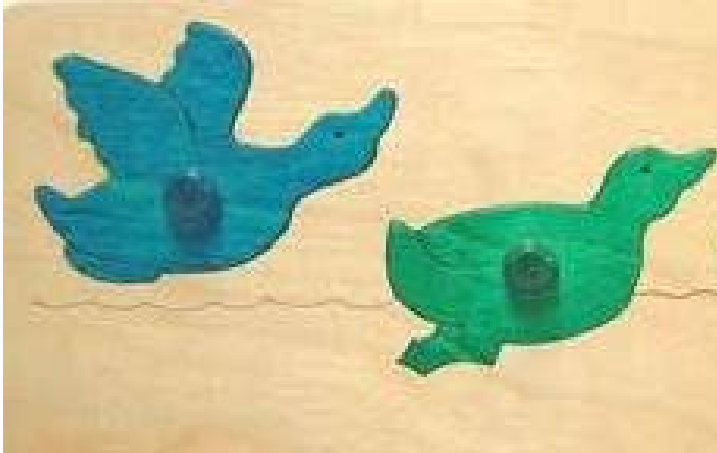
Moreover we don't think neuroscientists have discovered appropriate mechanisms either (in 2007).

See Arnold Trehub: *The Cognitive Brain* (1991) <http://www.people.umass.edu/trehub/>

For more detail on a theory of vision as involving running of simulations see
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0505>

We cannot do it all from birth

Causal reasoning adults find easy can be difficult for infants.



A child learns that it can lift a piece out of its recess, and generates a goal to put it back, either because it sees the task being done by others or because of an implicit assumption of reversibility. At first, even when the child has learnt which piece belongs in which recess there is no understanding of the need to line up the boundaries, so there is futile pressing. Later the child may succeed by chance, using nearly random movements, but the probability of success with random movements is **very** low. (Why?)

Memorising the position and orientation **with great accuracy** will allow toddlers to succeed: but there is no evidence that they have sufficiently precise memories or motor control.

Stacking cups compensate for that partly through use of symmetry, partly through sloping sides, so they are much easier.

Eventually a child understands that unless the boundaries are lined up the puzzle piece **cannot** be inserted. Likewise she learns how to place shaped cups so that one goes inside another or one stacks rigidly on another.

Conjecture:

Each such change requires the child to extend its ontology for representing objects, states and processes in the environment, and that ontology is used in a mental simulation capability. **HOW?**



Learning ontologies is a discontinuous process

- The process of extending competence is not continuous (like growing taller or stronger).
- The child has to learn about **new kinds** of
 - objects,
 - properties,
 - relations,
 - process structures,
 - constraints,...
- and these are different for
 - rigid objects,
 - flexible objects,
 - stretchable objects,
 - liquids,
 - sand,
 - mud,
 - treacle,
 - plasticine,
 - pieces of string,
 - sheets of paper,
 - construction kit components in Lego, Meccano, Tinkertoy, electronic kits...

We don't know how many different things of this sort have to be learnt, but it is easy to come up with many significantly different examples.

There is no fixed order in which things have to be learnt: there are many dependencies but not enough to generate a total ordering – each learner finds routes through several partially ordered graphs.

Think also of learners with physical and cognitive disabilities and abnormalities.

CONJECTURE

In the first five years

- a child learns to run at least hundreds,
- possibly thousands, of different sorts of simulations,
- using different ontologies
 - with different materials, objects, properties, relationships, constraints, causal interactions —
some opaque and Humean others transparent and Kantian.
- **and throughout this learning, perceptual capabilities are extended by adding new sub-systems to the visual architecture, including new simulation capabilities**

In the case of humans, things available to be learnt keep changing from one generation to another: provision of new kinds of playthings based on scientific and technological advances is a major form of communication across generations.

However, some three year olds today do things that require ontologies undreamed of even by their grandparents.

Some more examples are available in

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0601>

COSY-DP-0601 Orthogonal Competences Acquired by Altricial Species (Blanket, string and plywood).

Pushing and pulling

As toddlers learn to push, pull and pick things up, they find that some things 'hang together': if you move a part other parts move. But the growing ontology, and mechanisms for representing actions and their perceived effects need to allow for things that hang together in different ways.

If a group of bricks is lying on the floor, pushing a brick on the boundary towards the centre can make the whole group move, whereas pulling it in the opposite direction moves no other brick.

On the other hand if you push the edge of a blanket towards the centre most of the blanket does not move, whereas if you pull the edge away from the centre the blanket follows (in an orderly or disorderly fashion, depending on how you pull, with one or two hands, etc.).

A sheet of paper the same size as the blanket will typically behave differently: pushing and pulling will move the whole sheet, but the effect of pushing will be different from pushing a pile of bricks (in what ways?) and the effect of pulling will be different from pulling the blanket (in what ways?).

What they have in common includes the fact that if a toy is resting on the blanket or sheet of paper, pulling the edge towards you brings the toy closer too, whereas if you pull too fast, or if the toy is on the floor near the far edge, pulling will not have that effect. Why not?

The child's ontology has to allow not only for different kinds of stuff (cloth, wood, paper, string, etc.), but also different ways in which larger wholes can be assembled from smaller parts: which requires a grasp of relations of different kinds, including 'multi-strand relations', and the 'multi-strand processes' that occur during changes in multi-strand relations, as discussed in

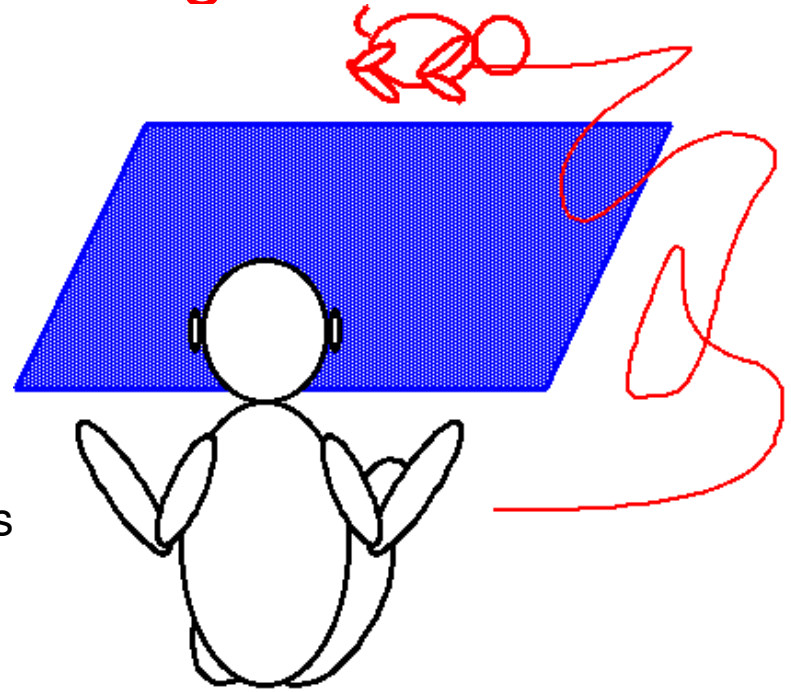
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0507>

Some of the understanding of causation in such processes may start off Humean (i.e. using only conditional probabilities) and then as the ontology is enriched to include properties like *rigid*, *flexible*, *impenetrable*, *elastic*, *inextensible*, and these are combined with shape and spatial relations, the understanding can become more Kantian, i.e. structure-based, generative and deterministic, supporting more creative exploration and discovery.

Blanket and String

If a toy is beyond a blanket, but a string attached to the toy is close at hand, a very young child whose understanding of causation involving blanket-pulling is still Humean, may try pulling the blanket to get the toy.

At a later stage the child may either have extended the ontology used in its conditional probabilities, or learnt to simulate the process of moving X when X supports Y, and as a result does not try pulling the blanket to get the toy lying just beyond it, but uses the string.



However the ontology of strings is a bag of worms, even before knots turn up.

Pulling the end of a string connected to the toy towards you will not move the toy if the string is too long: it will merely straighten part of the string.

The child needs to learn the requirement to produce a straight portion of string between the toy and the place where the string is grasped, so that the fact that string is inextensible can be used to move its far end by moving its near end (by pulling, though not by pushing).

Try analysing the different strategies that the child may learn to cope with a long string, and the perceptual, ontological and representational requirements for learning them.

Ontologies for getting at something

Understanding varieties of causation involved in learning how to get hold of a toy that is out of reach, resting on a blanket, or beyond it.

Some things to learn through play and exploration

Toy on short blanket **Grab edge and pull**

Toy on long blanket **Repeatedly scrunch and pull**

Toy on towel **Like blanket**

Toy on sheet of plywood

Pull if short(!), otherwise crawl over or round it

Toy on sheet of paper **Roll up?**

(But not thin tissue paper!)

Toy on slab of concrete **Crawl over or round**

Toy at end of taut string **Pull**

Toy at end of string with slack **Pull repeatedly**

String round chair-leg **Depends**

Elastic string

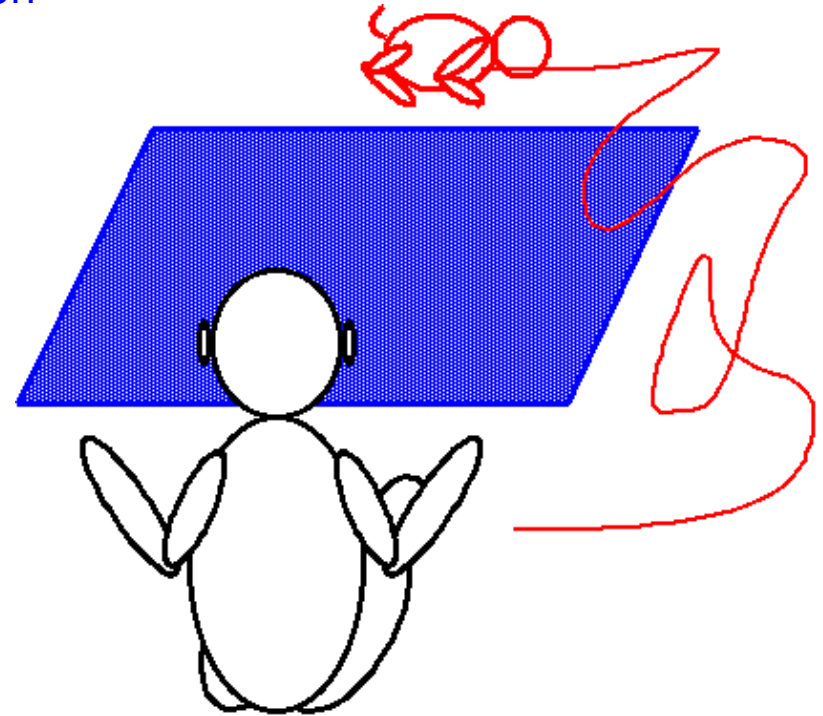
?????

See this discussion of learning orthogonal recombinable competences

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0601>

It takes a lot of learning to develop all the visual and reasoning competences required for seeing and understanding these affordances – including visualising what would have happened if you had done something different, or if someone else were to move something.

Our spatial and visual competence goes far beyond actually doing.



Tamarins at work

Jackie will talk about the following

Hauser, M.D., Kralik, J. & Botto-Mahan, C. (1999).

Problem solving and functional design features: Experiments with cotton-top tamarins (*Saguinus oedipus*). *Animal Behaviour* 57: 565-582.

Available online

<http://www.wjh.harvard.edu/~mnkylab/publications/learnconcepts/problemsolving.pdf>

They conclude

...First, cotton-top tamarins readily solve means-end tasks. Second, in solving such tasks, the tamarins attend to the functionally relevant features. Featural transformations that do not affect an object's functionality were readily tolerated with regard to performance in the means-end task. This shows that tamarins can discriminate between objects that show signs of good design and those that do not, and can use this knowledge to select an appropriate object for solving a problem. Third, some featural transformations that were clearly relevant to the functionality of the task proved difficult for the tamarins and required explicit training. Thus, although the tamarins appear to have a greater capacity for creating learning sets than revealed by some early experiments on the closely related marmoset (Miles & Meyer 1956; see also Rumbaugh et al. 1996), there are limitations. At present, it is not clear whether these limitations arise due to their relatively poor dexterity, to problems in making fine-grained perceptual discriminations (e.g. narrow gaps in the connected problem), to conceptually mediated problems associated with understanding means-end problems (e.g. understanding abstract-relational concepts), or to some combination of these factors. Fourth, although the tamarins were able to inhibit some actions, some problems proved difficult because of their inability to inhibit a reaching response. Future work must assess the degree to which problems of inhibition over-ride a species' capacity to solve problems conceptually and the extent to which interspecific differences in the prefrontal cortex contribute to interspecific differences in problem-solving ability.

The summary assumes that the tamarins have a fairly rich ontology.

We can ask whether that ontology was pre-configured or meta-configured, and a host of further questions arising, whichever the answer is, e.g. questions about precisely what information is represented, how the information is represented, how the representations are manipulated, how the results are used, etc.

These questions would have to be answered in designing robots with similar capabilities.

Cloth and Paper: Learning never ends



You have probably learnt many subtle things unconsciously, some as an infant or toddler, some later on, about the different sorts of materials you interact with (e.g. sheets of cloth, paper, cardboard, clingfilm, rubber, plywood).

That includes different ways in which actions can and cannot distort their shape.

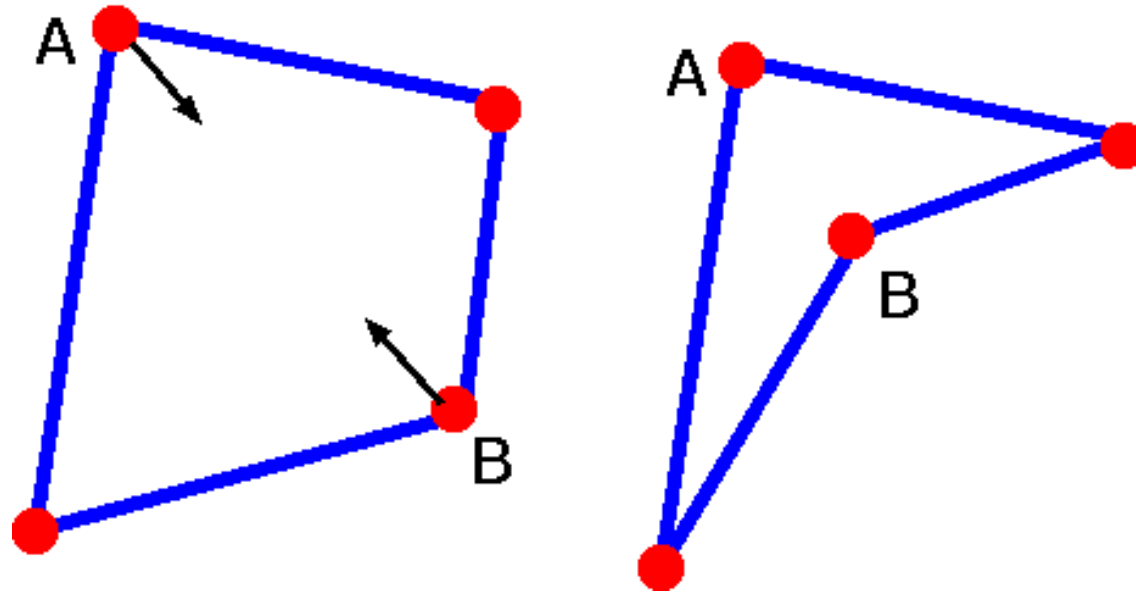
Lifting a handkerchief by its corner produces very different results from lifting a sheet of printer paper by its corner – and even if the handkerchief had been ironed first (what a waste of time) it would not have behaved like paper.

Most people cannot simulate the **precise** behaviours of such materials mentally but we can impose constraints on our simulations that enable us to deduce consequences.

In some cases the differences between paper and cloth will not affect the answer to a question, e.g. in questions about results of folding processes that depend only on shape, not material.

Simulating motion of rigid, flexibly jointed, rods

A Kantian example: on the left, what happens if joints A and B move together as indicated by the arrows, while everything moves in the same plane? Will the other two joints move together, move apart, stay where they are. ???



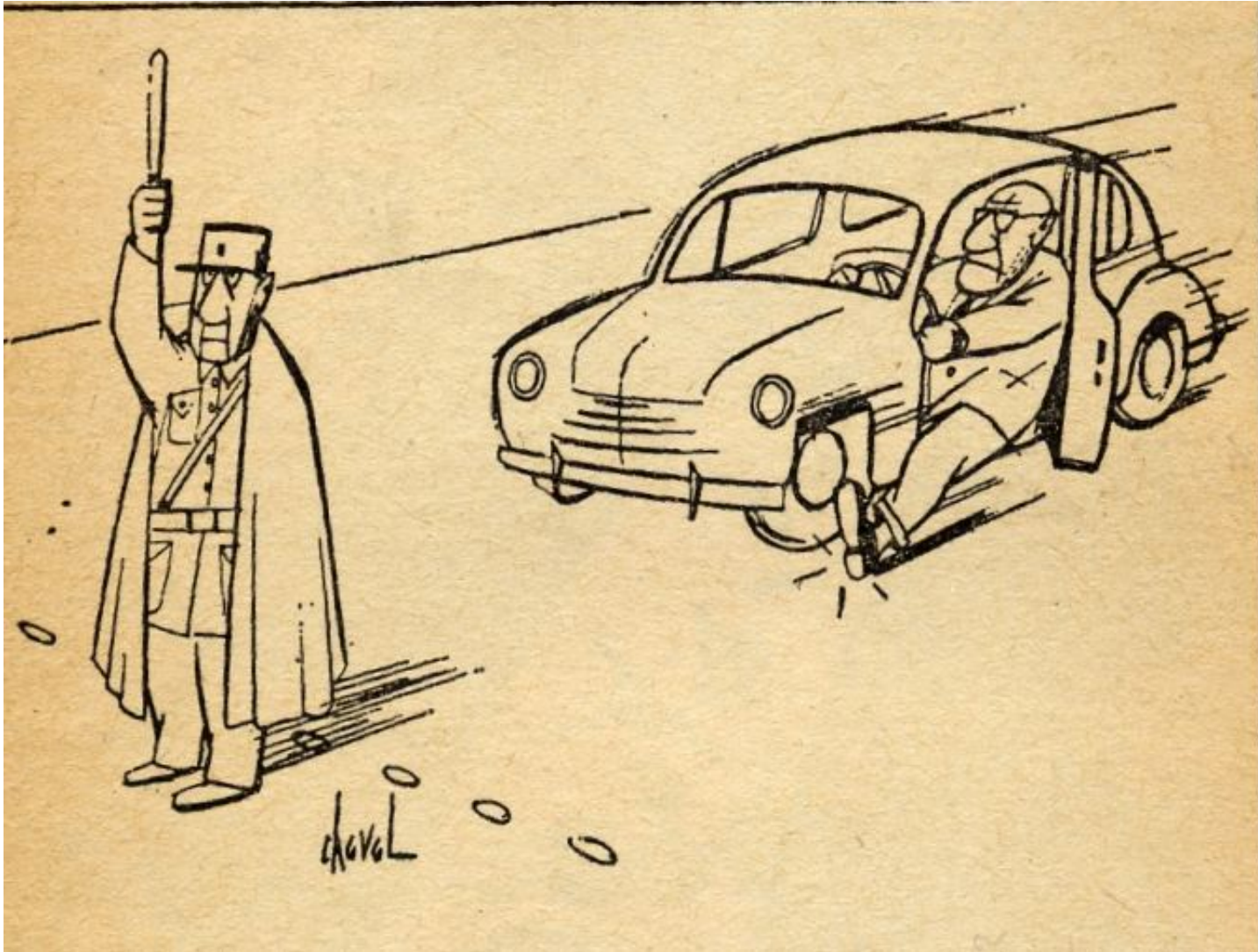
- What happens if one of the moved joints crosses the line joining the other two joints?
- This task is harder than the gears task (why?).
- We can change the constraints in our simulations: what can happen if the joints and rods are not constrained to remain in the original plane?
- What has to develop in a child before such tasks are doable?

Much, or most, perception is of processes

We live in an environment in which there is constantly motion of many kinds, including motion of perceivers, and causal relationships between structures and the various perceived motions.

Examples show that even interpretation of static pictures can include perception of processes, past or future, and causal relations involving those processes.

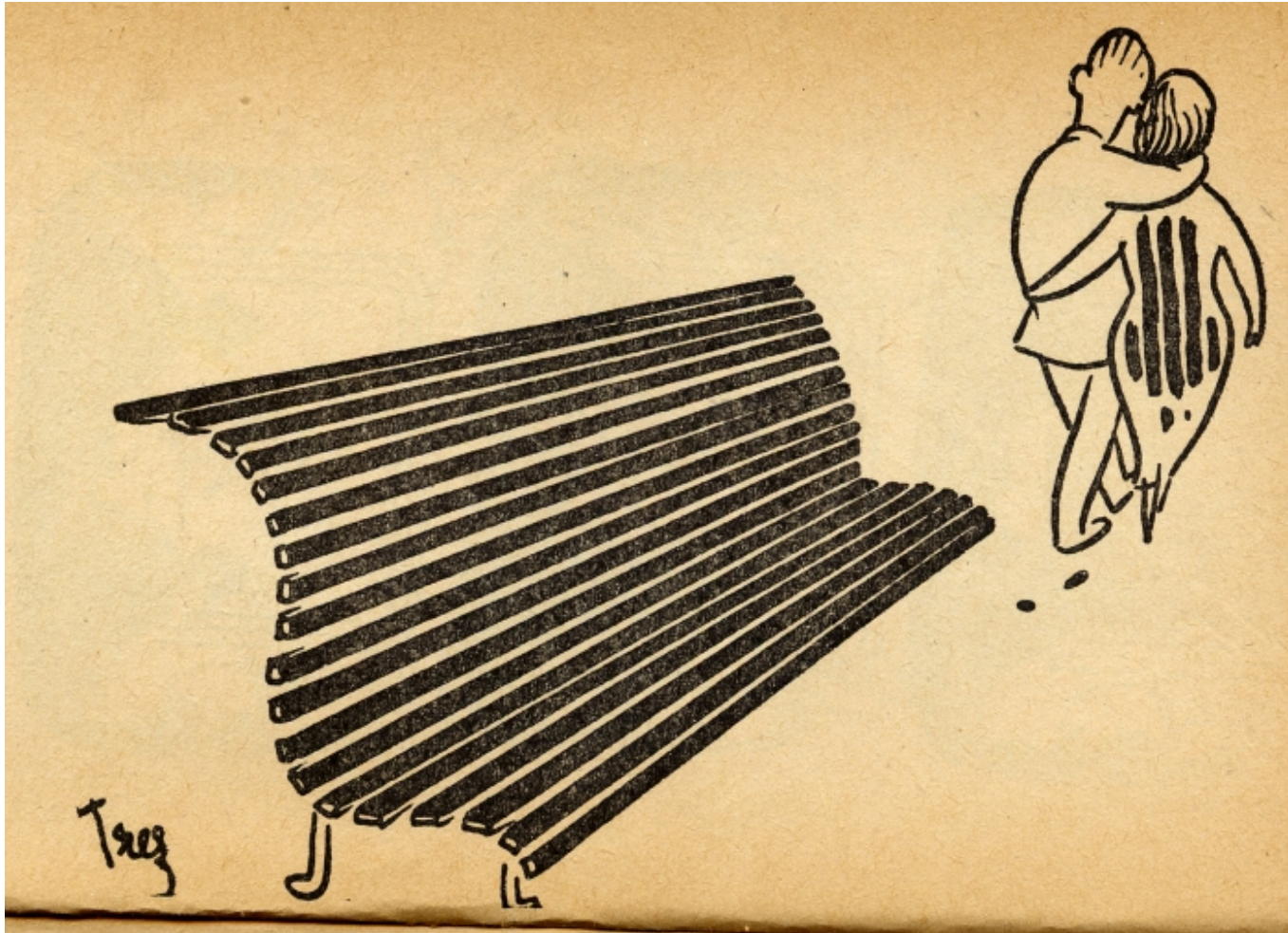
What do you see?



Perhaps you see a process extending to a future time?

And causal connections?

What do you see?



Various objects and relationships of different sorts

Perhaps you see a process starting at an earlier time?

And causal connections?

Visual reasoning about something unseen

If you turn the plastic shampoo container upside down to get shampoo out, why is it often better to wait before you squeeze?

In causal reasoning we often use runnable models that go beyond the sensory information: sometimes part of what is simulated cannot be seen –

a Kantian causal learner will constantly seek such models, as opposed to Humean (statistical) causal learners, who merely seek correlations.

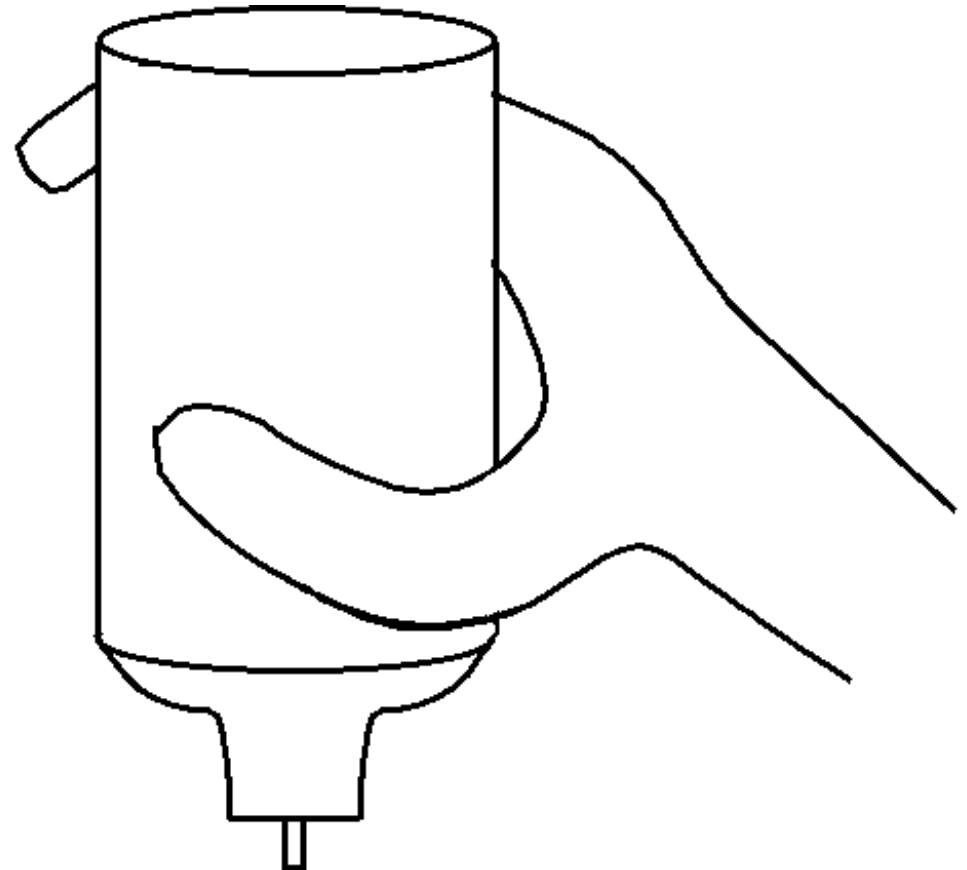
Dare we say this:

There are Kantian sciences and scientists and also Humean ones, who have no idea what they are missing.

Note that the model used here assumes incompressibility and viscosity.

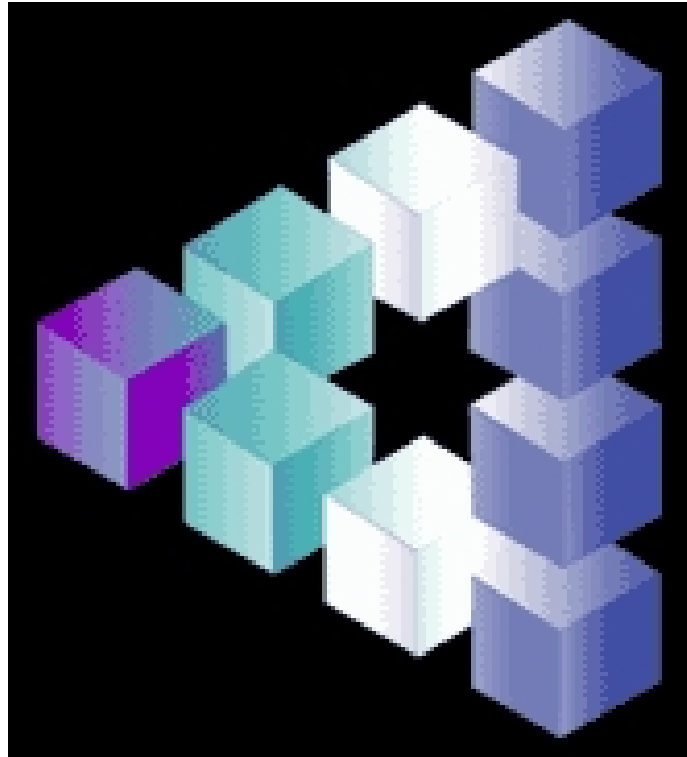
The model can also explain the fact that as more of the shampoo is used up you have to wait longer before squeezing.

The next slide shows that the word ‘model’ as used here allows inconsistency.



Note: a percept does not need to be consistent

A nice picture by Reutersvard – before Penrose



What are the implications of what you see?

Think of all the things you can do with or between the little cubes.

Seeing does not require a **consistent** percept

A model (in the normal sense) cannot be inconsistent: but a model of this scene would have to be.

Escher's Weird World

Many people have seen this picture by M.C. Escher:

a work of art, a mathematical exercise and a probe into the human visual system.

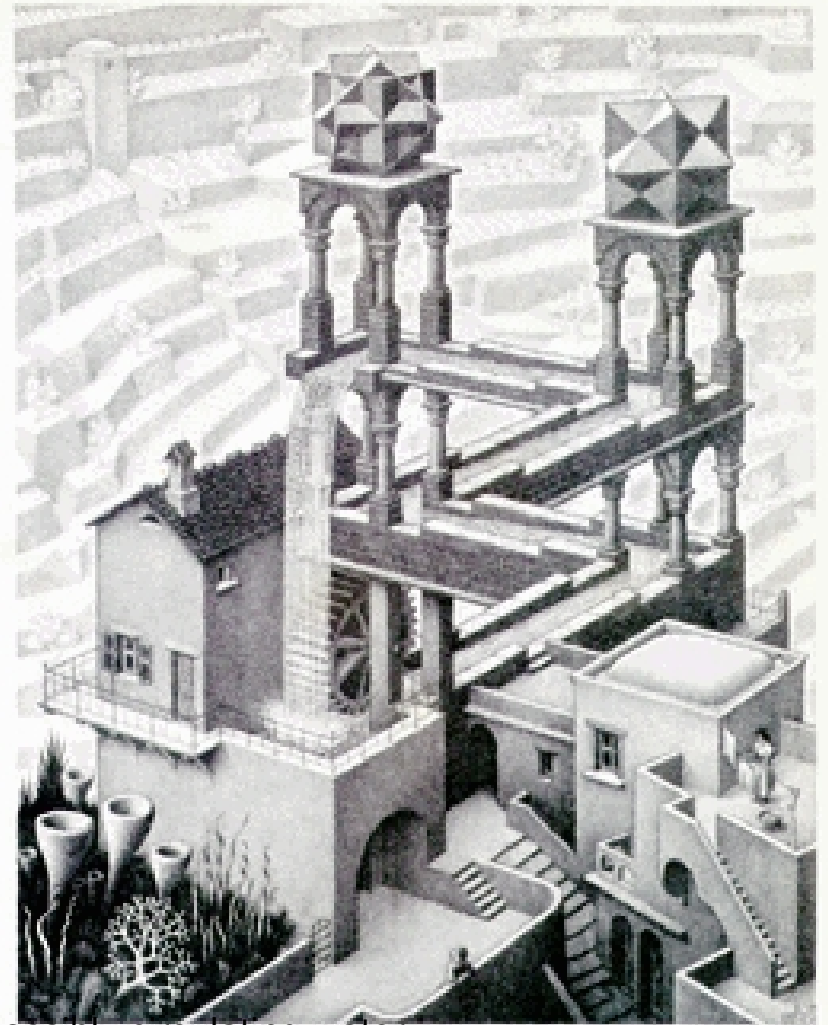
You probably see a variety of 3-D structures of various shapes and sizes, some in the distance and some nearby, some familiar, like flights of steps and a water wheel, others strange, e.g. some things in the 'garden'.

There are many parts you can imagine grasping, climbing over, leaning against, walking along, picking up, pushing over, etc.: you see both structure and affordances in the scene.

Yet all those internally consistent and intelligible details add up to a multiply contradictory global whole. What we see could not possibly exist.

There are several 'Penrose triangles' for instance, and impossibly circulating water.

Can you see the contradictions? They are not immediately obvious.



A visual percept cannot be a model

Models cannot be inconsistent

However if percepts are made up of fragments combined in a manner that does not correspond to full spatial integration then inconsistencies are possible.

E.g.

- A is bigger than B
- B is bigger than C
- C is bigger than A

or, more plausibly, a large collection of proto-affordances of different sorts, spatially located.

Why might the use of such a fragmented, though spatially related, collection of distinct interpretations of portions of the scene be desirable?

Because the very same scene needs to be perceivable in different ways, depending on current goals, interests, etc.

So it must be possible to switch different items of information in and out of the percept.

E.g. different affordances, different relationships, low level or high level details.

This is one form of attention switching.

Biological bootstrapping mechanisms

- There are some types of animals whose needs cannot be served by genetically determined (preconfigured) competences (using pre-designed architectures, forms of representation, ontologies, mechanisms, and stores of information about how to act so as to meet biological needs)

why not?

- Evolution seems to have ‘discovered’ that it is possible instead to provide a powerful **meta-level bootstrapping mechanism** for ‘**meta-configured**’ species:
 - a mechanism without **specific** information about things that exist in the environment (apart from very general features such as that it includes spatio-temporal structures and processes, causal connections, and opportunities to act and learn, and that the neonate has a body that is immersed in that environment)
 - but with specific information about things **to try doing**, things **to observe** things **to store**
 - and with specific information about **how to combine** the things done and records of things perceived into ever larger and more complex reusable structures,
 - sometimes extending its own architecture in the process (e.g. in order to cope with a substantial extension to its ontology)
 - And including a continually extendable ability to run simulations that can be used for planning, predicting and reasoning.

So there are preconfigured and metaconfigured species, or, to be precise species with different mixtures of preconfigured and metaconfigured competences.

Empiricism and Biology

- Empiricists tend to dislike Kantian theories – or more generally theories about ‘innate’ knowledge or innate cognitive competence.
- **But that may be because they don’t know enough biology.**
- The vast majority of animals have most (and in many cases all) of their cognitive competences pre-programmed innately
 - e.g. precocial species such as chickens, deer, and salmon, as well as most most non-vertebrates. Chicks can walk around and peck for food soon after hatching and some deer can run with the herd very soon after birth.
- Many of those can also learn using adaptive mechanisms that produce **relatively slow** kinds of learning based on the statistics of their interactions with the environment (e.g. reinforcement learning)
- But for some types (e.g. corvids, hunting mammals, primates), evolution found an alternative strategy, better suited for neonates starting off in very varied environments, or which require complex cognitive skills in adult life that cannot be provided in the genome
 - e.g. because there is not enough evolutionary time or opportunity to learn, or because of difficulty of encoding the design in available DNA.

At both extremes there is genetically determined competence: but one has **content** determined and the other has **information and mechanisms for acquiring content** pre-determined: the outcomes are very different.

Biological Nativism: Altricial/Precocial tradeoffs

- Evolution ‘discovered’ that for certain species which need to adapt relatively quickly to changing environmental pressures and which perform cognitively demanding tasks as adults, a kind of Kantian learning mechanism is possible that allows much faster and richer learning than is possible in systems that merely adjust probabilities on the basis of observed evidence (statistical data).
- The latter species, with more or less sophisticated forms of the Kantian mechanism, learn a great deal about the environment after birth and in some cases are able rapidly to develop capabilities none of their ancestors had (like young children playing with computer games).
- We conjecture that this uses an information-processing architecture which starts off with a collection of **primitive perceptual and action competences**, and also with a **mechanism for extending those competences by ‘syntactic’ composition**, as a result of play and exploration, which is done for its own sake, not to meet other biological needs (food, protection from hurt, warmth, etc.)
- The meta-level features of the mechanism and the initial competences are genetically determined, but the kinds of composite competences that are built are largely a function of the environment.
- This requires forms of learning that are not simply adjustments of probabilities, but involve continual creation of new useful structures.

Terminological problem

- The labels 'altricial' and 'precocial' are used by biologists with a rather narrow meaning, relating to state at birth.
- We are talking about patterns of cognitive development that seem to be correlated with those differences at birth.

Insects, fish, reptiles, grazing mammals, chickens, are precocial (born/hatched physiologically developed and behaviourally competent),

whereas

hunting mammals, primates, crows, humans are altricial (born/hatched underdeveloped and incompetent), but achieve deeper and broader cognitive competence as adults, with more rapid and creative learning.

- We need new terminology for the cognitive differences.

Perhaps a distinction between

– **preconfigured** (relatively rigid) cognitive development
and

– **non-preconfigured** (relatively flexible and fast path-building?) cognitive development using **metaconfigured** capabilities.

However, organisms with the latter may also include many preconfigured competences e.g. some 'waiting' for puberty in humans.

Alison Gopnik's Work

Alison Gopnik

Alison Gopnik and Laura Schulz (2004)

'Mechanisms of theory formation in young children' in *TRENDS in Cognitive Sciences* Vol.8 No.8

http://ihd.berkeley.edu/gopnik_tics.pdf

Their conclusion

“Although much more research is necessary it seems that infants and young children can detect patterns of conditional probability, understand the nature of their own and others interventions, and to at least some extent, integrate conditional probability and intervention information spontaneously and without reinforcement.”

There are different theses here:

- Children learn spontaneously (through play and exploration)
- To some extent they are like young scientists discovering how things work (an old idea: Piaget, Kuhn, etc.)
- Their understanding of causality involves learning conditional probabilities as expressed in Bayesian nets (a topic of much recent research in AI).

The first two points relate closely to the theory we have been developing:

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

The Altricial-Precocial Spectrum for Robots (IJCAI 2005)

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

Natural and artificial meta-configured altricial information-processing systems (IJUC 2007)

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0702>

Computational Cognitive Epigenetics (BBS 2007)

The third point leaves out the kind of causal understanding, which we think occurs in some altricial species:

Kantian (not only Humean) causation.

HYPOTHESIS

- In nature, fluid, flexible, metaconfigured, cognitive development (using particular sorts of architectures, mechanisms and forms of representation), is generally found only in species that biologists call 'altricial' – i.e. born/hatched under-developed and cognitively incompetent

However, (a) the converse does not follow, and (b) the link is contingent:

Are elephants exceptions?

- This underdevelopment and incompetence at birth may not be necessarily a feature of metaconfigured artificial systems with flexible cognitive development – perhaps some machines, or animals on some other planet, can be 'born' fully formed and fairly competent as well as possessing the competence to learn qualitatively new things by other means than slow statistics gathering.
- Nevertheless there may be design features that are required by both artificial and natural rapid and flexible learners, capable of spontaneously developing new ontologies and new combinations of old competences – e.g. if brain development has to be staggered or 'cascaded', then at birth infants are more likely to be incompetent.
Compare cascade correlation neural network systems.
- We need to understand the design principles if we wish to develop machines capable of human-like understanding of the environment and rapid, flexible cognitive development.
- There can be different competences in the same animal or robot – some more rigid (precocial, genetically determined) some more flexible (derived creatively from exploration and play).
- We need to understand relations between environmental and task constraints that favour different combinations of pre-configured and metaconfigured development.

Summary so far:

There is an important sub-class of animals in which competences are not all pre-configured, whose development makes use of:

- Genetically determined actions, perceptual capabilities and representations,
- Genetically determined play/exploration mechanisms driving learning which **extends** those actions, etc., using abilities to **chunk**, **recombine** and **store**
 - new more complex **action fragments**
 - new more complex **perceptual structures**,
 - new more complex **goals**,
- Creating new ontologies, theories, competences (cognitive and behavioural),
 - i.e. new more complex **thinking resources**,
- Thereby extending abilities to search in a space built on larger chunks:
solving ever more complex problems quickly.
 - (unlike most statistical forms of learning)
- **Humans are able to apply this mechanism to itself – producing new forms of self-awareness and new forms of self-understanding, including mathematical knowledge.**

For AI systems this will require us to discover new architectures and learning mechanisms.

Two 'altricial' species and a pointer to a third

- Movie: Betty, the hook-making New Caledonian crow.

<http://news.bbc.co.uk/1/hi/sci/tech/2178920.stm>

or give to google three words: betty crow hook

- Movie: An infant (11.5 month) yogurt-manipulator experimenting with a bit of his world made up of spoon, hands, thighs, mouth, carpet, yogurt, tub — detecting interesting happenings and trying to understand and replicate/modify them.

<http://www.cs.bham.ac.uk/axs/fig/yog-small.mpg>

Like Betty he later tried to learn about hooks, but went through a stage of not understanding, shown here

http://www.jonathans.me.uk/josh/movies/josh34_0096.mpg

(We need many more videos of such infant exploratory play to study – in humans and other animals.)

- The key ideas are quite old – e.g. Piaget.
- Compare Oliver Selfridge's program that learns to 'count'

Reimplemented in Pop11 <http://www.cs.bham.ac.uk/research/poplog/teach/finger>

(describes Pop-11 program written over 20 years ago on the basis of an idea described to me by Oliver Selfridge.)

See:

O. G. Selfridge, "The Gardens of Learning", *AI Magazine* 14(2) (1993) 36*48

<http://aaai.org/Papers/Magazine/Vol14/14-02/AIMag14-02-005.pdf>

Partly like Case-based or Explanation-based learning.

Selfridge's metaconfigured Finger/Count program

RUNNING THE POP11 VERSION.

Initial state

Counter: 20

```
          v
    [] [] [] [] [] [] []
    0 1  2 3  4 5  6 7
```

There is a 'finger' adjacent to a row of blocks. It has two actions

- goright
- goleft

and a 'counter' that has two actions

- increment
- decrement

Actions can be composed in various ways

- in sequences
- loops

Loops terminate when either the finger or the counter hits a 'boundary'.

Example snapshots of the program working

Initial state

Counter: 20

```
          v
    [] [] [] [] [] [] []
   0 1  2  3  4  5  6  7
```

The program asks for a goal state and I type in '1 1'.

Target finger position and target counter value? 1 1

It then searches for a combination of moves that will produce a state with both counter and finger registering 1 — but it fails.

```
.....
.....
I give up on this one
```

Each dot represents a tested combination of actions.

It gives up after trying 120 different systematically varied actions.

After carefully selected training examples

Because successful chunks are stored as new action units, the set of available 'basic actions' increases:

- (1) [goright increment]
- (2) [[repeat goleft] [repeat decrement]]
- (3) goleft
- (4) goright
- (5) increment
- (6) decrement

After a few example tasks it gets to this collection:

- (1) [[repeat goleft] [repeat decrement] [repeat [goright increment]]]
- (2) [goright increment]
- (3) [[repeat goleft] [repeat decrement]]
- (4) goleft
- (5) goright
- (6) increment
- (7) decrement

Now it can always solve the 'counting' problem

No matter what the starting configuration, if given the 'count blocks' goal
(same target number for finger and counter),

it always solves the problem using only **one** stored action.

E.g. I give it the goal **17 17** in this configuration

Counter: 20

```

                                     v
    [] [] [] [] [] [] [] [] [] [] [] [] [] [] [] [] []
    0 1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17
```

Target finger position and target counter value? 17 17

•
Plan was: [[repeat goleft] [repeat decrement] [repeat [goright increment]]]

A single complex action reliably solves the problem which previously was found too difficult.

The program was not (like precocial animals) pre-configured with the ability to solve this class of problems. But it was metaconfigured with the ability to configure itself to solve such problems, given a carefully selected training sequence ('scaffolding' by the teacher).

Conjecture

We conjecture that rapid learning in altricial species depends on similar mechanisms, where the metaconfigured learner spontaneously attempts things without requiring a teacher.

This depends crucially on discretisation (chunking) of continuous domains, to provide ontological and representational units that are capable of being combined in ever more complex discrete structures.

Learn the easy things first, and some hard things become easy.

It is nearly impossible to learn anything that is hard to learn.

Oliver Selfridge: AI Magazine

The Gardens of Learning: A Vision for AI

14(2): Summer 1993, 36-48

However these ideas need to be expanded to include mechanisms that support **substantive** (i.e. non-definitional) ontology-extension.

I.e. abduction with introduction of new undefined symbols.

How to control the search for such extensions is a major problem: partly solved by evolution's meta-configured competences.

If all this is correct it also undermines **symbol-grounding theory** and the **sensory-motor contingencies** theory of cognition.

Limitations of the 'Finger' program

The program is obviously very limited

- very simple actions
- very simple kinds of perception
- no conditionals
- no parameters
- only very simple loop terminations
- very restricted kinds of goals
- it is essentially passive: goals must come from outside
- very simple 'environment' – e.g. no 3-D rotatable structures
- very restricted ways of composing actions
- no parallelism
- very few actions: and no need for action-selection mechanisms

All of these limitations could be removed in more complex programs.

Some requirements for extending the theory

Expanding this sort of mechanism to account for ‘altricial’ (flexible, creative, constructive, metaconfigured path-makers) will not be easy.

- It requires a host of specialised (probably genetically determined) mechanisms including mechanisms generating playful exploratory behaviour
- It needs recursive (?) syntactic competence and meta-semantic competence.
Meta-semantic competence is needed in mechanisms that can represent systems which themselves have representational capabilities – in the same agent or in others
- Some of the required elements seem to exist in AI developments of the last 40 years (many of them forgotten and not taught to students alas).
E.g. Sussman’s HACKER program (MIT, circa 1973?), and various kinds of symbolic learning mechanisms, including concept learning, rule learning, mechanisms (e.g. Explanation-based learning), as well as the more statistical mechanisms that now get most attention.
- The bootstrapping process needs
precocial (pre-configured) meta-level capabilities
- *Perhaps the hardest problem:*
we need to find spatial forms of representation in which complex structures and processes can be embedded and used for reasoning, planning, and learning.

Implications for theories of meaning

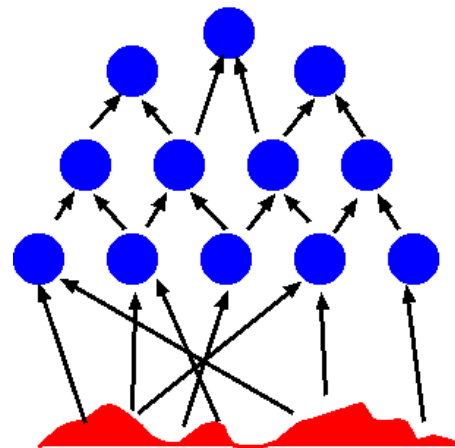
The existence of precocial species refutes 'symbol-grounding' theory

(Otherwise known as 'concept empiricism' – the theory that all meaning has to be derived by processes of abstraction from sensory experiences, which is clearly not required for precocial species that are competent at birth).

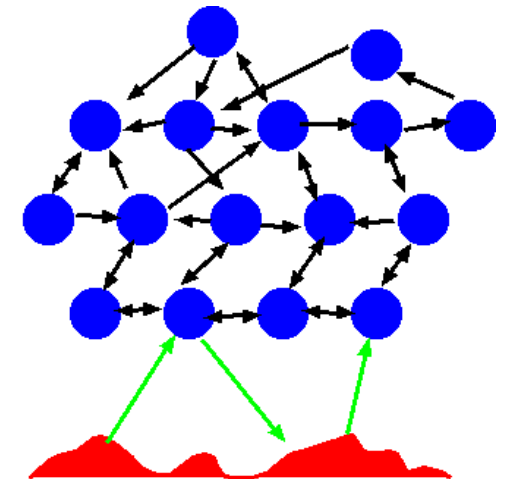
In our 2005 IJCAI paper we distinguish two sources of meaning

- the structure of a theory in which 'undefined terms' occur
(where the structure limits the class of possible models/interpretations)
- links to sensing and acting
(where the links – e.g. Carnapian 'meaning postulates' further reduce the set of possible interpretations, tethering the interpretation – though there is always residual indeterminacy.)

The second picture seems to represent how scientific theories get their meaning, so why not toddler theories?



Symbol Grounding



Symbol Tethering

How do you 'tell by looking'?

The examples of understanding (Kantian) deterministic causation in gears, links, shampoo containers, etc. presupposed that we sometimes can understand propagation of changes through changing structural relationships.

How it is done is far from clear, and it is far from clear how to implement such things in artificial systems.

The answer may be closely related to a theory of visual perception, according to which seeing involves running a collection of simulations at different levels of abstraction, partly, but not entirely, driven by the visual data.

Summary available here:

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0505>

- The simulation that you do makes use of not just perceived shape, but also **unperceived constraints**: rigidity and impenetrability.
- These need to be part of the perceiver's ontology and integrated into the simulations, for the simulation to be deterministic.

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*. His idea may have been based on something like this simulation theory, as follows:

It is obvious that the equivalence below is preserved if you spatially rearrange the blobs within their groups:

$$\begin{array}{r} 000 \\ 000 \\ 0 \end{array} + \begin{array}{r} 0 \\ 0 \\ 000 \end{array} = \begin{array}{r} 0000 \\ 0000 \\ 0000 \end{array}$$

Or is it?

How can it be obvious?

Can you see such a general fact?

How?

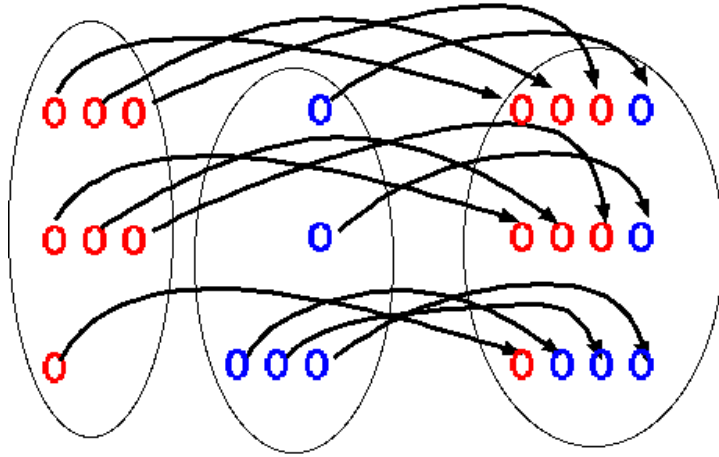
What sort of equivalence are we talking about?

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.

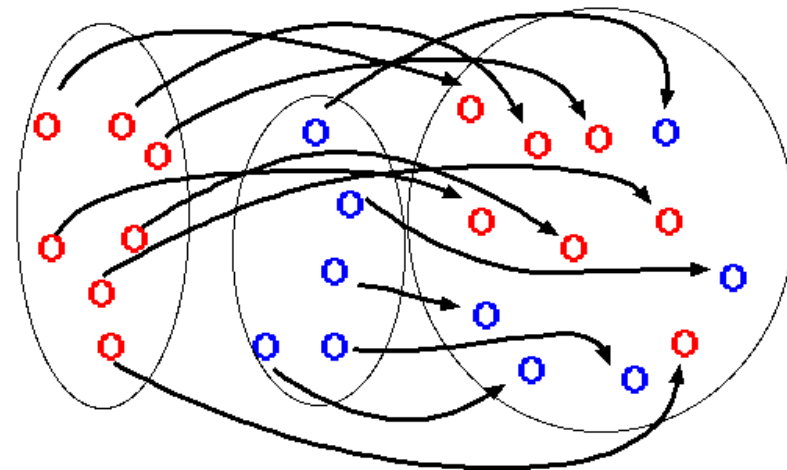
SEEING that $7 + 5 = 12$



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.

But that's a topic for another occasion.

Different kinds of learning

- We have made it sound as if some kinds of learning, such as learning about structure-based causation, or about mathematics, happen only in one way,
- However there are many things that are learnt by thinking explicitly, using something like the mechanism we have been describing (and probably others), after which that competence is used repeatedly in such a way that another part of the system, a 'reactive' layer gets trained to do the same task by going automatically from task or problem to solution, using a stored association, instead of working out the required behaviour.
- This can allow tasks using highly trained subsystems to run in parallel, while the deliberative structure-manipulating creative learning subsystem does something else.
- There are many examples, some physical (e.g. learning to play musical scales at high speed or learning to ride a bicycle or drive a car), and some mental, such as finding out numerical facts and then memorising them so that they are instantly available.
- Much learning of language seems to have the two strands: structure based explicit and relatively slow on the one hand and fast and fluent on the other.

The latter fools some researchers into thinking it's all statistical.
- Thus we should never ask 'How do humans do X?', for there may be many different ways humans do X (walk, talk, sing, plan, see, think, learn, reason, predict, explain).

Conclusion

- We have been emphasising the growth of understanding of the environment as based on a Kantian notion of causation – but only for some altricial species.
- This accounts for many of the most distinctive features of human life – and many causes of death also, when we act on incomplete or erroneous theories.
- However we are not claiming that **all** or even **most** of our information about causation is based on explanatory knowledge about the underlying structures.
- In particular, most of what a child learns about itself is Humean, including how to control its movements, then later much of how its mind works.
- Much self-knowledge, about body and mind, is incomplete, and liable to error.
- Alongside growth of insight into how physical things work a child also gradually bootstraps theories about how minds work, its own and others – child science includes psychology as well as mechanics and physics.

Both can produce errors (including religion and superstition) that persist in adult life. The errors will depend on how good the genetically determined and subsequently developed learning mechanisms are – and how far the understanding and teaching of science and engineering have progressed in the culture.

‘Know thyself’ Socrates is reputed to have said.

But understanding what is probably the most complex machine on earth, including many coexisting, interacting virtual machines within it, is easier said than done.

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