How to experience the world: some not so simple ways

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

Based on work done with Jackie Chappell concerning animal cognition, work with the CoSy team on requirements for human-like robots, and discussions with Dean Petters about babies.
See http://www.cs.bham.ac.uk/research/projects/cosy/

Closely related papers and talks:
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0601
On learning about orthogonal recombinable competences
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0505
A (possibly) new theory of vision
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0506
Two views of child as scientist: Humean and Kantian

Some questions: Let's have a vote!

- **Is a fish conscious?**

Some questions: Let's have a vote!

- Is a fish conscious?
- Is a fly conscious of the fly-swatter zooming down at it?

Some questions: Let's have a vote!

- Is a fish conscious?
- Is a fly conscious of the fly-swatter zooming down at it?
- Are you conscious when you are dreaming?

Some questions: Let's have a vote!

- Is a fish conscious?
- Is a fly conscious of the fly-swatter zooming down at it?
- Are you conscious when you are dreaming?
- Are you conscious when you are asleep?

Some questions: Let's have a vote!

- Is a fish conscious?
- Is a fly conscious of the fly-swatter zooming down at it?
- Are you conscious when you are dreaming?
- Are you conscious when you are asleep?
- Is the file protection system in an operating system conscious of attempts to violate access permissions?

Some questions: Let’s have a vote!

- Is a fish conscious?
- Is a fly conscious of the fly-swatter zooming down at it?
- Are you conscious when you are dreaming?
- Are you conscious when you are asleep?
- Is the file protection system in an operating system conscious of attempts to violate access permissions?
- Is a soccer-playing robot conscious?

Problems with ‘consciousness’ in science and philosopy

- We all know about a wide range of phenonena that we associate with the label ‘consciousness’
- Our beliefs about them are not all consistent
- As people develop scientific and philosophical theories intended to illuminate the phenomena they tend to ignore more and more of what they originally knew
- And end up looking closely only at the phenomena illuminated by their theory.
- Does that remind you of anything?

Thanks to Murray Shanahan for helping me formulate this summary.

My own papers using the word ‘consciousness’ all point out that there is no unitary phenomenon to be explained.

Elephants

Before asking what consciousness is, let's ask what an elephant is.

The Parable of the Blind Men and the Elephant
John Godfrey Saxe (1816-1887)

http://www.wvu.edu/~lawfac/jelkins/lp-2001/saxe.html

It was six men of Indostan
To learning much inclined,
Who went to see the Elephant
(Though all of them were blind),
That each by observation
Might satisfy his mind.

The First approached the Elephant
And, happening to fall
Against his broad and sturdy side,
At once began to bawl:
"God bless me, but the Elephant
Is very like a wall!"

The Second, feeling of the tusk,
Cried, "Ho! what have we here
So very round and smooth and sharp?
To me 'tis very clear
This wonder of an Elephant
Is very like a spear!"

The Third approached the animal
And, happening to take
The squirming trunk within his hands,
Thus boldly up he spake:
“I see,” quoth he, “The Elephant
Is very like a snake!”

The Fourth reached out an eager hand,
And felt about the knee:
“What most the wondrous beast is like
Is very plain,” quoth he;
“Tis clear enough the Elephant
Is very like a tree!”

The Fifth, who chanced to touch the ear,
Said, "Even the blindest man
Can tell what this resembles most;
Deny the fact who can:
This marvel of an elephant
Is very like a fan!"

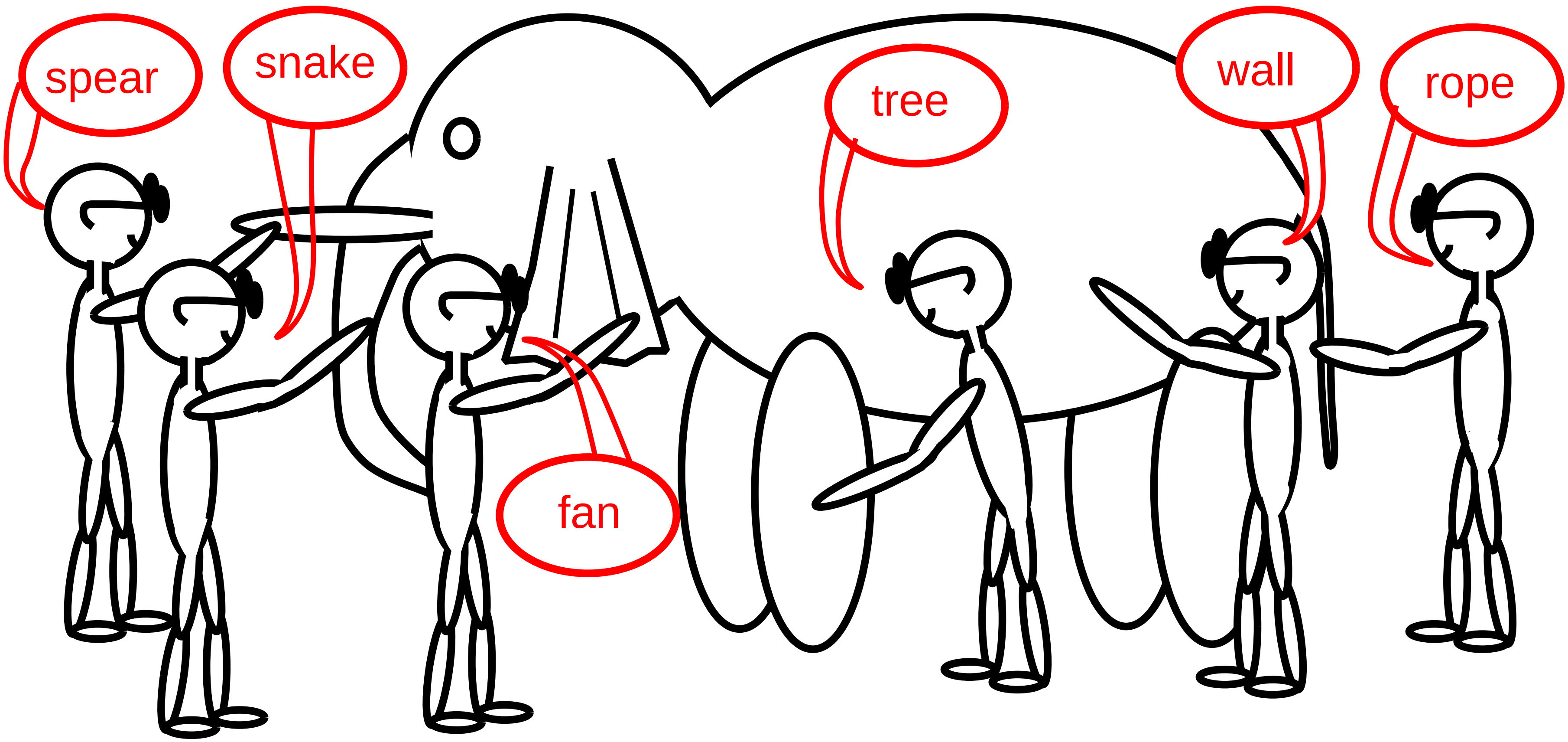
The Sixth no sooner had begun
About the beast to grope
Than, seizing on the swinging tail
That fell within his scope,
“I see,” quoth he, “the Elephant
Is very like a rope!”

And so these men of Indostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong.
Though each was partly in the right,
They all were in the wrong!

(Use Google to search for "blind men elephant")

What is an Elephant?

See: “The Parable of the Blind Men and the Elephant”
by John Godfrey Saxe (1816-1887)
http://www.wvu.edu/˜lawfac/jelkins/lp-2001/saxe.html



These blind men can claim in their defence that they are not systematically ignoring lots of things they have previously encountered because they don’t fit the theory.

MORAL

**Consciousness is
a huge elephant
studied by
many blind and forgetful men
(and women)**

**(Actually several elephants,
corresponding to different kinds
of candidates for the label 'consciousness'
in different sorts of animals and machines...)**

Instead of focusing on legs, or tusks, or tail, ... lets collect information about all major aspects of the whole beast: the whole information-processing system.

Then there will be no further question about what it really is.

Some people are starting to do this then spoiling things by saying they are discussing consciousness.

Blind men describing consciousness

Different consciousness theorists contract what they say about 'it' — and even contradict themselves sometimes. Maybe there's no unique 'it'?

- **It's indefinable, knowable only through having it**
- **It's what it is like to be something (hungry, in pain, happy, a bat...)**
  What on earth does that mean????
  (Compare http://www.cs.bham.ac.uk/~axs/misc/like_to_be_a_rock/)
- **You lose consciousness when you are asleep**
- **You are conscious when you dream**
- **Consciousness is essential for processes to be mental**
- **Many mental processes are inaccessible to consciousness**
- **It causes human decisions and actions**
- **It has no causal powers (it is epiphenomenal)**
- **It can exist independently of physical matter (e.g. in an after-life)**
- **It's a special kind of stuff somehow produced by physical stuff**
- **It's just a collection of behavioural dispositions**
- **It's just a collection of brain states and processes**

...continued

- It's an aspect of a neutral reality which has both physical and mental aspects
- It's just a myth invented by philosophers: best ignored
- It's got something to do with talking to yourself (Dennett?)
- It's something you either have or don't have
- It's just a matter of degree (of something or other)
- Consciousness requires a public (human) language
- Animals without language can have it
- All animals have it to some degree
- Humans are the only animals that have it
- It's located in specific regions or processes in brains
- Talk about a location for consciousness is a "category mistake"

...continued

- Specific conscious events must have specific neural correlates
- Specific mental events are all multiply realisable, and therefore need not have fixed neural correlates.
- No machine could have it
- A machine that was indistinguishable from humans would have it
- Zombies are possible: machines that are indistinguishable from us could lack consciousness
- A machine that had exactly the same internal information processing capabilities as humans would necessarily have it.

.... and so on and so on

Suggestion:

The only thing common to everything investigated in consciousnes studies is
acquisition and use of information about something

— and that's something even microbes do.

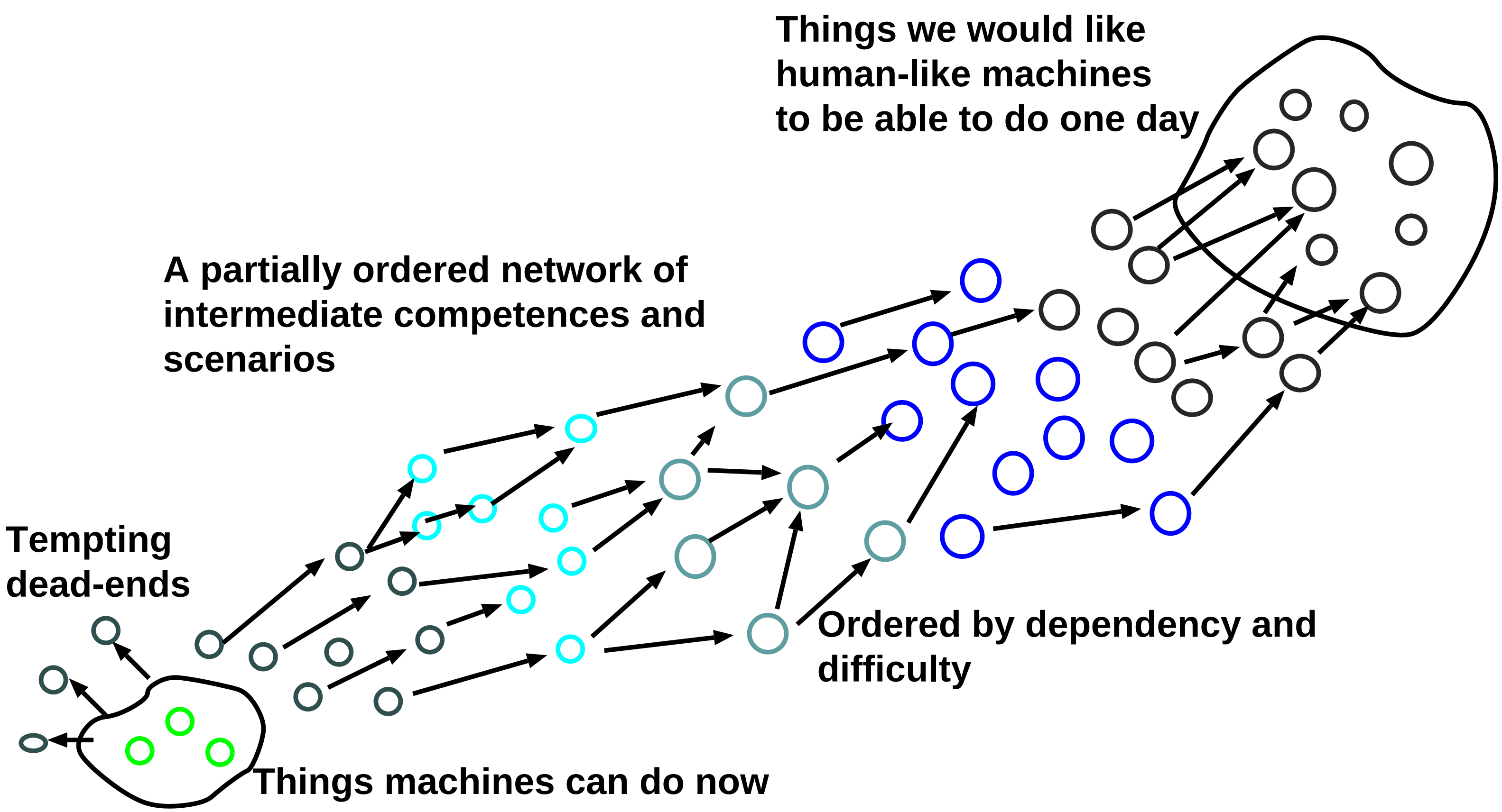
Summary

- **The best way to extend our scientific understanding of consciousness is to stop using the noun and investigate all the many mental processes that can and do occur in humans and other animals and future robots in very great detail and explain how they are possible.**
- **Then everything of substance about consciousness will have been covered, and the vacuous, incoherent unanswered questions generated in philosophical discussions will remain unanswered as they should be, because they are unanswerable.**
- **This point has to be understood in the context of a long term research programme, that I shall try to characterise.**
- **If there's anything new here it is that both the body and things in the world have complex changeable structure involving multiple concurrent changes (subject to many different sorts of constraints) requiring simulation (old idea) of many concurrent processes at different levels of abstraction.**
- **Evolution discovered the importance (for a subset of species) of depending not on sensorimotor contingencies but on condition/consequence contingencies – i.e. to some extent disregarding embodiment.**
- **Mirror neurons should be called abstraction/objective neurons.**

To make this precise requires many data-points

We need to get very clear about requirements if we start producing designs

Steps towards a research roadmap



Forward chaining research asks: how can we improve what we have already done?
Backward chaining research asks: what is needed to achieve our long term goals?

See the introduction to GC5 in the booklet and on the web: researchers don't put nearly enough effort into analysing requirements based on backward chaining from detailed analysis of distant scenarios.

Methods and tools to help build roadmaps

Many people find it very difficult to think up a systematic and comprehensive collection of future scenarios of the kind required. We have been working on a methodology to help with development of this network of roadmaps.

A three-dimensional grid of competences. One dimension represented by columns in the figure is concerned with types of *entity* to which competences can be applied (e.g. 2-D and 3-D spatial locations, regions, routes, inert objects, mobile objects, objects with goals, perception, and action, and more abstract entities such as beliefs, proofs, numbers, plans, concepts). Another dimension (rows) is concerned with types of *competence* that can be applied to instances of some or all of the types of entities; for instance competences like perceiving, manipulating, referring to in thought, referring to in language, constructing, ...

Entity-types / Competences	E1	E2	E3	E4	E5	E6	E7	E8
C1								
C2								
C3								
C4								
C5								
C6								
C7								

The third dimension could be thought of as the depth of the boxes – difficulty of the competence (implying time required to produce working systems).

Actually a more complex topology than a rectangular grid is required: refinements and elaborations of the grid are topics for future research. (For more detail see the introduction to GC5 symposium in proceedings or website http://www.cs.bham.ac.uk/research/cogaff/gc/aisb06/).

The space of sets of requirements: ‘niche space’ for biological and non-biological machines

- Types of entity
- Types of competence
- Types of combined competence-type+entity-type
- Some entities are abstract and internal to the individual
- Some are in other individuals or in the environment (e.g. causal relations)
- Types of complex competence based on combinations of simpler competences.

  E.g. seeing or imagining or describing a hippo swallowing a fly.

The recent concerns about embodiment, sensorimotor contingencies, symbol grounding, dynamical systems all arise from a consideration of only a subset of the requirements for a human- (or chimp- or crow-) like information processing machine, namely the subset shared with microbes, insects, fishes, reptiles, etc.

The space of designs

We need a meta-theory of types of

- **mechanisms**
- **forms of representation**
- **types of sub-functions**
- **architectures in which the above can be combined**

Evolution produced a wide variety of which we still understand only a tiny subset.

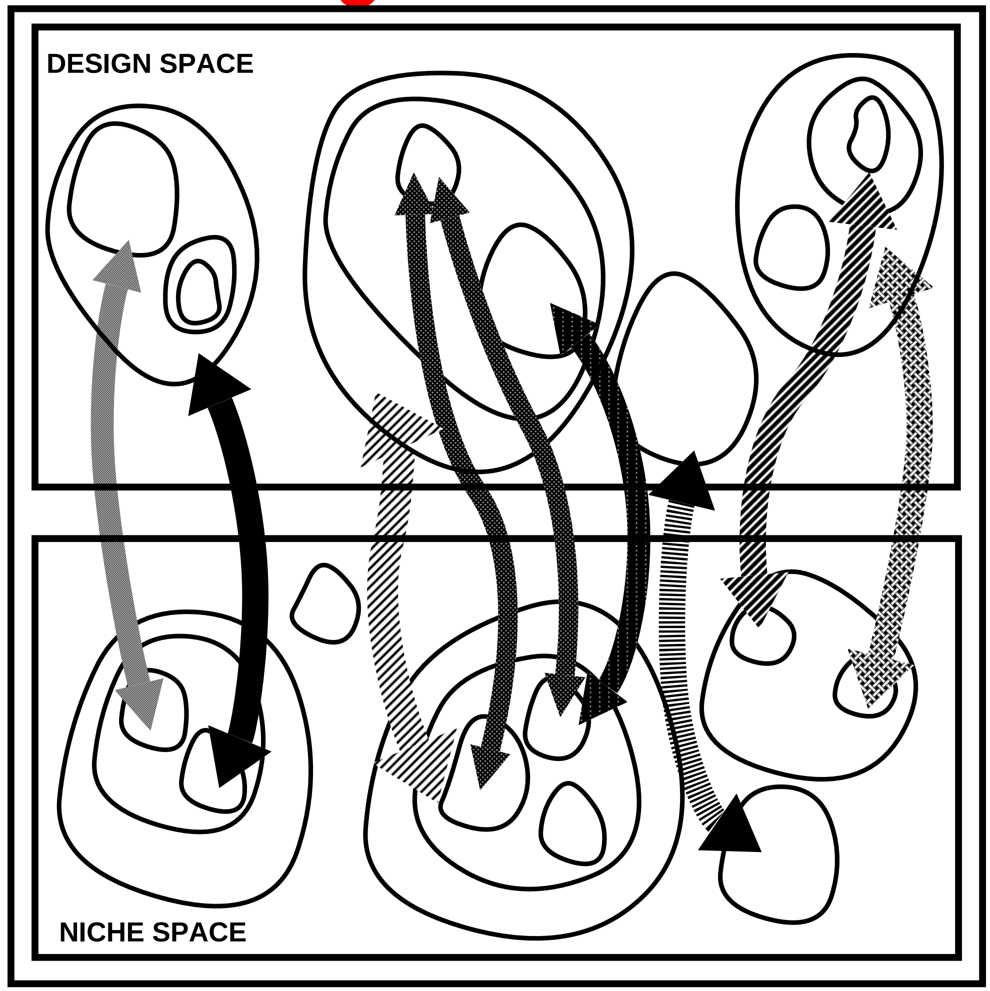
For example there are ill-informed debates about whether things do or do not use representations, which need to be replaced with investigations into the variety of types of information acquired, manipulated, stored, combined, transformed, derived, used ....,

That requires investigating types of ways in which information structure can differ and can change = types of 'syntax'.

Mappings between sets of requirements (niches) and designs

Any set of requirements may be satisfi ed (in different ways) by different designs (see Which? consumer reports).

Any design will related in different ways to different sets of requirements (niches)



Trajectories in spaces of niches and designs

I-trajectories: individual learning and development

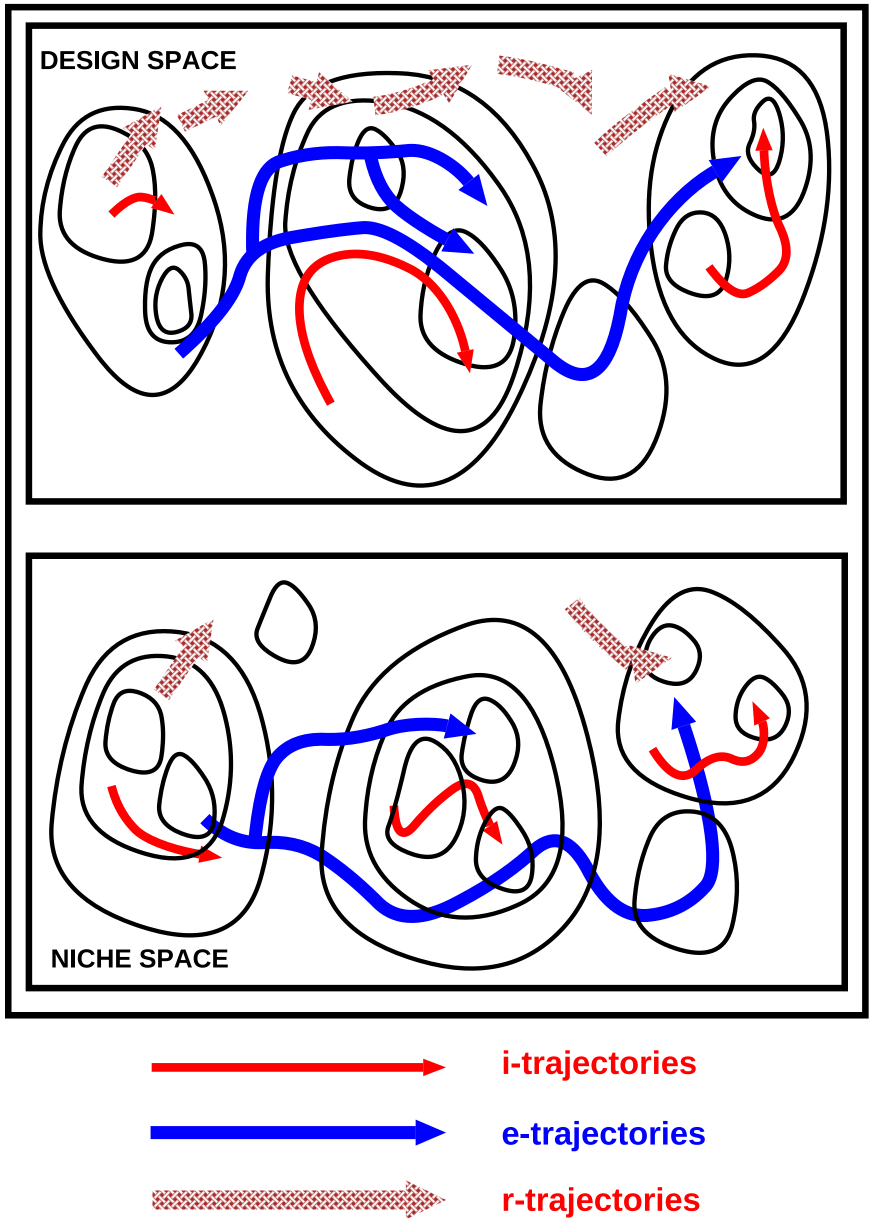
E-trajectories: evolutionary developments

S-trajectories: Social/Cultural changes

R-trajectories:
External Repair, Rearrangement, Replacement, Re-design

An ecosystem: a virtual machine in which multiple niches and multiple design-types coexist, change concurrently, interact, generating multiple concurrent feedback loops of different kinds.

We probably need new kinds of mathematics to characterise such systems precisely – though they may be inherently non-deterministic, e.g. because of chaos.



Trajectories leave traces

- Two complex architectures that fit a particular niche well differ greatly in how they work if they have different developmental and/or evolutionary trajectories.
- In particular humans don't just have one architecture but have layers of architecture that evolved at different times.
- Whether we can replicate that functionality in artificial systems with very different architectures is an open question. (Which may depend on properties of the physical universe.)
- Whether the same architecture can be implemented in different physical machines is an open question.

We need a meta-ontology for types of architectures

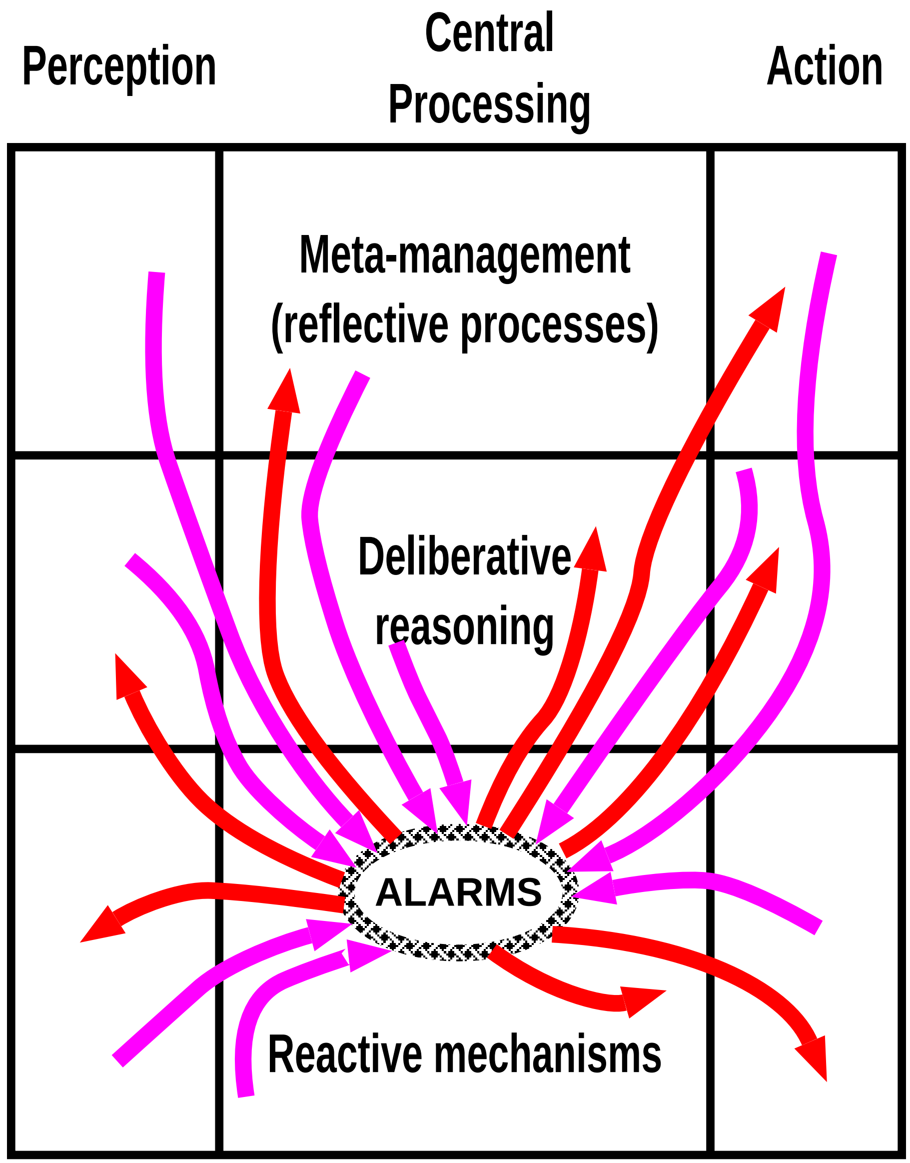
The cogaff schema is crude first example.

These are very crude divisions, and there are probably many intermediate cases between the categories presented here through which the evolutionary trajectory passed, many of those intermediate mechanisms may still be present – e.g. proto-deliberative mechanisms.

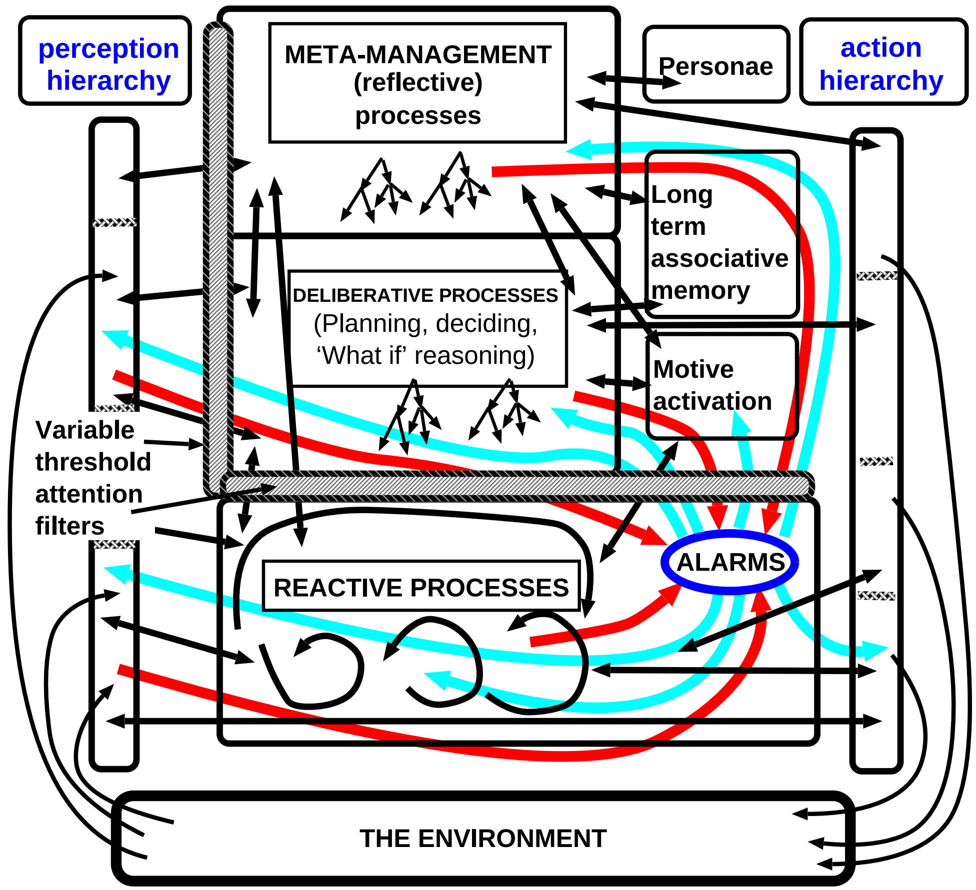
Perception	Central Processing	Action
	Meta-management (reflective processes) (newest)	
	Deliberative reasoning ("what if" mechanisms) (older)	
	Reactive mechanisms (oldest)	

One elaboration of CogAff: Alarms

Some types of phenomena commonly labelled using the word ‘emotion’ can be explain by the functioning of such mechanisms.



Another elaboration of CogAff: HCogaff



This is an instance (or specialised sub-class) of the architectures covered by a generic schema called "CogAff".

Many required sub-systems are not shown.

Different kinds of process may go on in different parts of the architecture – some very old and widely shared, some relatively new and found in very few species.

So there may be different kinds of perception, of recognition, of decision-making.

Embodiment

- Birds could not have evolved with their current ability to travel by flying if they had not previously used ground or water-based locomotion.
- But tha does not mean that birds necessarily use ground or water-based locomotion now.
- Likewise many aspects of embodiment (e.g. as a four limbed two-eyed mammal) were crucial for the evolution of human minds, but it does not follow that something human-like now has to be embodied with four limbs and two eyes.
- Everyone knows humans can be born blind or limb-less (e.g. because of thalidomide) yet develop a human mind.
- Given more time I would argue that many aspects of human intelligence could exist in a disembodied form, including the ability to learn many things, do mathematics, understand and use human language, and have certain sorts of emotional and other affective states which do not depend on bodies, contrary to popular myths.

An old, tempting, and mistaken theory

Concept empiricism is an old, very tempting, and mistaken theory, recently re-invented as "symbol-grounding" theory and endorsed by many researchers in AI and cognitive science, even though it was refuted long ago by the philosopher Immanuel Kant (1781).

Roughly, concept empiricism states:

- **All simple concepts have to be abstracted from experience of instances**
- **All non-simple (i.e. complex) concepts can be defined in terms of simple concepts using logical methods of composition.**

  E.g. if red and line are simple concepts then red line can be defined in terms of them using conjunction

Symbol grounding theories may add extra requirements, such as that the experience of instances must use sensors that provide information in a structure that is close to the structure of the things sensed.

People are tempted by concept empiricism (whatever it is called) because they cannot imagine any way of coming to understand notions like
red line sweet pain pleasure etc.
except by experiencing instances.

KANT: YOU CAN'T HAVE EXPERIENCES UNLESS YOU ALREADY HAVE CONCEPTS.

Alternatives to concept empiricism and symbol grounding

We'll present some alternatives making use of the following ideas:

- **Meanings can be to a considerable extent determined by structural relations between sets of concepts** (i.e. theories can determine or at least constrain meaning)
- **Sensory links can reduce residual indeterminacy of meaning without being the sole basis of meaning:** we call this symbol attachment in contrast with symbol grounding
- **Millions of years of evolution can produce individuals that have some concepts from birth**

  e.g. precocial species such as deer, that can see and run with the herd shortly after birth
- **Genetically determined bootstrapping mechanisms can constrain what is learnt by mechanisms that develop concepts "from experience" –**

  i.e. what is learned by interacting with the environment may include some innate and some empirical content, in varying proportions.

  Research is needed on varieties of bootstrapping mechanisms and different kinds and amounts of innate conceptual information that suffice for different sorts of organisms or machines.

Before elaborating on that we need to survey some general ideas about meaning.

Origins of meaning: the myth of grounding

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

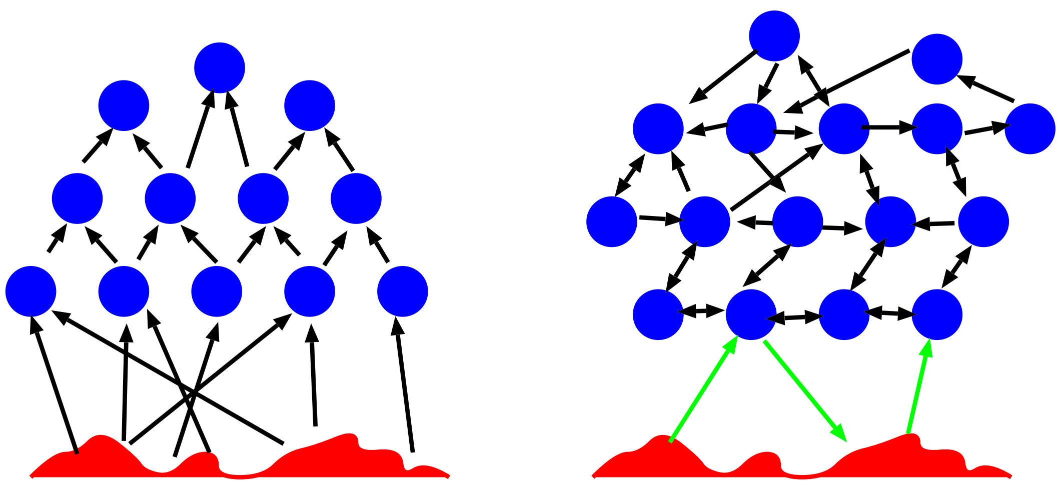
(The theory that all meaning has to be derived by processes of abstraction from sensory experiences).

We distinguish two main sources of meaning

- **the structure of a theory in which 'undefined terms' occur**

  This determines a class of possible models.

- **links to sensing and acting which 'tether' the system to a portion of reality, thereby selecting a particular interpretation (which may be partly indeterminate, allowing future development).**



Symbol Grounding **Symbol Tethering**

Switching experiences

My talk is an illustration of a small part of this project, starting from a comment made by Wittgenstein when discussing the experience of ambiguous figures. He wrote:

> The substratum of this experience is the mastery of a technique.

I don’t really know what he meant by that, but those words slightly modified thus:

> The substratum of any experience is mastery of a large collection of techniques available and ready to be deployed if required, possibly in new combinations.

could be used to express a theory of what is involved in perceiving objects in our environment

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 ???

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.

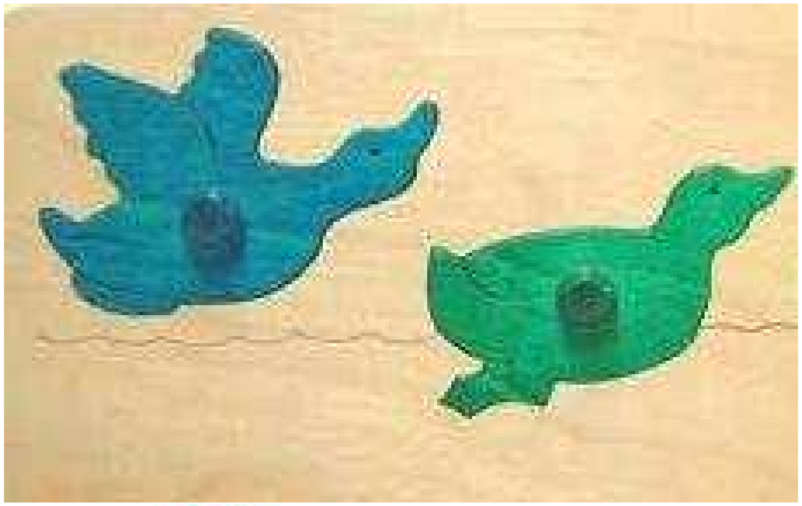
The detailed representational, algorithmic, mechanistic and architectural requirements to support such learning, and the growth of the ontology involved, require much deeper analysis than I can give at present.

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

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.

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 the child's mental simulation mechanism. HOW?

A succession of stages

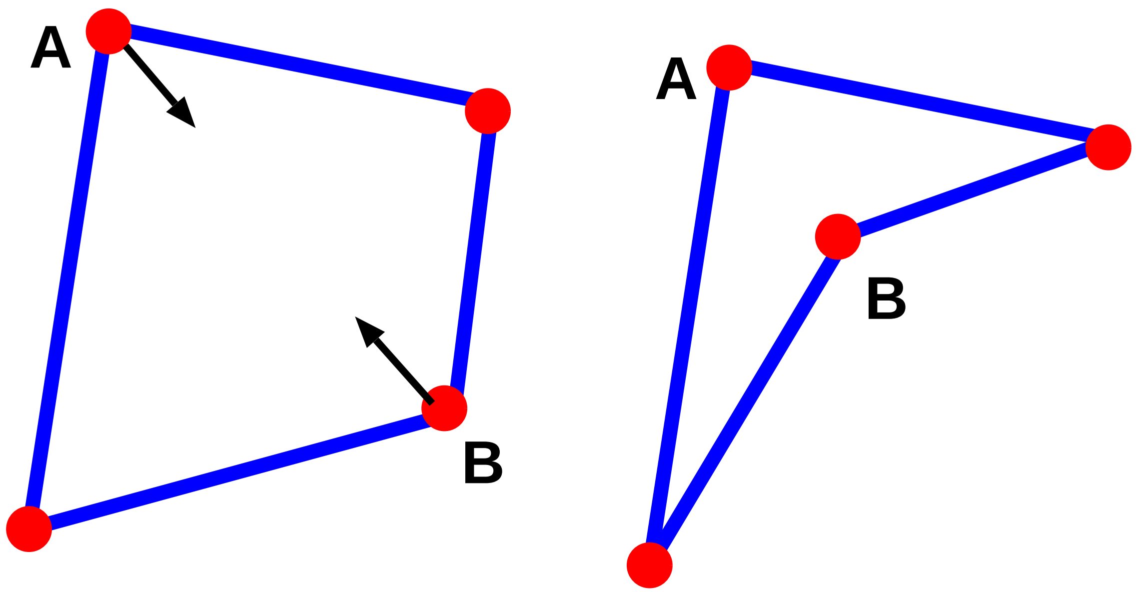
- The process of extending competence is not continuous (like growing taller or stronger):
- The child has to learn about
  - distinct new kinds of objects, properties, relations, process structures, e.g. for rigid objects, flexible objects, stretchable objects, liquids, sand, treacle, plasticine, pieces 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 recent acquisitions.
- The word 'stage' can mislead: 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.
- 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.
- 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.

  CONJECTURE:

  in the first five years 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, causal interactions – some opaque and Humean others transparent and Kantian.

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?**

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 I had ironed the handkerchief 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 the folding examples, coming later.

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 fbor, 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?).

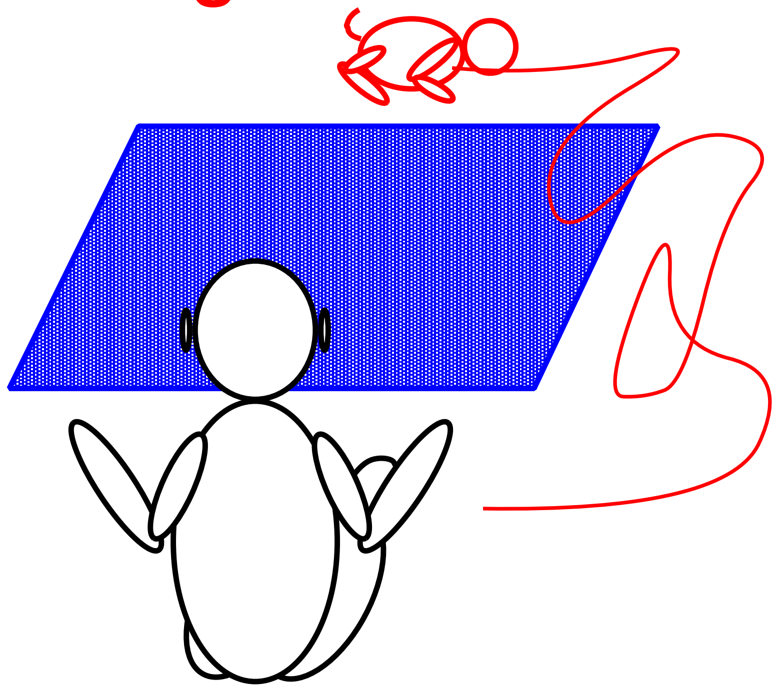
What they have in common includes the fact that if a toy is resting on the blanket or sheet of paper, pulling the edge towards you brings the toy closer too, whereas if you pull too fast, or if the toy is on the fbor near the far edge, pulling will not have that effect. Why not?

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.

Contributors to simulation features

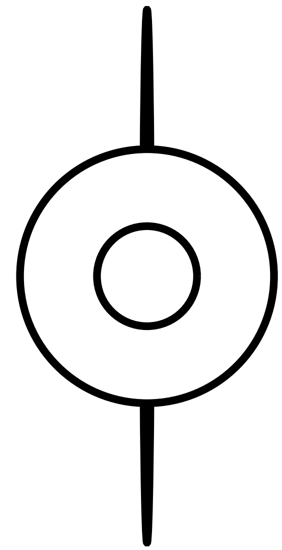
- **We have so far seen that both shape and material can contribute to features of a simulation, including the constraints on what can and cannot change and what the consequences of change are.**

- **Another thing that can be important is viewpoint.**

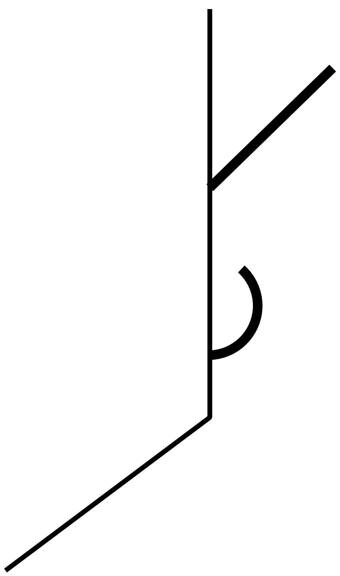
  E.g. viewpoint can interact with opacity of materials, as well as with the mathematics of projection from 3-D to 2-D.

Sometimes a simulation includes a viewpoint

Droodles illustrate our ability to generate a simulation (possibly of a static scene) from limited sensory information (sometimes requiring an additional cue, such as a phrase ('Mexican riding a bicycle', or 'Soldier with rifle taking his dog for a walk').



In both of these two cases the perceiver is implicitly involved: one involves a perceiver looking down from above the cycling person, whereas the other involves the perceiver looking approximately horizontally at a corner of a wall or building.



In both cases the interpretation includes not only what is seen but also occluded objects: the simulation depends on knowing about opacity.

This does not imply that we have opaque objects in our brains: merely that opacity is one of the things that can play a role in the simulations, just as rigidity and impenetrability can.

The general idea may or may not be innate, but creative exploration is required to learn about the details.

Viewpoint matters - some viewpoints are ‘vicarious’

The importance of viewpoint is obvious for any animal that moves, for self-motion can change the appearance of objects in a manner than depends on the shape of the object, its material, the lighting, the type of motion and what else is in the environment (actual or potential occluders).

What is not so obvious is that a part of the body, e.g. a grasping hand, may have a ‘viewpoint’ that is different from the visual viewpoint and which changes differently, as the hand moves or as something in the environment moves. E.g. something moving can block the eye’s view of an object while leaving the hand’s ‘view’ (route to the object) intact, and vice versa.

Likewise another person (or a child that needs help) may have a different and changing viewpoint.

So an intelligent animal or robot may need to be able to construct and reason about, or simulate properties of, ‘vicarious viewpoints’, i.e. viewpoints for others, or for different parts of oneself.

Likewise viewpoints for oneself in the future (predicting what you will be able to do) and in the past (thinking about what you could have done).

I don’t know how many other species can do these things.

I don’t think a new born baby can.

How does that ability develop/grow – and what has to change in the baby when it does?

We can see things from more than one viewpoint

- **Vicarious affordances:** a parent watching a child needs to be able to see what is and is not possible in relation to the child's needs, actions, possible intentions, etc.
- This may include such things as visualising the scene from the child's viewpoint, including working out what the child can and cannot see – and the possible consequences of the child seeing some things and not seeing others.
- Some people can draw pictures of how things look from some other place than their current location.
- This ability to contemplate the world from multiple viewpoints, not just one's own current viewpoint, is essential for planning, since at some future state in the plan one's location and orientation could be very different from what it is now, yet it still needs to be reasoned about in extending the plan.
- The ability to perceive and use information about 'vicarious' affordances (affordances for others) and the ability to perceive affordances for oneself in the past (e.g. thinking about a missed opportunity) or future (planning to use opportunities that have yet to be created) may use the same mechanisms **because both are disconnected from current viewpoint**.

  Could that be the main point of substance behind all the fuss about "mirror neurones"?

Seeing things from the viewpoint of your hand

The importance of hand-eye uncoordination!

- **The evolution of body-parts for manipulation that can move independently of a major sensor perceiving what's happening (hands *vs* beak or mouth) had profound implications for processing requirements.**
- **Most animals are restricted to doing most of their manipulation with a mouth or beak, which cannot move much without the eyes moving too.**
- **If your eyes move as your gripper moves, because they are closely physically connected, then the sensory-motor contingencies linking actions and their sensory consequences will have strong, useful regularities that can be learnt and used.**
- **If a gripper can move independently of the eyes then the variety of relationships between actions and sensed consequences explodes.**

  The explosion can be reduced by modelling action at a level of abstraction removed from sensory changes: e.g. by representing actions as altering 3-D structures and processes (including subsequent actions), independently of how they are sensed.
- **The mapping between sensory data and what is perceived becomes very indirect, and there may need to be several intermediate layers of interpretation: perception becomes akin to constructing a structured theory to explain complex data. (Compare the 'dotty picture' example, above.)**
- **This is one of many reasons for NOT regarding perception as simply concerned with detecting sensory-motor contingencies.**

Sensory-motor vs action-consequence 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 internal motor signals and internal sensor states (subject to prior conditions),**

to

> **a more ‘objective’ understanding of the environment in terms of action-consequence contingencies linking changes in the environment to consequences in the environment,**

followed 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 other than the mouth that could manipulate objects and be seen to do so.

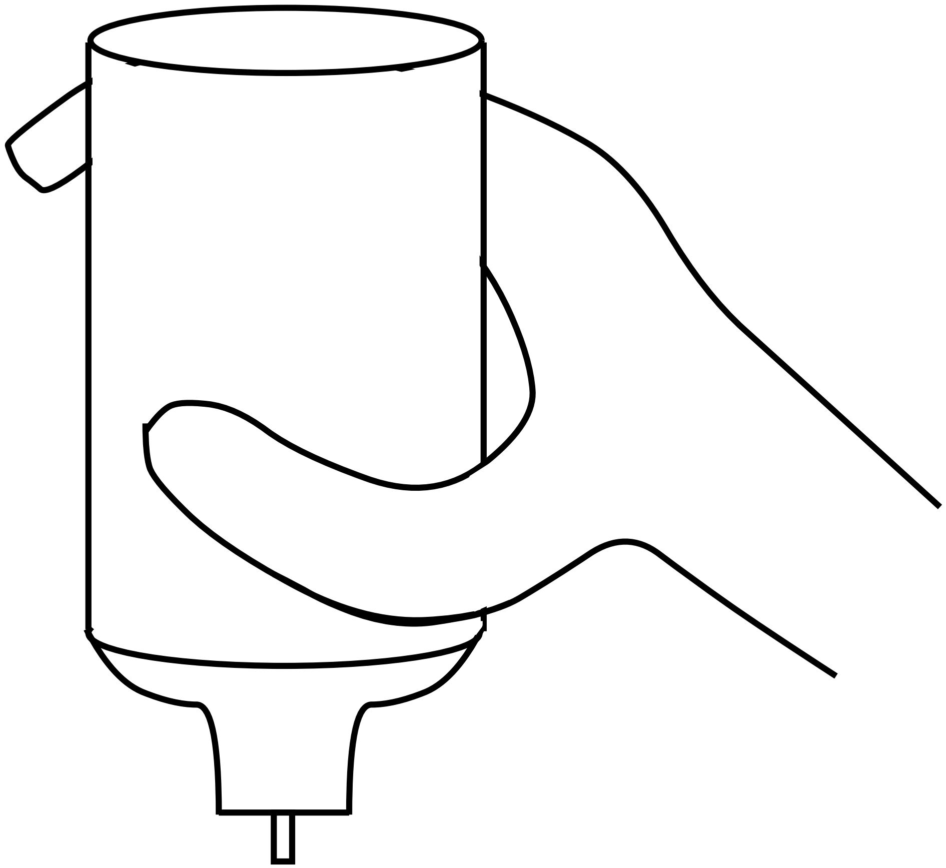
> **However the cognitive developments were not inevitable consequences: e.g. crabs that use their claws to manipulate food do not necessarily have the generative competence.**

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.**



Note that the model used here assumes uncompressability rather than rigidity.

Also, our ability to simulate what is going on can also explain why as more of the shampoo is used up you have to wait longer before squeezing.

Beware of intellectual fashions

There is a wide-spread fashion (partly fuelled by Rodney Brooks, who has recently changed his mind) which disparages

- **GOFAI,**
- **planning,**
- **symbolic representations,**
- **Reasoning with Fregean representations**

and emphasises

- **dynamical systems,**
- **embodiment,**
- **sensory-motor contingencies.**
- **neural mechanisms**

This fashion provides an excellent framework for studying insects, and maybe many other types of animals (precocial species), but certainly not humans, and probably not nest-building birds, hunting mammals, primates and possibly other altricial species.

Most animals are far more embodied than humans are.

However, evolution discovered the need for disembodied competences long before humans ever thought about them.

Exercises for disembodied minds

Example: Elastic band – Making an H (REF de Sauvy)

Can you make an outline capital H using an elastic band and pins on a pin board.

Various sub-tasks.

Several questions. E.g. why can’t young children do this, and what changes when they learn to do it?

Some of the evolved mechanisms had ‘unintended’ consequences, i.e. consequences for which they were not selected.

Transfinite ordinals

Can you imagine a never ending row of dots going off to the right?

....

What about an infinite set of positive integers?

0 1 2 3 4 5 6 7 8 9 10 11 12 ....

Now imagine the even numbers being lifted out of the row

and then moved to the end on the right and appended to the others

or reversed before being appended

What comes first along the modified row of numbers, 3,489,267 or 24 ?

Good things can have bad side-effects

Not all ‘unintended’ consequences of selected mechanisms are good:

compare nationalism, racialism, positive feedback loops causing family rows, philosophical muddles (e.g. which way is the Universe moving?) and all the harm done by religion, including mind-binding inflcted on young children, cruelly restricting development of orthogonal competences, as foot-binding cruelly restricts development of feet.

Biological bootstrapping mechanisms

- **There are some species 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 biological species have most (and in many cases all) of their cognitive competences pre-programmed innately (e.g. precocial species such as chickens, deer, reptiles, fish and most non-vertebrates).
- E.g. 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 species (e.g. corvids, hunting mammals, primates) that was not adequate – and 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).
- In both sorts of species 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 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.

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 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.**

- **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 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.**

KEY IDEA

In addition to physical growth –

biological organisms also grow their own

information-processing architectures

(which are virtual machines, not physical machines,

and therefore can change their structure more radically.)

There are probably many more ways this can happen in nature or in machines than we have thought of.

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/fi g/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/fi nger
  (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**

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

A spectrum of competences

- Every organism is a mixture of both kinds of capabilities:

  pre-configured — constructed

- Not all of the first kind are manifested at birth/hatching – many are 'time-bombs'.
- Architectures for altricial species can do many things that are not directly biologically useful:

  including (possibly dangerous) exploration of a space of possibilities.

- Architectures 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 favourite, 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

(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 IJCAI paper we distinguish two sources of meaning

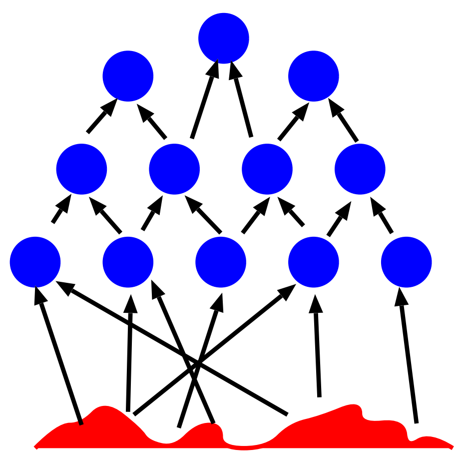
- **the structure of a theory in which 'undefined terms' occur**

  (where the structure limits the class of possible models/interpretations)

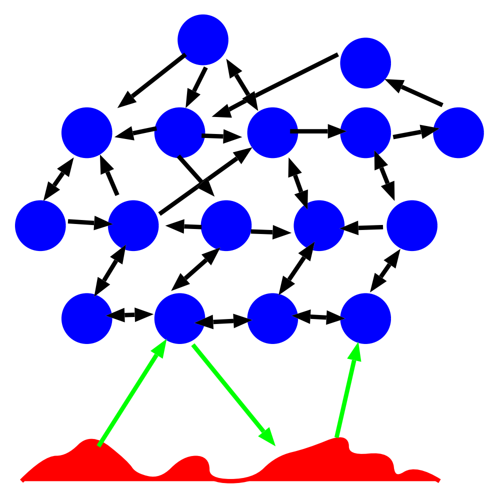
- **links to sensing and acting**

  (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. presuppose 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*. I think his idea was based on something like this simulation theory.

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

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

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

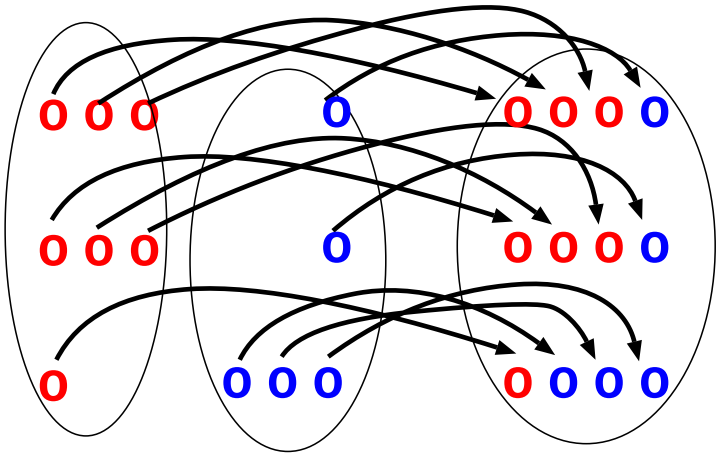
What sort of equivalence are we talking about?

I.e. what does “=” mean here?

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

That can be defined logically, but the idea can also be understood by people who do not yet grasp the logical apparatus required to define the notion of a bijection.

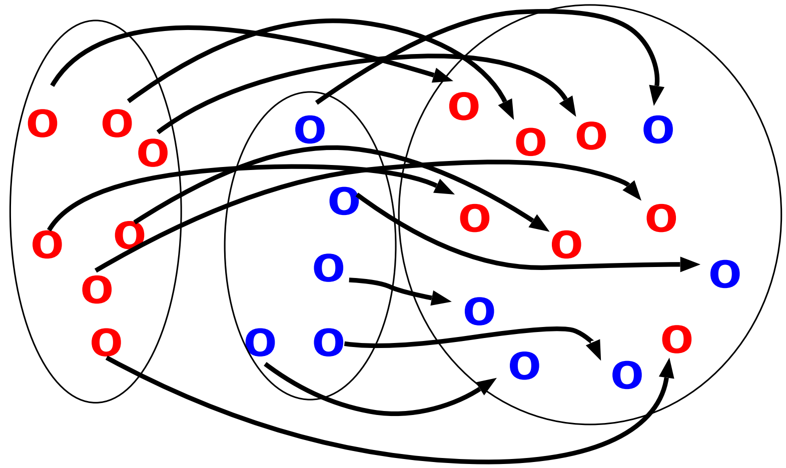
SEEING that 7 + 5 = 12



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

- I've 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 I 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 ....).

additional points

- We need to find out how many different kinds of simulative capabilities a child, a chimp, a nest-building bird, a domestic robot, needs to acquire.
- 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.
- There will not be a simple step: understanding causation
- 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.

Conclusion

- I 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 I am 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/