

How an animal or robot with 3-D manipulation skills experiences the world

Aaron Sloman, Jackie Chappell
and the CoSy team.
University of Birmingham, United Kingdom

**‘The substratum of this experience is the mastery of a
technique’ (Wittgenstein)**

We look at how the ability to experience grows as an architecture grows itself along with growing the ontology used to experience, understand and act in the environment.

The design-based approach to visual experience

Many people assume that the nature of visual experience can be adequately characterised by combining results of introspection and laboratory experiments in which people report what they see, either verbally or through actions they perform.

Some philosophers combine the above approach with conceptual analysis.

Some neuroscientists attempt to add further detail by investigating which neural mechanisms are active when various visual experiences occur.

These empirical and conceptual approaches on their own are inadequate: in order to find deep explanations we need to identify what exactly needs to be explained (e.g. which processes), and we need forms of explanation describing **mechanisms capable of explaining the processes that occur**. Some of the processes and mechanisms are not physical physiological but occur in abstract virtual machines.

We should also consider:

- How the phenomena to be explained **vary across species**,
- How the phenomena to be explained within a species **vary across individuals**,
- How the phenomena to be explained within an individual **vary during development (or after damage and disease)**.
- What can occur in future machines, and how to design them.

Many philosophers and psychologists, focus only on (normal) adult humans – a tiny subset.

NB. Young humans and many other animals see and manipulate things they cannot talk about: so the abilities do not require human language – though language may enhance them.

Example: watch toddlers and children, and ask: how could we design something that does that?



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

Video available at <http://www.cs.bham.ac.uk/~axs/fig/yog.mpg>

J discovered he could transfer yogurt to his leg, and picture 1 shows him trying to transfer more. His ontology seems not yet to include the orientation of the bowl. Picture 2 shows J trying to place a piece of yogurt picked up from the carpet into the spoon, prior to transferring it into his mouth. Picture 3 shows him trying, and failing, to put another piece of yogurt on the carpet, still apparently not experiencing the orientation of the bowl. Later J manages to transfer his grasp of the spoon handle from one hand to another. What mechanisms would allow a robot to learn like this?

J seems to experiment with his hands, legs, spoon, yogurt and the carpet. He **sees opportunities** and **tries them out**, **notices things** and **tries to recreate them** (often unsuccessfully). His ontology is quite rich but some gaps are evident.

He probably doesn't know he is doing all this! That would require a sophisticated self-monitoring architecture that is probably still being constructed. A baby is not just a tiny adult!

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 ???

What mechanisms explain this?

What exactly changes as a result?

How does the child' ontology grow?

What are the nature-nurture trade-offs?

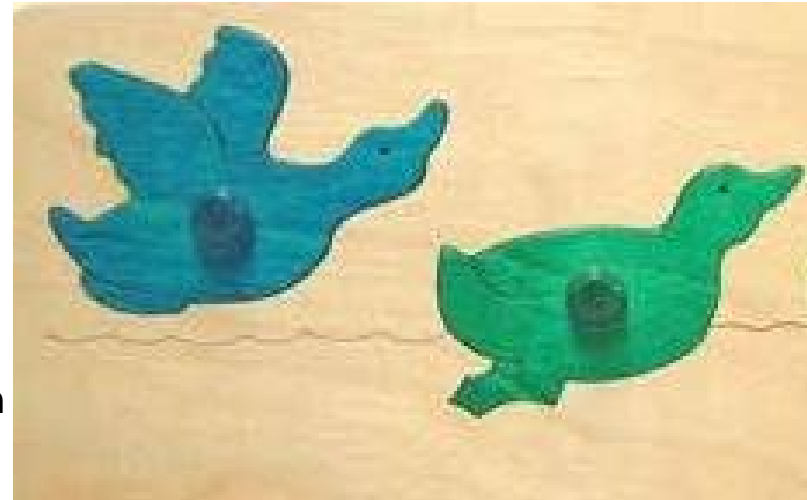
No current AI system comes close.

AI vision systems are restricted to much simpler tasks,
e.g. recognition, tracking, prediction,
point-wise reconstruction of surfaces, etc..

We cannot do it all from birth

Infants may not see causal relations adults experience as obvious

A child C learns that she can lift a piece out of its recess, and generates a goal to put it back, either because C sees the task being done by others or because of an implicit assumption of reversibility. At first, even when C has learnt which piece belongs in which recess there is no perception of unaligned boundaries, so there is only futile pressing. Later C 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** might allow toddlers to succeed: but there is no evidence that they can memorise precise orientation and location of an irregular shape. Can you?

Stacking cups simplify the cognitive task, partly through use of symmetry, partly through sloping sides — so they are much easier.

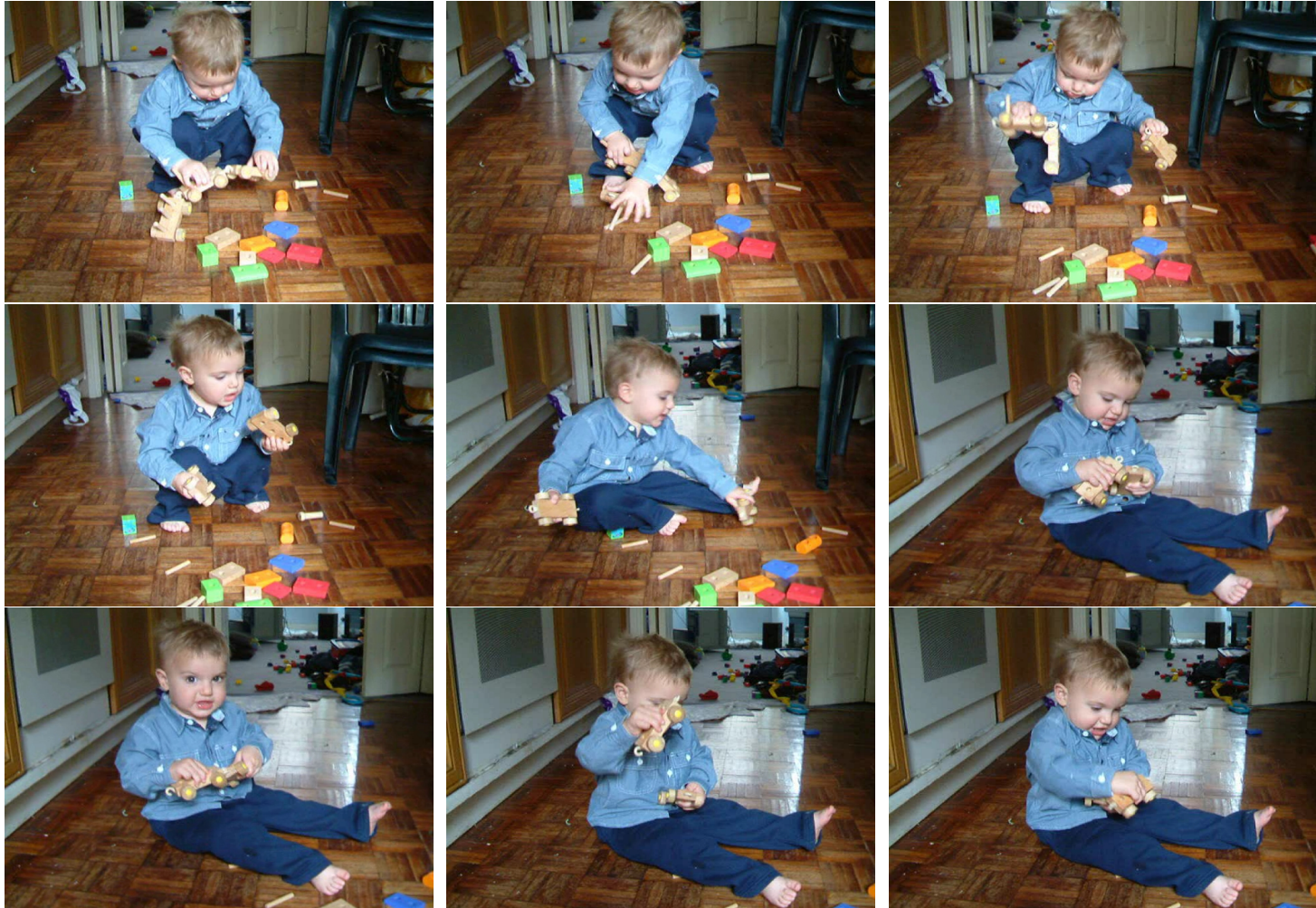
Eventually C's (still pre-linguistic) ontology includes something like 'boundary' and 'alignment'. Only then can she learn that if the boundaries are not aligned the puzzle piece **cannot** be inserted — probably some time after learning how to cope with symmetric stacking cups.

Conjecture: many changes in perception and action competence require the child to extend its ontology for representing objects, states and processes in the environment. The enriched ontology is used by the child's pre-linguistic perception and problem-solving mechanisms. HOW?



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.)

Last revised: August 7, 2006

A partially ordered network of stages

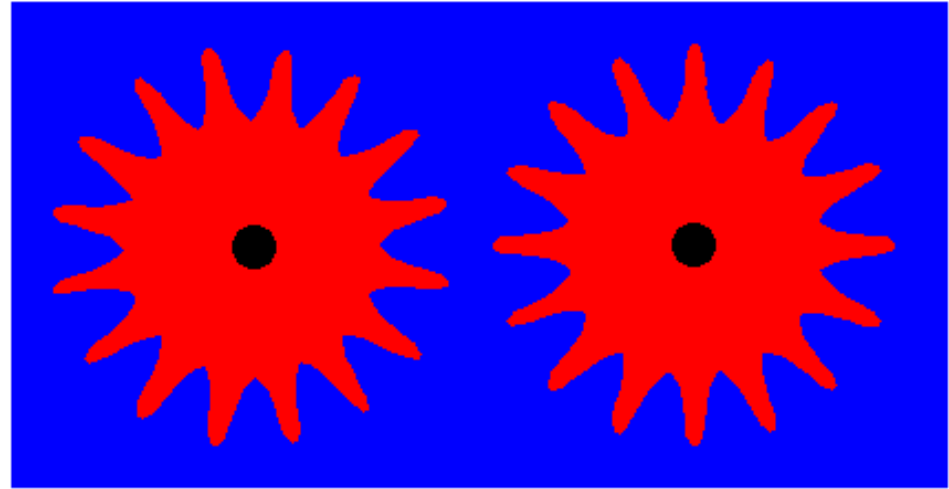
- **The process of extending competence is not continuous (like growing taller):**
- **The child has to learn about**
 - distinct new kinds of objects, properties, relations, process structures, e.g. for rigid objects, flexible objects, stretchable objects, liquids, sand, treacle, plasticine, pieces of string, sheets of paper, construction kit components in Lego, Meccano, Tinkertoy, electronic kits...
 - new forms of representation, new kinds of transformations, new constraints on transformations, new applications of previously acquired information.
- **There are not fixed stages: there is no 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.**
(Compare: Waddington's epigenetic landscape)
- **I don't know how many different things of this sort have to be learnt, but it is easy to come up with hundreds of significantly different examples.**
- **What can be learnt keeps 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. (Not for all species!)**
CONJECTURE: in the first five years, using a genetically determined bootstrapping mechanism, a child learns to run at least least hundreds, possibly thousands, of different sorts of simulations, using different ontologies – with different materials, objects, properties, relationships, constraints.
This involves acquiring two kinds of causal knowledge: Humean and Kantian.

Kinds of Causation: 1 (Humean)

Two gear wheels attached to a box with hidden contents.

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

Not really!



- You can tell by experimenting: you may or may not discover a correlation – 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.
- Humean causation involves **unintelligible sets of conditional probabilities** (e.g. Bayesian nets?). Knowledge of Humean causation amounts to having a convenient summary of raw data.

Kinds of Causation: 2 (Kantian)

Different gear wheels (still pivoted at their centres):

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

How? Assuming rigidity and impenetrability, you can **simulate** rotations and 'inspect' the consequences.

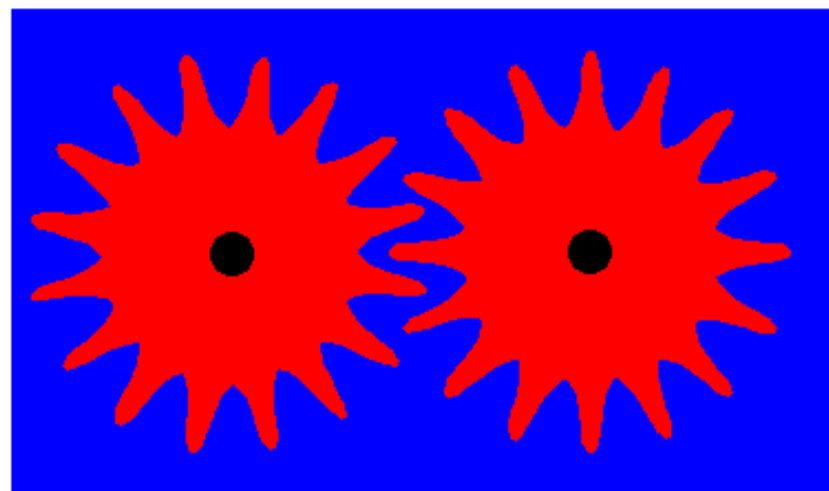
What you can see includes this:

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.

(I am not claiming that children need to reason **verbally** like this: consciously or unconsciously. It's a deeper, pre-linguistic ability at first. Probably some other animals grasp Kantian causality, e.g. nest-building birds?)

We seem to be able to run simulations using not only perceived shape, but also **unperceived constraints**: in this case rigidity and impenetrability.

Such constraints must be part of the perceiver's ontology and used in the simulations, for the simulation to be deterministic. How do 'rigidity' & 'impenetrability' enter a toddler's, or chimp's, or crow's ontology? The constraints and the processes using the constraints need not be conscious, or expressed in a logical, or verbal, form.



The ability to do causal reasoning in different domains (e.g. use of hooks) has to be learnt.

The ability to **work out consequences** in novel contexts 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, or simulating the interaction between hook and ring in pulling.

Appropriate constraints ensure the right counterfactual conditionals (and therefore causal relations) 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 I can give at present.

Part of the point of the CoSy 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.

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>

(Compare R.Grush BBS, 2004. He uses the label 'emulation'.)

Pushing and pulling in different contexts

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 whole 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 below and 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.

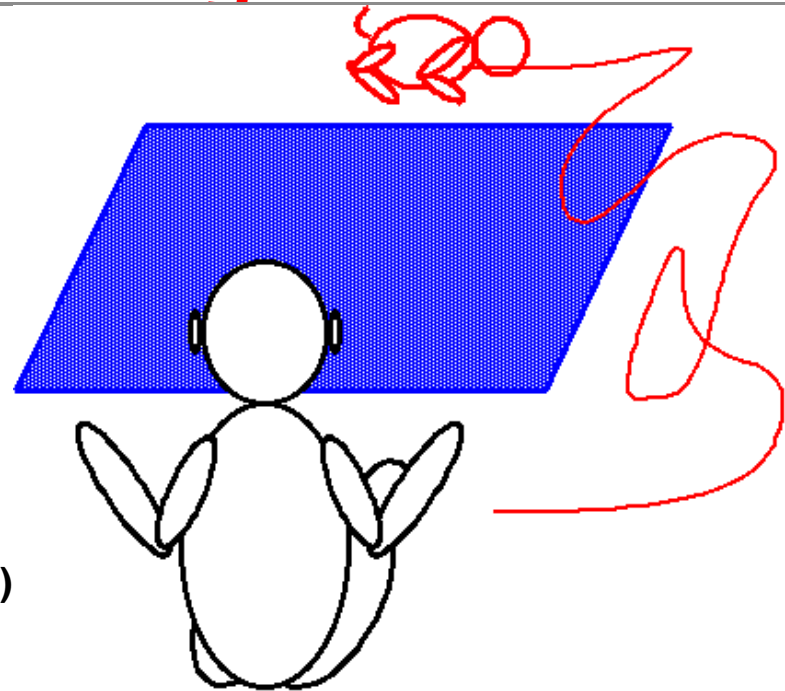
Later the child may **either** have extended the ontology used in its conditional probabilities, **or** have learnt to simulate the process of moving X when X supports Y. (Kantian causation.)

As a result of either, he doesn't try pulling the blanket to get the toy lying just beyond it, but uses the string.

If he has a Kantian understanding he can solve a wider range of problems.

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

Learning how to get hold of a toy that is out of reach.

Some things to learn by exploration

Short blanket

Grab edge and pull

Long blanket

Repeatedly scrunch and pull

Towel

Like blanket

Sheet of plywood

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

Sheet of paper

Roll up? (But not thin tissue paper!)

Slab of concrete

Crawl over or round

Taut string

Pull

String with slack

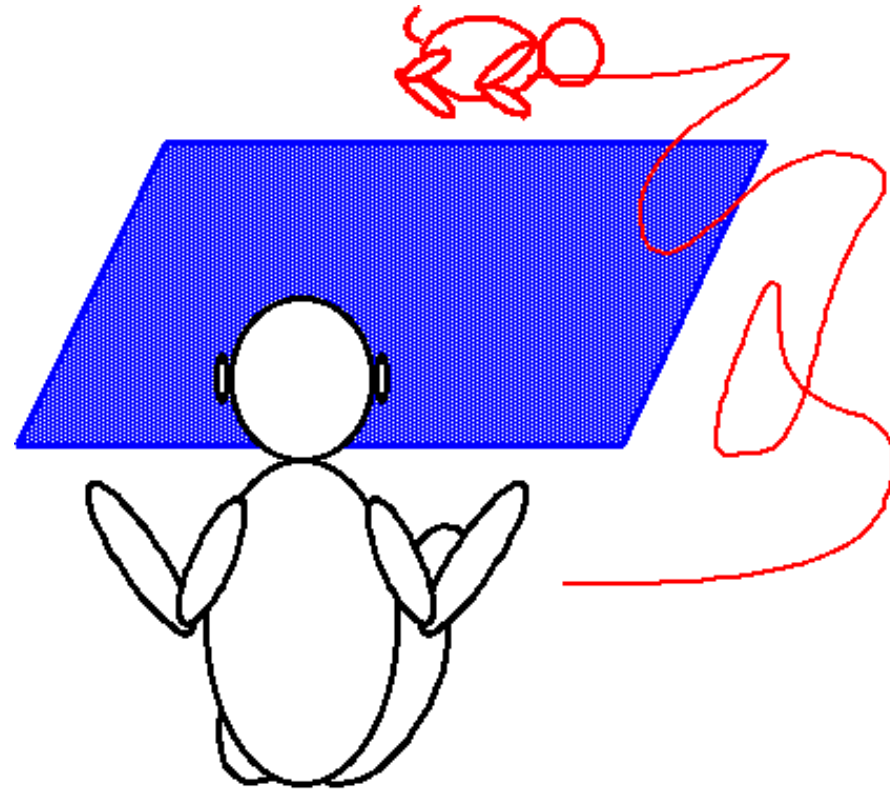
Pull repeatedly

String round chair-leg

Depends

Elastic string

?????



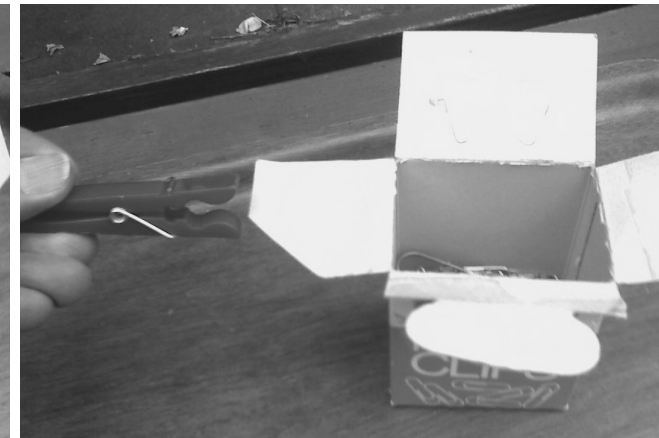
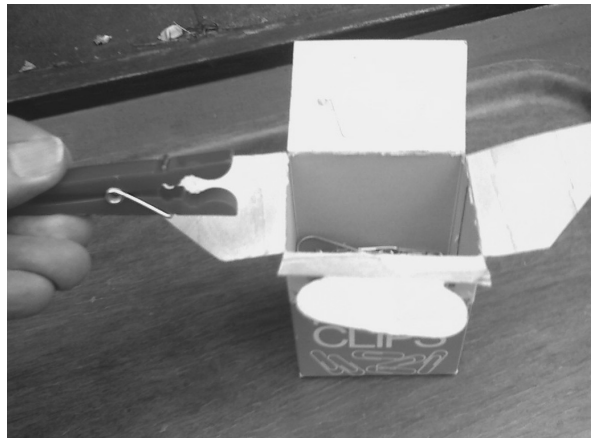
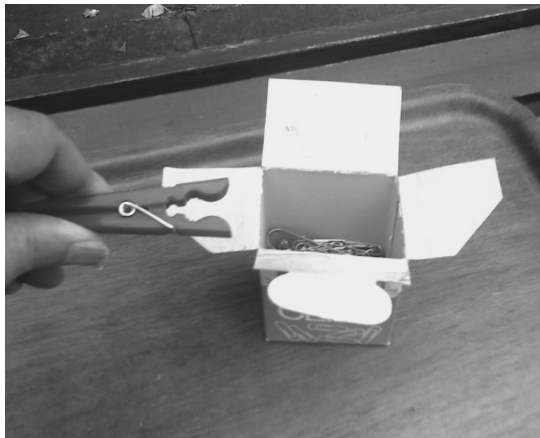
See this discussion of learning orthogonal recombinable competences
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0601>

It takes a lot of learning to acquire mastery of the 'techniques' required for seeing and understanding these affordances.

Manipulative competences

What information do you need to acquire (consciously or unconsciously) and use, in getting the banana into the glass?

What possibilities do you need to see? How many different concurrently changing relationships are involved? Metrical, topological, qualitative,



Getting a clothes peg to clip onto a particular flap on a box, involves **several simultaneous changes in relationships** (position, orientation, orientation of the gap, aligning the gap with the flap's edge.)

Multi-strand relations change throughout the action. Some changes are produced by motors, others are perceived consequences. Many changes are **structural**, not just changing values in vectors.

Managing a flexible hand



How many points of contact do you need to lift a cup without it tipping over?
How many distinct processes are involved in producing those relationships?

NB: saying things need to be seen does not mean that they are all processed in the same part of the total system architecture or that we know we are seeing them.

- Some bits of the architecture **do visual servoing** using continuously changing visual information: some are 3-D changes (e.g. grasping), some 2-D image based (narrowing a gap).
- Some mechanisms **process more abstract qualitative relationships**, e.g. whether a whole object is moving roughly in the right direction, so as to avoid a collision, or rotating in the right way.
- Some **process topological changes**, such as one thing coming into contact with something else, or entering a region, e.g. the convex hull of the glass, or penetrating the empty disk-shaped volume inside a closed circular handle.
- There are different sorts of information and different functions for which information is used.
(**Forget crude 'what vs where' and 'what vs how' distinctions!! Think about what has to work!**)

What needs to be explained is far from obvious

We can identify some of what needs to be explained by observing many things and adopting the design-stance to produce **detailed requirements** for non-vacuous explanations.

It can be very hard to do!

A good explanation of something that occurs should be usable in producing a detailed simulation, or better, a replication, of a variety of different but connected phenomena – otherwise it may just be a re-description, or worse.

(The difference between replication and simulation may sometimes be a matter of dispute.)

Typically the process of working on a design brings up features of the performance that were not noticed previously – **giving new insights into requirements.**

When I started (in the EC-funded CoSy project) working in detail on requirements for a (partly) human-like robot able to manipulate 3-D objects, I noticed for the first time what should have been obvious to me long before, namely

- that structured objects have **'multi-strand' relationships** not expressible simply in the familiar form $R(X, Y)$, because the relation between X and Y usually involves **many relations between parts of X and parts of Y .**
- the objects need not be hierarchically decomposable in a unique way into parts: rather there are many different decompositions – sometimes used simultaneously.
- while some parts may be separable objects, others may be surface fragments, e.g. part of the rim of a cup, a groove or a dent. (Relationships usually form graphs, not trees.)
- These multi-strand relations need not have familiar names in ordinary language, and understanding some of them may occur in pre-linguistic children and non-linguistic animals.

Perceiving diverse multi-strand processes

Motion of structured objects with multi-strand relations that change during the motion involves 'multi-strand' (concurrent) processes.

- That is, many relationships change in parallel – e.g. faces, edges, corners of one block may all be simultaneously changing their relationships to faces edges and corners of another.
- Moreover, not only do changes occur in parallel, but some cause others (remember the gears).
- Things get more complex when objects are flexible, e.g. seeing a hand peeling a banana or a child putting a sweater on, or dough being kneaded.
- What are the design requirements for seeing such things and what mechanisms can satisfy those requirements?

Ontological and representational requirements for experiencing commonality.

What is common to performing the same sort of grasping action (bringing two surfaces together to clasp part of an object) using

- your left hand,
- your right hand,
- two hands grasping a large object,
- one hand and your chest,
- your teeth?

How much of what you need in order to see and do those things also applies to thinking about past or future occurrences, or similar actions performed by others?

Summary so far

Analysing requirements for human-like perception and production of 3-D manipulations shows that mechanisms are needed which produce concurrent ('multi-strand') processes representing concurrent structural and metrical changes at different levels of abstraction, in **partial** registration with the optic array.

(NOT in registration with the **retina**, since saccades, etc., occur frequently).

This involves being able to run several 'simulations' concurrently, for different (coexisting) purposes: low level control (visual servoing), planning, predicting, plan-checking, visualising consequences....

So vision can have a kind of complexity that appears not to have been noticed by many researchers because of too much focus on static objects and much simpler visual tasks like recognition, tracking and prediction.

Nothing in AI comes close to modelling all this (though there are fragments).

It seems likely that it will be hard to explain in terms of known neural mechanisms especially since the mapping between key sub-processes and the retina is constantly changing.

(Compare Arnold Trehub *The Cognitive Brain* (MIT Press, 1991) and Grush, BBS, 2004)

There are deep implications for human (and animal) development, understanding of causation, and computational modelling.

This expands Wittgenstein's observation (in *Philosophical Investigations*):

'The substratum of this experience is the mastery of a technique'

Not quite right: visual experiences use multiple 'techniques' concurrently!

Yet more are **available** if needed, even if not actually invoked. (Compare global workspace theory.)

Objective sensorimotor contingencies??

How are structures and processes represented? Not all as sensorimotor-patterns!

A child or robot that has to manipulate 3-D objects in its environment would face a combinatorial explosion if all possible situations have to be learnt about separately. This could take evolutionary time-scales – as happened for precocial species, like deer that run with the herd soon after birth, and chicks pecking for food.

Humans and some other altricial species seem to use innate mechanisms for decomposing situations into components that can be explicitly learnt about, and stored so that the competence can be re-used in combination with others, in perceiving novel situations, planning and performing novel actions.

That includes learning about kinds of surface fragments (e.g. varieties of curvature and surface discontinuities), kinds of surface properties (e.g. texture, hardness, etc.), kinds of material (rigid, flexible in different ways), kinds of objects composed of materials and shapes, kinds of relationships, kinds of changes in relationships, kinds of causal connections between changes.

These sometimes need to be represented in a manner that is independent of precise sensory and motor signals when they are perceived so that knowledge about them can be used in planning future actions, thinking about the past, and comparing actions using different hands, or hands or mouth in different positions.

This implies a use of ‘objective’ representations (e.g. of 3-D structure) which can then also be used in perceiving ‘vicarious’ affordances (for others). Is this what mirror neurons really do???

Somatic vs Exosomatic contingencies

Two evolutionary 'gestalt switches'?

The preceding discussion implies that during biological evolution there was a switch (perhaps more than once) from

insect-like understanding of the environment in terms of **sensory-motor contingencies** linking somatic patterns in **internal** motor signals and sensor states at various levels of abstraction (subject to prior conditions),

to

a more 'objective' (amodal, exosomatic, objective) understanding of the environment in terms of **action-consequence contingencies** linking **changes in the environment** to **consequences in the environment**,

followed by or accompanied by

a further development that allowed a **generative** representation of the principles underlying those contingencies, so that novel examples could be predicted and understood, instead of everything having to be based on statistical extrapolation.

A major driver for this development could be evolution of body parts that can manipulate objects and be seen to do so and move independently of eyes.

However the cognitive developments were not **inevitable** consequences: e.g. crabs that use their claws to manipulate food do not necessarily have the generative competence.

(Biologist Holk Cruse tells me the above is unfair to some insects. Also Portia spiders?)

NOTE: Terminology changed Aug 2006: intrasomatic/extrasomatic \Rightarrow somatic/exosomatic

Affordances

The notion of affordance introduced by J.J.Gibson is now widely referred to, but it turns out that not everyone understands it in the same way.

E.g. many people ignore the fact that affordances are not necessarily concerned with whole objects but can be associated with portions of the surface of an object, e.g. dents, grooves, cracks, edges, openings, part of a handle etc.

And few appear to have realised the importance of the ability to perceive ‘vicarious affordances’ – affordances for others, including conspecifics, prey, predators.

The kind of generalisation of form of representation that makes vicarious affordances possible, i.e. representation that is independent of the sensory and motor signals involved in perceiving and using the affordance, may also have been important for self-affordances in the context of

- Predicting them
- Planning actions involving them
- Reflecting on past actions involving them
- Generalising across different ways of using them – e.g. different ways of grasping a graspable object (with left hand, with right hand, with both hands, with mouth, held to chest, etc. etc.)

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

Conjecture: the evolution of grasping devices that move independently of eyes (i.e. hands instead of mouth or beak) had profound implications for the need to represent actions while abstracting away from the somatic sensorimotor details. The switch to an exosomatic (amodal) representation enabled enormous economy of resources, based on compact re-usable generalisations.

Misunderstood neurons

- An implication of all this is that so-called mirror neurons might more accurately have been called ‘abstraction neurons’: they are probably linked to mechanisms that have a much deeper and more general role than somehow facilitating the self-other transfer of information.
- Abstraction is escaping from the limitations of sensorimotor information so as to represent and use an (amodal) exosomatic ontology.
- Multi-modal learning can happen without this, e.g. learning correlation between certain motor signals to neck and changes in optical flow.
- Learning exosomatic contingencies made possible perception and use of ‘vicarious affordances’, the ability to think about future plans, past causes and remote or unobserved processes.
- There are many implications, for robotics, psychology and neuroscience.

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

In particular, abstracting from intrasomatic (even multimodal) phenomena to use an amodal exosomatic ontology referring to materials, shapes, processes, constraints in the physical environment allows deeper kinds of causal learning than mere conditional probabilities linking sensory and motor signals.

It allows a transition from Humean (probabilistic, correlation-based) causation to Kantian (deterministic, structure-based) causation.

Biological bootstrapping mechanisms

- There are some species whose needs cannot be served by genetically determined (preconfigured) competences based on 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)
 - with specific information about **types of**: 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 more precise, species with different mixtures of preconfigured and metaconfigured competences.

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, expanding the ontology used**. (Simple computer demo available.)

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 primitive actions, perceptual capabilities and representations,
- Genetically determined play/exploration mechanisms driving learning that **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**,
- Not restricted to intrasomatic sensorimotor ontologies.
- 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.

A spectrum of competences

- Every organism is a mixture of both kinds of capabilities:
pre-configured — constructed (meta-configured)
- Not all of the first kind are manifested at birth/hatching – many are ‘time-bombs’
(e.g. ticking away, waiting for sexual maturity, or for the season to hibernate, or migrate).
- Architectures for altricial species can do many things that are not directly biologically useful, but may provide reusable information:
including (possibly dangerous) exploration of a space of possibilities.
- Architectures can change over time.
- Ontologies used can change over time.
- Forms of representation used can change over time.
- **Altricial architectures are virtual machines that grow themselves.**
But we have over-simple ideas about how: e.g. the notion of a knowledge-free, general-purpose, learning system is current popular but inadequate mechanism.

See our (Sloman & Chappell) IJCAI paper

<http://www.cs.bham.ac.uk/research/cogaff/05.html#200502>

and the H-CogAff architecture described on the Cognition and Affect web site:

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

Implications for theories of meaning

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

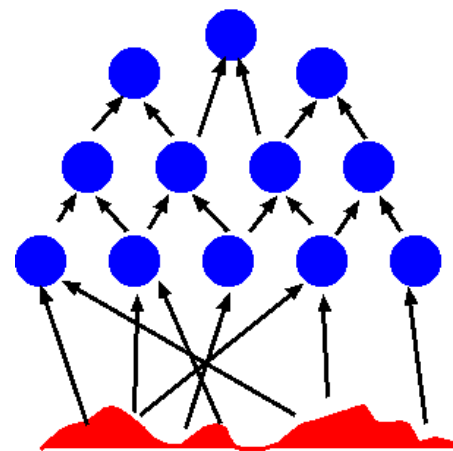
(One version of 'concept empiricism' – the theory that all meaning has to be derived by processes of abstraction from sensory experiences, which is clearly not required for precocial species that are competent at birth).

In our IJCAI paper we distinguish two sources of meaning

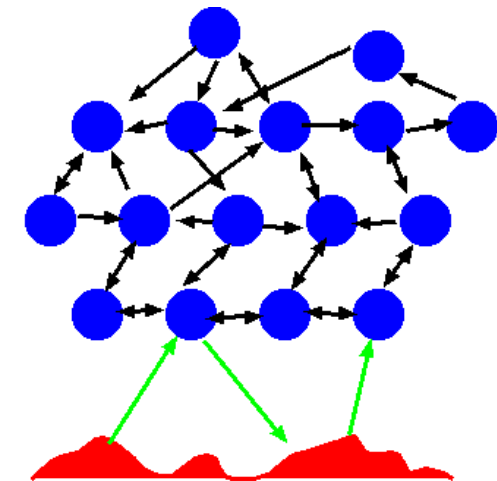
- the structure of a theory in which 'undefined terms' occur
(where the structure limits the class of possible models/interpretations)
- links to sensing and acting
(where the links – e.g. Carnapian 'meaning postulates' further reduce the set of possible interpretations, tethering the interpretation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

The second picture seems to represent how terms in scientific theories get their meaning, i.e. largely from the structure the theory. So why not concepts in toddler theories?

Compare 20th century philosophy of science after crude empiricism was shown to be wrong: Popper, Carnap, Hempel, Pap,



Symbol Grounding



Symbol Tethering

additional points

- We need to find out how many different kinds of simulative capabilities are typically acquired by: a child, a chimp, a nest-building bird, a useful domestic robot.
- We need to understand what sorts of forms of representation, mechanisms and architectures, can produce those developments.
- The process can involve creation of new ontologies and new forms of representation.
- Many different kinds of cognitive competence relevant to understanding different kinds of structures and processes grow during our life time.
- Different people grow different subsets (why?)
- Scientific research is just an extension of this – though too many scientists restrict their research to accumulation of correlations (like learning in precocial species?).
- When the ability we are discussing is applied to itself we get activities like mathematics and philosophy.

Avoid the noun 'consciousness': explain processes

Usage of the word 'consciousness' is full of confusion (with many conflicting views and definitions): so it is not a fit term to use in formulating scientific questions, e.g.

- how did it evolve?
- why did it evolve?
- which animals have it?
- which physical mechanisms can produce it and how?

If there is no well-defined concept (no unambiguous 'it') these are not questions for which there can be right or wrong answers. We need to **sharpen** the questions: different things evolved at different times.

See Aaron Sloman and Ron Chrisley, (2003) Virtual machines and consciousness, *JCS* 10,
<http://www.cs.bham.ac.uk/research/cogaff/03.html#200302>

In order to make progress, scientists should drop use of the muddle-inducing noun 'consciousness' and try to explain **everything else**, in great detail, preferably focusing on verbs, adjectives and adverbs to describe phenomena of many kinds, each of which has many variants with different explanations: e.g. perceiving, experiencing, learning, inferring, deciding, wanting, enjoying, disliking, noticing, recalling, forgetting, being surprised, caring about something, dreaming, etc. etc. (But don't focus **only** on things for which we already have names.)

We should also consider how the phenomena to be explained **vary across species**, and within a species **across individuals**, and within an individual **during development** – and their roles in future machines. Don't study only (undamaged) adult humans – a tiny population.

Conclusion

- We have been emphasising the growth of understanding of the environment as based on a deep Kantian notion of causation in an external world, extending shallow Humean causation (correlations) – but not for all 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** of our information about causation is based on explanatory knowledge about the underlying structures.
- 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 – unhampered by religious indoctrination and mind-binding.

- **Most of what a child learns about itself is only Humean, alas, including how to control its movements. Kantian self-understanding is rare, and difficult.**

‘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.

Further Reading

A longer presentation of some of these ideas can be found at:

<http://www.cs.bham.ac.uk/research/cogaff/talks/aisb06-ortho.pdf>

A web site exploring some of the issues in more detail is here:

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

Orthogonal Recombinable Competences Acquired by Altricial Species (HTML)

A (possibly) new theory of vision related to all this is presented in:

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

A discussion of ways of conceiving of and learning about causality:

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0506> (PDF)

Two views of child as scientist: Humean and Kantian

A paper co-authored by Jackie Chappell, presented at IJCAI'05:

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

The Altricial-Precocial Spectrum for Robots

A discussion of the distinction between intrasomatic and exosomatic interpretations of the notion of sensorimotor contingencies:

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

Sensorimotor vs objective contingencies

Partial analysis of the notion of a 'fully deliberative' system and the intermediate cases between proto-deliberative and fully deliberative systems:

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604>

More presentations, including stuff on virtual machines, architectures, emotions, etc.:

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