# Meta-Morphogenesis:

## Evolution and development of mechanisms for producing minds

An approach to biology-based computational cognitive science.

Evolution, individual development, learning, and cultural change

producing new mechanisms of

evolution, individual development, learning, and cultural change

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

This will be added to my talks directory (PDF):
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk100

See also
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html

# Still under construction.

Please send comments and suggestions to

Aaron Sloman a.sloman@cs.bham.ac.uk

# Abstract for Cambridge presentation

I'll present an idea, that could help to bring different kinds of research together in fruitful cooperation: The idea is that many of the developments in biological evolution that are so far not understood, and in some cases have gone unnoticed, were concerned with changes in information processing. The same is true of changes in individual development and learning: they often produce new forms of information processing.

Examples include:

learning new ways to learn
development of new forms of development
evolution of new types of evolution
evolution of new types of learning
evolution of new forms of development
development of new forms of learning
how new forms of learning support new forms of evolution ....
Then add evolution of culture to the mix...

Much current research in AI/Robotics, psychology, neuroscience and biology assumes that the main function of brains is to control movement. In contrast I'll argue that there's a wide variety of types of control and some of the forms of control that evolved later have very little to do with control of movement, though many are related to understanding what kinds of structures and processes can and cannot exist in the world. But understanding the environment need not be motivated by a need to produce or prevent motion.

Later developments, such as the development of mathematical and philosophical interests, thoughts and discussions are even more remote from any requirement to produce motion, even if some of them were originally provoked by problems of coping with a complex environment.

Contrast designing

- a chess player that always tries to win
- a chess player that moves the pieces on the board
- a chess player that can play against a novice in a helpful way.

# Hoped-for effects of this presentation to Cambridge students

**Some of today's students will be tomorrow's teachers. What should you teach?**

There is currently much discussion of computing education in schools and this is leading to major new developments, including less emphasis on teaching students to use existing software tools (ICT) and more emphasis on teaching students to design and implement their own working systems (Software Engineering).

At present most of the educational emphasis is on engineering (making and doing), not science (understanding how the world works and building models to demonstrate and test our theories and explanations).

That emphasis is justified by the shortage of applicants for computer science degree courses, and the shortage of well qualified applicants for IT jobs in industry.

All of that is important, but in the long term it is just as important to educate youngsters to use computational ideas in trying to understand how things work, e.g. how animals perceive, learn, act, communicate; how biological evolution works; how social and economic systems work; how diseases like cancer work, and how child learning develops.

These scientific goals are all concerned with information processing systems. Learning to think about such systems after leaving school is too late, like learning play a violin only after leaving school.

So we must ensure that not just our future engineers but also our future scientists and other thinkers are intellectually equipped for the task of understanding existing complex information-processing systems, not only building new ones that are useful or fun.

# Themes (1): The universe, information and evolution

This talk presents a subset of themes from an ambitious long-term multi-disciplinary research project.

- The universe contains matter, energy, and information.

  For a partial answer to "What is information?" see Sloman (2011b).

  Like the concepts "matter", and "energy", "information" cannot be defined explicitly without circularity.

  **NB** "Information" is not used here in Shannon's (syntactic) sense, but the ordinary sense: information is about something, actual or possible.

  It's not explicitly definable, but (mostly) implicitly defined by our theories about information.

- Information, like matter and energy, is usable, in different ways:

  The key use is for control – control of physical processes or control of information processing.
  (Often the potential use does not produce actual use – for various reasons.)

- The variety of types of information and types of use and influence, have been steadily increasing as a result of biological evolution:

  it produces new users, with new needs, new capabilities, new ways of acquiring and using information.
  It can also produce new predators, new prey, new competitors for resources, and new collaborators, ...

  All of this depends on new mechanisms, including new information-processing mechanisms.

- For the earliest organisms almost all information processing was molecular (chemical) – there were no brains or nerves.

  Chemical (molecular) information processing remains important for all life, including building brains.

  We probably don't yet understand the full potential of chemical information processing.

- We still have much to understand about ways in which information-processing changed, over millions of years, and in shorter periods.

# Themes (2): Challenges for a science of information

- Despite many advances, we still understand only a small subset of what sorts of information-processing can occur.

  We have learnt about a subset of

  – types of information content
    E.g. in animal vision
    See `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#gibson`

  – types of information bearer (forms of representation)

  – types of information manipulation
    (construction, derivation, comparison, analysis, storage, retrieval, application, ...).

  – types of information processing mechanism (natural and man-made).

  – types of multi-mechanism information processing architecture

  – types of self-constructing, self-monitoring, self-modifying information processing architecture

  – types of virtual machinery (including “layered” virtual machines).
    Computer systems engineers have learnt about and developed many different types of virtual machinery and mechanisms to support them and their interactions: some running on a single computer, some distributed over networks, with various kinds of interconnection with other machines or with the physical environment or underlying physical mechanisms (e.g. mechanisms supporting “virtual” memories.

    I conjecture that biological evolution “discovered” the need much earlier, and produced many types of virtual machinery, including human minds. (Sloman, 2010, 2011a)

  There’s much work to be done by future researchers – I’ll try to give some pointers.

# Themes (3): What the physical world makes possible

- All of this biological diversity depends on the underlying combinatorial richness of the physical world, at many levels.

  E.g. sub-atomic, atomic, molecular, in organisms, in ecosystems, in social systems, ...

- Although much has been learnt about physical substances, physical structures, physical processes, types of physical interaction, ... we are still in the early stages of learning about ways in which physical structures and processes can be related to types of information-processing, including information-processing in virtual machinery.

  See (Sloman, 2011b), and
  UK Grand Challenge 7: `http://www.bcs.org/content/conMediaFile/9556`

- Most human-made information processing systems (especially during the last half century) make use only of streams of discrete changes in collections of switches (and other bi-stable, or multi-stable, devices),

- In contrast, biological information processing systems seem to use far more varied mechanisms supported by physical matter and physical processes (including different types of virtual machinery).

# Themes (4): What we have and have not learnt since Darwin

- We now already know far more about information processing systems than scientists and philosophers did at the time of Darwin:

  He was unable to answer some of his strongest critics – but we can give new answers, that are not yet widely understood – and are ignored by most philosophers.

  This can be used to shed new light on some very old problems about the nature of mind, self-consciousness and mind-body relationships. Sloman (2010)

  E.g. since the 1940s, the notion of non-physical machinery, virtual machinery, that can **do** things has developed far more sophistication than most people have noticed

- **NOTE:**

  **Ada Lovelace understood some of this a century before Turing.**

  She wrote in 1842:

  "The operating mechanism can even be thrown into action independently of any object to operate upon (although of course no result could then be developed). Again, it might act upon other things besides number, were objects found whose mutual fundamental relations could be expressed by those of the abstract science of operations, and which should be also susceptible of adaptations to the action of the operating notation and mechanism of the engine. Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent." (In Note A of Lovelace (1842).)

  Clearly she had some understanding of the idea of virtual machinery.
  I don't know whether she appreciated that such capabilities could also be linked up with sensory and motor apparatus in a machine interacting with its environment.

# Themes(5): Implications for education

- **Education for doing vs education for explaining and understanding**

  Computing education can serve many needs, including producing people who are good at **changing** the world, to meet **practical** needs.

- But the equally important need to produce **thinkers** who can help us **understand** our world, is not being served well by the almost exclusive emphasis on teaching about applications of computing, and how to create them.

  In part this is a result of the limited computational education that current computing teachers, and most computing "experts" in academe and industry, have experienced in the last few decades.

  **NOTE:** The opposite complaint used to be made (e.g. by The RSA `http://www.thersa.org/`) a few decades ago: our educational system was not enough concerned with teaching people how to **make** and **do** things: education was alleged to be too intellectual and not practical enough. We now seem to be paying the price of the swing to the opposite extreme, at least in computing?

- Fred Brooks wrote that scientists build to understand and engineers understand to build, and seemed to argue that there is no computer science, only computer engineering (Brooks, 1996) `http://www.cs.unc.edu/~brooks/Toolsmith-CACM.pdf`

  But that ignores all the non-engineering work done by computer scientists,

  including those trying to understand natural and artificial forms of computation.
  Compare John McCarthy: "Towards a Mathematical Science of Computation" (IFIPS 1962)
  `http://www-formal.stanford.edu/jmc/towards/towards.html`

- A good education for the 21st century should include computation as **science** and bring out the deep two-way connections between making models and understanding.

# Themes(6): model-building and understanding

Education on programming for understanding can start from attempts to identify things various animals (including humans) can do, and then introduce ways of attempting to replicate or model aspects of those competences:

- Start with relatively simple competences and progress to increasingly complex ones

  Competences that evolved later are not necessarily all harder to model, if we allow models that simplify. For example, very few animals can play noughts and crosses (Tic-Tac-Toe), but it's easier to get a computer to play it than to do things that much older species do – including ants, termites, nest-building birds and others. Could such a program think like a human? Do you know how you think?

- Try to identify and model different **sorts** of competence. If the tasks are too difficult, at least try to characterise what is difficult about them and what would be needed to model them. E.g.

  1. Think about making machines that
     - perceive things (using touch, smell, resistance to force, sound, vision, or other sensors)
     - want things or want to be in certain states or want to avoid certain objects or states, so that they produce or avoid behaviours in accordance with what the want
     - have various ways of moving which they can control to achieve goals,
     - can learn about the world they inhabit, including building explanatory theories, which they use in planning or executing actions
     - can be aware of their own states of mind,
     - can understand a simple language and use it to think with or communicate with.
  2. When you have your models working, consider how you could test whether they are good models of how humans or other animals work? E.g. in which environments would they fail?

- This involves studying different mechanisms, architectures, representations, and more.

# Themes(7): Products of expanded education

**Our educational system needs to produce more people who can think deeply about such things:**

It can't be left to psychologists, neuroscientists, biologists, philosophers –

they need to be educated about natural and artificial varieties of computational mechanisms – including ones we don't understand yet.

Making progress is very hard:

but it helps if we think about the right questions,

many of which are unobvious – and cross disciplines.

## So, how can we make progress?

I'll illustrate questions to be asked, by giving examples of animal competences.

A long term collaborative project can extend the collection of examples, organise them, find intermediate examples, and search for explanatory mechanisms, with help from several disciplines.

A fruitful approach could consider transitions (competence changes) in these trajectories:

1. Evolutionary trajectories for a species, or succession of species;
2. Developmental and learning trajectories for an individual;
3. Social/Cultural trajectories for a community;
4. Trajectories for an ecosystem

Identifying observed, inferred, or conjectured transitions can help us get closer to modelling "before" states, "after" states, or both.

# Change gear

## Now consider biological information processing

A vast amount has been learnt, in many different disciplines or sub-disciplines.

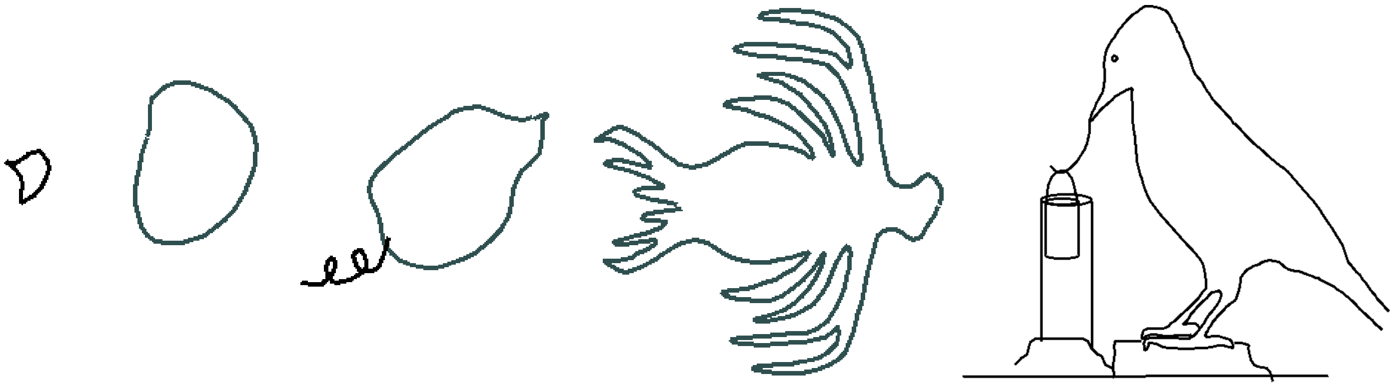
I'll point to only some very general, relatively well known, ways in which organisms vary.

Some of the changes are impossible unless others have occurred first: e.g. visual perception of distant objects requires optical sensors.

The subdivisions I draw attention to are very coarse-grained: empirical research has revealed and is likely to reveal many more more detailed sub-divisions.

For now a broad-brush picture will have to suffice to characterise the variety of research sub-goals that contribute to the overall goals of this meta-morphogenesis project (explained more fully below).

# All organisms are information-processors but the information to be processed, the uses of the information, and the means of processing have varied enormously, between the earliest microbes and sophisticated modern animals.



Evolution isn't just a uniform process in which natural selection uses random chemical changes:

Many different things can influence evolution, including

- What's in the environment

  Physical structures and processes, which vary across the planet and at different times.

  Other information-processing systems – prey, predators, competitors, conspecifics, offspring, ...

- Products of previous evolution

  Previous physical developments and previous computational developments in the species, or in other species.

  These provide new requirements – e.g. new requirements to control articulated body parts, or to get the most benefit from new sensory mechanisms.

  They also provide new opportunities: new platforms for further development.

# Changing environmental influences on evolution

## Types of environment with different information-processing requirements

What information processing mechanisms could be useful in the following epochs?

- Microbes in a chemical soup
- Microbes in a soup with detectable chemical gradients
- Soup plus some stable structures (places with good stuff, bad stuff, obstacles, supports, shelters)
- Things in the environment that have to be manipulated to be eaten (e.g. disassembled)
- Organisms with controllable manipulators - for eating, building, fighting...
  (products of previous evolution: providing new opportunities for information-based control)
- Food/prey/predators/mates with detectable aromas (more chemical information bearers)
- Environments with food that tries to escape (prey).
- Organisms in environments with things that try to eat them (predators).
- Environments with and without places to hide from predators
- Environments with and without places for prey to hide
- Prey, predators, collaborators, whose behaviour can reveal intentions, interests, knowledge...
- Mates with preferences
- Competitors for food and mates
- Collaborators that need, or can supply, information.
- and so on .....

Can we analyse the changes in information processing capabilities made useful/possible by the above changes, so as to form a collection of "dependency trees" showing possible evolutionary and developmental trajectories?

(A sort of generalisation of Turing's 1952 paper on physical morphogenesis?)

# Changing requirements

- How do the information-processing requirements change across these cases (microbes, insects, vertebrates .... in various environments)?

  Identifying the changing niche-pressures can be very hard, but by trying to design working models of the organisms, and finding the problems we can discover some.

- How can we learn to see what problems young children and other animals are encountering?

  Often, problems, and opportunities, only become visible when you try to build a machine to do what the animal or infant does.

  We (adults) can also make ourselves a bit like younger learners by putting ourselves in new situations where we play, then discover new structures and regularities

  E.g. in activities proposed in (Sauvy & Sauvy, 1974).
  E.g. see the puzzles in `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler`

- It's hard to notice the right questions to ask about what is learnt.

  Noticing behavioural changes may not tell us what forms of perception and thinking led to the change: e.g. perceiving and reasoning about affordances (possibilities and constraints), deliberating about what to do, remembering, ... (Sloman, 2011c, 2006)

- How do the phenomena of self-awareness (including qualia as contents of consciousness) fit into this framework?

  I think the answer is connected with evolution of species that grow virtual machinery, including virtual machines for self-monitoring and self-control.
  But that's another talk. `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk95`

  An explanation for the ineffability of qualia is suggested in (Sloman & Chrisley, 2003)

# Animal intelligence is not all about control of movement

There's a huge amount of research on trying to get robots to walk, run, pick things up, go smoothly through doorways, move so as to avoid obstacles.

Those are important, non-trivial research targets relevant to explaining the behaviour of many types of animal, including insects, fishes, birds, squirrels, grazing mammals, etc.

The ability to control real time interaction with an environment could be called "on-line" intelligence: it requires a lot of information to be continually acquired through sensors (including haptic and proprioceptive receptors, and in some cases the vestibular system – semi-circular canals), stored transiently and used in feedback control. The information is continually replaced during and after use.

There are other uses to which information about the environment can be put, which requires information acquired to be stored for longer times, and to be integrated with other stored information.

One example is incrementally storing and integrating metrical and or topological information acquired while moving around in the environment, about places, routes, space occupants, obstacles etc..

This is called SLAM, "Simultaneous Localisation and Mapping".
http://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping

Online control needs different mechanisms from deliberative control, detecting and reasoning about affordances, formulating goals, making plans to achieve the goals, executing the plans, and evaluating progress.

Much of the functionality of deliberative mechanisms produces only internal changes, without any external movement or action. See (Sloman, 2006) and

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

This has implications regarding functions of vision, which are misunderstood by many.

A visual system that merely labels recognized objects, or merely builds a model of what's in the environment does not necessarily understand what is in the model nor what the implications are.

# Varieties of visual/cognitive competences

## Conjecture:

Widespread animal abilities to perceive and reason about affordances, and to form intentions, discover plans, and store those plans, seem to be precursors to the human ability to discover and reason about mathematical properties of space and time.

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

And also precursors to human language
`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang`

Currently there are no good working models of these phenomena.

# Some specially important transitions

Forms of biological information processing that exist now are products of biological information processing over many stages of evolution and development, including cultural evolution in the case of humans.

- Some forms of information processing are concerned only with online control of interactions with the immediate sensed environment (Only sensorymotor processing)
- Simple organisms can represent only their own sensor and motor signals and internal states: they use only somatic semantics. Others can refer beyond the skin, using exosomatic semantic contents.
- Some exosomatic contents refer only to physical objects, properties, relations, processes, states of affairs, etc. (Including physical objects that happen to be organisms.)
- Some exsomatic contents refer to the past, the future, remote locations, invisible entities, insides of things (e.g. kinds of material). (When are these useful??)
- Others have meta-semantic competences they can refer to things that refer and the contents referred to:

  This requires the ability to deal with "referential opacity", "intensionality", etc.
    ("Theory of mind" studies in developmental psychology refer to special cases (Apperly, 2010).)

  Some of these competences involve reference to the individual's own contents
    (self-awareness of perceptual, thinking, reasoning, motivational states and processes)
- Some individuals can do only Humean causal reasoning – using associations, correlations, statistics, probabilities – what always or usually happens.
- Some can do Kantian causal reasoning: using the ability to work out what must happen? – An important aspect of human development. (Karmiloff-Smith, 1992; Sloman, 2008)

# What can be learnt by interrogating the environment?

Topics for further investigation (a few examples):

- some of the ways nature can be interrogated, e.g.
  - perceiving
  - acting while perceiving (e.g. exploring)
  - getting information from others who have already acquired information
  - combining information acquired in different places, at different times, in different ways. How?
- some of the kinds of information that can be acquired by such interrogation, e.g.
  - about what **particular things** and what **types of things** exist in the environment
  - about possibilities for change and limitations of possibilities (what can and what cannot happen)
  - about the directly accessible portions of the environment
  - about remote places, past events, future events
  - generalisations about what happens under what conditions (laws)
  - limitations and benefits of **particular forms of representation**
  - the need to **modify or extend current ontologies** – notions of what can exist (Sloman, 1978, Ch 2)
- some of the things that can be done with the information, e.g.
  - achieving practical goals (controlling actions, changing the environment, including online control)
  - formulating new kinds of theory and new kinds of goal
  - understanding causation and making correct predictions
  - explaining WHY things are as they are, in two ways:

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

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

# Concepts and ontologies

Philosophers tend to distinguish propositions, which are capable of being true or false, from concepts, which are neither, but can be combined to form true or false (or in some cases incoherent) propositions (questions, etc.)

- Concepts are combinable meaning-fragments used for expressing semantic contents:

  motor signals, sensor infomration, thoughts, questions, commands, intentions, plans, theories, predictions, explanations, stories, ... (depending on type of organism)

  The concepts “horse” and “ant” are neither true nor false, whereas “Every full grown horse is bigger than every ant” is true, and “An ant can swallow a horse” is false.

- We tend to assume that all concepts used by humans correspond to words or phrases in human languages, but an individual can use non-verbal meaning fragments, for example, in percepts or action control.

  E.g. before a cat jumps from the ground to rest on a narrow wall, she needs to take in information about how high the wall is, how it is related to her starting location, how wide the top surface is, in which direction the wall is, where it wishes to end up, what kind of force it needs to exert on the ground and how quickly, etc. Cats certainly cannot express all that in a communicative language but they must encode the information in some form internally in order to be able to use it, in deciding to jump, and deciding how to jump, etc. The information they use need not be expressible in English.

- The set of concepts (recombinable meaning fragments) and forms of meaning composition used by an individual implicitly specify sets of possible entities, states, processes, causal interactions, generalisations, etc. to which the individual has access: the concepts available determine the ontology usable by the individual, i.e. the implicitly presupposed theory about what sorts of things can exist.
- We know very little about the kinds of meaning fragments, or ontologies most species (including humans) have access to, nor how they vary between individuals, or over time.

# Ontological extensions in species and individuals

**Different organisms are capable of referring to different things.**

- What they refer to may not be expressible using human language, but there is a clear difference between what a new chick can look for and react to (food to peck, a hen to imprint on), and a new foal, which somehow struggles to support itself on four legs, then finds and sucks a mare's nipples – or runs with the herd if attacked:
  These species clearly access and use different kinds of environmental information.

- How does this differ from the information their evolutionary precursors were able to acquire, represent, store, manipulate and use?
  Can we identify the ontological and representational transitions?

- Evolution produces representational differences across generations or species;
  Learning and development change meaning capabilities within individuals. How?
  The information a human neonate can represent seems to be very different from what an adult carpenter, historian, violinist, mathematician, athlete, programmer or quantum physicist can represent.
  What changes, and how? Not all species have so much potential for conceptual change in individuals.

- Concepts are meaning fragments that can be combined in different ways to form goals, questions, plans, hypotheses, intentions, predictions, control signals etc. In humans (and which other species?), information contents developed later are richer, more varied, and more complex than those available to neonates.
  Research question: how many different sorts of minimal meaning unit are used in organisms, and in how many different ways can they be combined to form larger meaning units?

- **What happens when individuals acquire new "meaning units" – new concepts?**

# Ontological extensions: trivial and non-trivial

"Trivial" ontological extensions introduce new terms definable in terms of old ones (E.g. A pentagon is a polygon with five sides.)

Use of explicitly defined concepts allows nothing new to be expressed: though mode of expression may be much more economical and much easier to understand.

"Non-trivial" ontological extensions introduce terms that are not explicitly definable.

E.g. Newton's concept of mass, the concept of a gene, an animal's concepts of 3-D objects or object fragments that can exist independently of whether and how they are perceived, a child's concepts of the kinds of **stuff** of which objects are composed, etc.

Often such concepts are not definable in terms of sensory and/or motor signals, since they don't refer to contents of such signals, but to external causes. They are exo-somatic, often a-modal, concepts.

Concepts of spatial location, distance, size, orientation, adjacency, straightness, containment, causation, and other spatio-temporal concepts necessary for representing and reasoning about our world are all exo-somatic: they refer to what's outside the perceiver's body.

Including concepts used by robots in SLAM: Simultaneous Localisation and Mapping

New terms are often implicitly defined by the theories that use them, but not totally defined by those theories.

So theories and modes of observation and measurement can change while old concepts remain in use, slightly modified.

A child or animal learning to think about things that exist independently of being perceived or acted on uses an ontology not definable in terms of patterns in its sensorymotor signals. I.e. an exosomatic ontology referring to things in the environment.

(Concept empiricism and "symbol grounding" theory erroneously assume this is impossible. (Sloman, 2007))

# Ontology extension – beyond the skin

- Trivial ontological extensions merely introduce new labels defined in terms of old ones – enabling nothing new to be expressed, although economy of expression may be increased usefully.

  Example: An organism with sensors that detect changes in pressure and changes in light intensity, and can represent co-occurrence, might find it useful to label events when pressure increases and intensity decreases, e.g. as “ZZZZ” events.

  That abbreviation does not expand what it can perceive, think about, reason about, learn about, compared with using old labels for pressure and intensity changes, with conjunction.
  But (like scientists) if it can invent **new concepts that are not definable in terms of old ones**, it may be able to express explanatory theories that were previously inexpressible (and so unthinkable), and also formulate questions that drive investigation of the environment. (Sloman, 2007)

  Such an organism, (using abductive mechanisms provided by evolution) might postulate the existence of external entities which it labels “ZZZZs” and which, when approached closely, can cause tactile pressure to increase while illumination decreases. It could then formulate tasks of finding out where ZZZZs are located, whether they move around, where they come from, what other effects they have, how big they are, whether they are edible, whether they are dangerous, etc.

  The concept of a ZZZZ would then be an exosomatic concept, unlike the somatic concepts of sensed pressure and sensed illumination, which refer only to states of the perceiver’s sensors.

- Using exosomatic concepts, allows organisms to refer to things outside them, instead of only their own sensory states and their correlations.

  Some can then develop models theories, beliefs, percepts, questions, plans, and goals referring deep into the environment, to the past, the future, remote places...

# Conjectures:

The first exosomatic ontologies were probably produced by evolution,

E.g. mobile organisms that construct primitive maps recording spatial occupancy, and organisms that manipulate small objects to learn about the objects rather than about its own sensorimotor correlations, use exosomatic ontologies.

Later on, mechanisms evolved that enabled individuals to create their own ontologies triggered by interacting with the environment, and in some cases with one another.

Are humans genetically predisposed to develop a concept of space indefinitely extended in all directions, or is that somehow a response to interacting with an ever-expanding local environment?

What mechanisms and forms of representation could enable such ontological creativity?

Compare: John McCarthy, "The well-designed child". (McCarthy, 2008) (Compare (Sloman, 2008))

Later still, ontologies were transmitted explicitly within a culture, especially to young learners – enormously speeding up ontology development.

NB: such transmission cannot all be based on explicit definition.

It may require stimulating learners to create their own explanatory theories.
(Conscious or unconscious scaffolding.)

See also (Karmiloff-Smith, 1992), and:
*The Computer Revolution in Philosophy*, 1978. Now online, revised, especially chapter 2.
`http://www.cs.bham.ac.uk/research/projects/cogaff/crp/`

# Non-trival ontology extension powers must have evolved.

concepts of kind of matter, with different properties (e.g. rigidity, flexibility (of different kinds), elasticity, viscosity, ...).

concepts of kinds of structural relationship

- topological
- metrical
- causal
- functional

Concepts of different kinds of agency

- physical forces (levers, gravity)
- chemical processes (decomposition, combustion)
- biological control (growth, repair, homeostasis)
- reactive behaviours (innate or acquired reflexes)
- deliberative capabilities (using hypothetical reasoning, searching)
- meta-semantic capabilities (referring to things that refer)
- concepts of kinds of merit, value, goodness, badness

Compare developments in computer systems engineering.

**We need careful study of actual and possible trajectories and changes in competence, in biological evolution; along with changes in individuals.**

# Some provisional, partial, conclusions

It's important to consider many different sorts of animal competence, including competence-changing competences.

- Some are closely related to the structure of the immediate physical environment.
- As organisms become more complex, the requirements for change depend on
  **both** their existing morphology and competences
  **and** the environment (which may include other information-processors)
- The features of virtual machinery that make design, monitoring, control debugging more tractable for human engineers, may also have influenced biological virtual machinery.
- Self-monitoring, reflective, machines may reach incorrect conclusions about what they are and how they work (e.g. reinventing old philosophical mind/body theories)
- There's a large amount of research on embodied cognition that assumes all intelligence, including learning, is concerned with patterns in sensory-motor signals: this ignores the need for some organisms and robots to use exosomatic ontologies.
- How to control development of exosomatic ontologies is a hard problem: there may be evolved constraints that limit the environments in which we can do this.
  "Symbol grounding" (concept empiricism) is a serious error, refuted by Kant, and philosophers of science.
- The ability of older individuals to help younger ones learn can speed up evolution – cultural evolution. Preceded in evolution by self-debugging and self-extension?
- Humans (and some others) have forms of development not yet modelled in AI, e.g. precursors of mathematical competences. (I have some rough ideas)

# Morphogenesis: types of change

For any of the above biological changes B1, B2, etc. and for any environmental change E, there can be influences of the forms

- E changes B

  Some stable structures (e.g. food sources) in the environment make controlled locomotion, exploratory movements and spatial memory useful – e.g. acquiring and using a terrain map

- B changes E

  New microbes can change earth's atmosphere, providing opportunities for lungs to be useful. Animals consuming plants can produce deserts, changing requirements for viability – e.g. finding buried food.

- Bi changes Bj

  Evolution of articulated manipulators provides opportunities for evolution of new perceptual and control mechanisms – e.g. using information about positions and relations of effectors and other objects.

- Ei changes Ej

  Volcanic eruptions can lead to cloud blankets shutting out solar radiation, leading to glaciation ....

## Mutually orchestrated parallel influences

Combinations of Ei, Bj, ... influence changes in others

- in evolution (across generations – changes in niches, and in designs – genomes)
- in individuals (over time)
- in societies
- in ecosystems

Understanding and modelling these concurrent mutually interacting changes and their influences on evolution of new forms of information-processing is a long term challenge.

# Meta-Morphogenesis (MM)

**Things that cause changes can produce new things that cause changes – eventually creating huge differences from early examples.**

## Old goals may be achieved in new ways

Effects can include both new kinds of information acquired and new ways of acquiring and using information: contents, functions and mechanisms can change.

(Try to think of examples of each kind of influence.)

## New information-processing mechanisms can produce new biological phenomena

Organisms that can monitor and modulate their own information processing, in new ways

Contrast feedback loops where information is transient and used transiently (e.g. homeostasis, servo-control) with building re-usable information records (e.g. records of structural and causal relationships and changes)

Examples needed of

Organisms that can discover what they have learnt, and use that to change some of their reasoning.

Organisms that discover that they have made discoveries

Organisms that can communicate what they have discovered

Organisms that can debug their conclusions, their reasoning, their forms of representation, their mis-uses of information...

Organisms that make and use predictions about new phenomena. (Craik, 1943; Karmiloff-Smith, 1992)

Organisms that can reason about perception, inference, and learning in others. (Apperly, 2010)

Instead of “organisms” substitute “societies” or “cultures”:

Compare the history of information-processing science and technology especially over the last century. (Dyson, 1997)

# The need for careful study of many examples

I'll show some examples of transitions that include major reorganisation of some body of knowledge, introducing new concepts and theories that could not be derived from the transport mechanisms used.

Some of these reorganisations seem to be examples of what Annette Karmiloff-Smith called "Representational Redescription". See Karmiloff-Smith (1992).

`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/beyond-modularity.html`

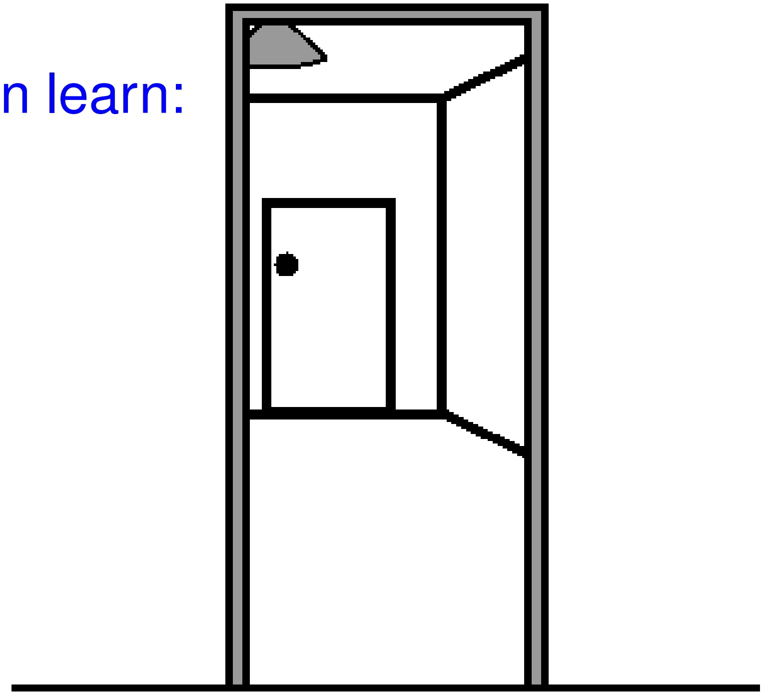
Some examples of discovery involving something like "Representational redescription" follow.

It's not clear when this happens in humans – perhaps far more often than anyone realises.

This could be part of the explanation of human mathematical competences.

# Getting information about the world from the world

## Toddler theorems:
## Things you probably know – and children can learn:



- You can get new information about the contents of a room from outside an open doorway

  (a) if you move closer to the doorway,

  (b) if you keep your distance but move sideways.

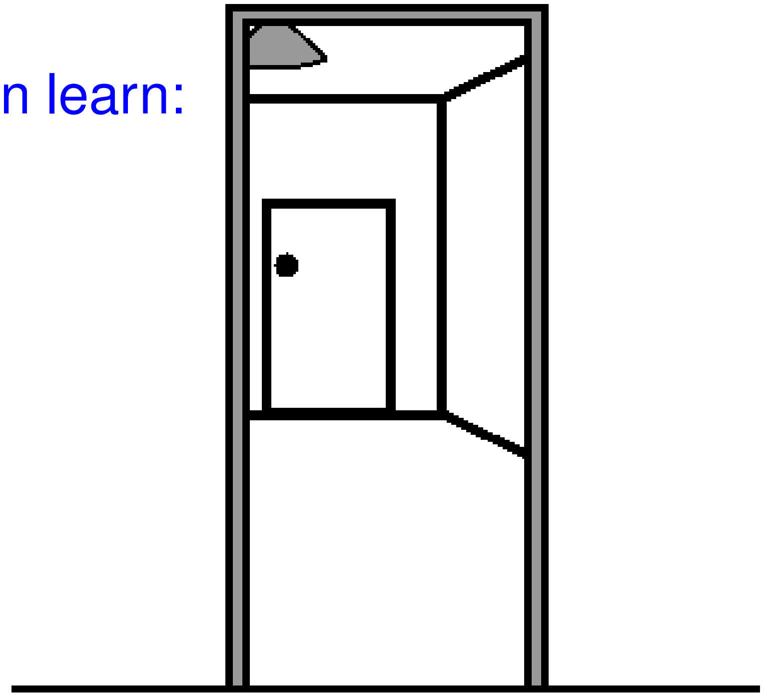
  Why do those procedures work? How do they differ?

  **(Theory: visual information travels in straight lines.)**

There's lots more...

# Getting information about the world from the world

## Toddler theorems:
## Things you probably know – and children can learn:



- You can get new information about the contents of a room from outside an open doorway

  (a) if you move closer to the doorway,

  (b) if you keep your distance but move sideways.

  Why do those procedures work? How do they differ?

  **(Theory: visual information travels in straight lines.)**

- Why do perceived aspect-ratios of visible objects change as you change your viewpoint?

  A circle acquires an elliptical appearance, with changing ratio of lengths of major/minor axes.

  A rectangle appears like a parallelogram or a trapezium.

- Why do you see different parts of an object as you move round it?
- What happens to the information available to you if an object in front of you rotates?
- What happens to information available about an object
  - If you place it inside an open box with the box facing upwards?
  - If you then place that inside an open box with the box facing upwards?
  - If you place an open box upside down over it?
  - If you then place all the previous items inside a box open at the top? (Why try that?)

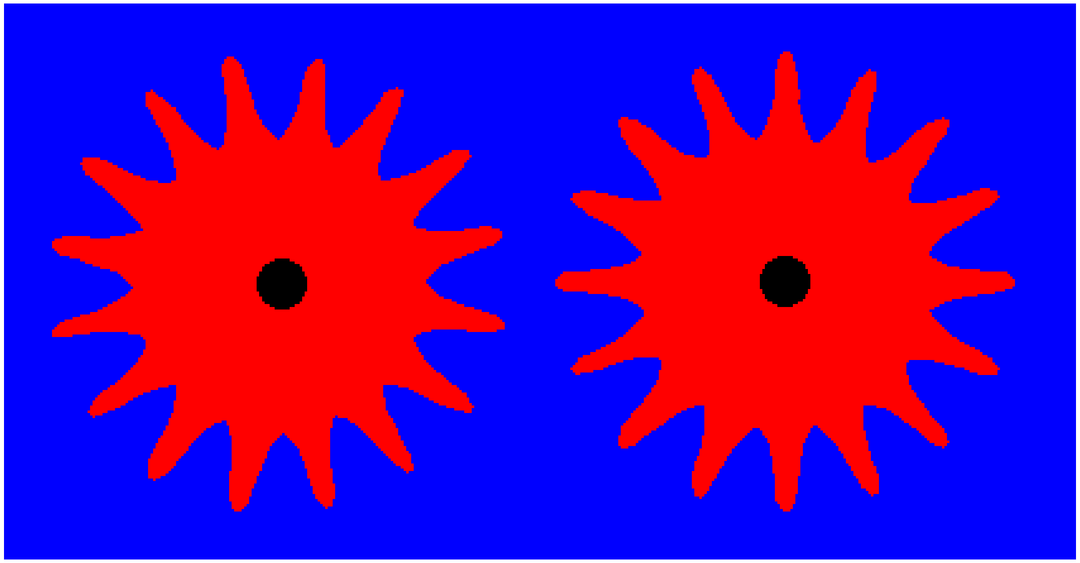
These are special cases of questions about affordances, which Gibson might have asked.

# Testing understanding of gears

## Kinds of Causation: (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?

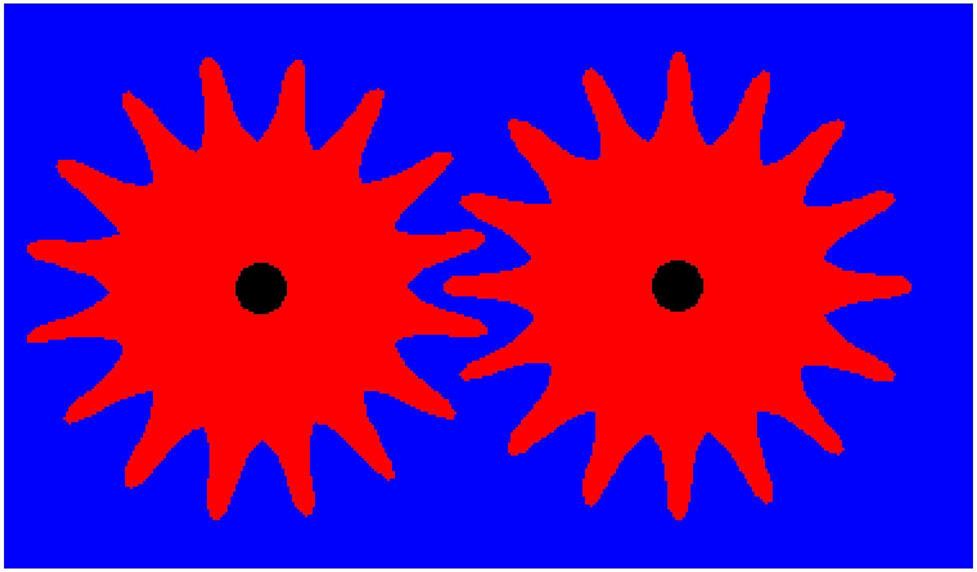


- You can tell by experimenting: you may or may not discover a correlation – depending on what is inside the box.
- In more complex cases there might be a knob or lever on the box, and you might discover a dependence: the position of the knob or lever determines how the first wheel's rotation affects the second wheel's rotation.
  (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.

# Meshed gear wheels are different

## Kinds of Causation: 2 (Kantian)

Two more gear wheels:



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

How? You can simulate rotations and observe the consequences.
(Making assumptions about the kind of stuff the wheels are made of: rigid and impenetrable)

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 the first tooth continues moving, it will push the other in the same direction, causing its wheel to rotate.

(I am not claiming that children need to reason verbally like this, consciously or unconsciously.)

NB: The simulations that you run can make use of not just perceived shape, but also unperceived constraints: in this case rigidity and impenetrability.

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

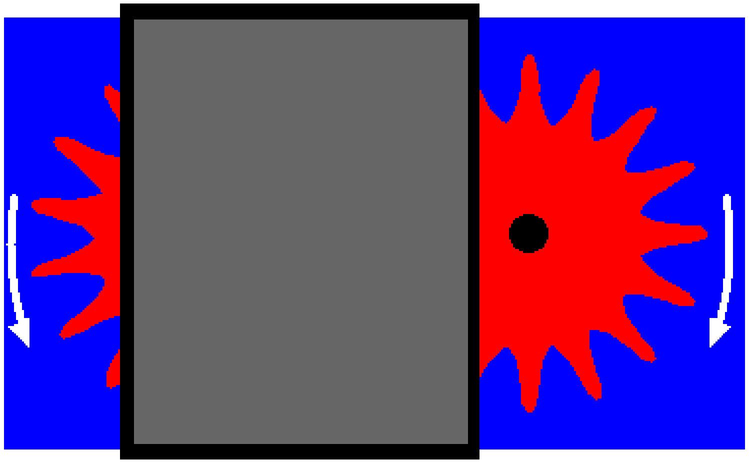
The constraints and processes using them need not be conscious, or expressed in linguistic or logical form: how all this works remains to be explained.

# Reasoning about partially observable mechanisms

## Kantian causal explanations

Reasoning that is used to predict what will be observed can also be used to explain what is observed, if mechanisms are partly hidden.

A similar kind of reasoning can be used to explain observed motion: e.g. if the left wheel is moved as shown by the left arrow, why does the right wheel move as shown by the right arrow?



There is no way to infer the correct explanation by reasoning from what is observed,
But if an appropriate theory can be constructed, and added as a "hidden premiss" (abductive reasoning) then it can be tested by working out its consequences and comparing them with what is observed.

Humean (nowadays Bayesian) causal reasoning based on statistical information about perceived correlations, is the only kind of causal reasoning available **when mechanisms are not understood.**

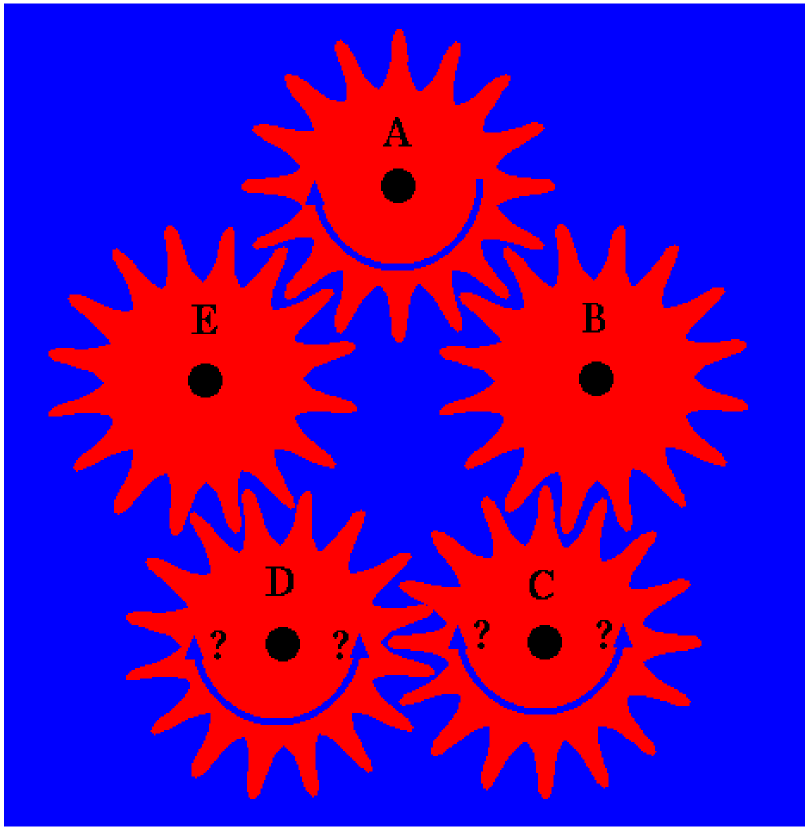
Kantian (structure-based, deterministic) causal reasoning is sometimes available, and is often more useful, but it requires a richer ontology and more general reasoning abilities that can make direct use of structural constraints in complex configurations.

# Testing a child's ability to do such reasoning

Perhaps the following configuration could be used to test whether a child had moved beyond the empirical understanding of gear-wheel interaction.

**If wheel A moves as shown,**
**which way will wheels C and D turn?**

**This can test adults too!**



# Sorting problem (investigated by Piaget)

How can a child learn about ways of producing ordered structures, e.g. arranging blocks in order of increasing or decreasing height?

Various stages

Partial successes and partial failures can trigger changes – but only if the right information processing mechanisms for self-monitoring and debugging are available.

What's involved in noticing

1. that you have made a mistake
2. that the mistake can be corrected
3. that there was information available at the time the mistake was made that could have been used but was not used
4. that there is a way of proceeding in future that uses such information?

# Gain some first hand experience of toddlerhood I

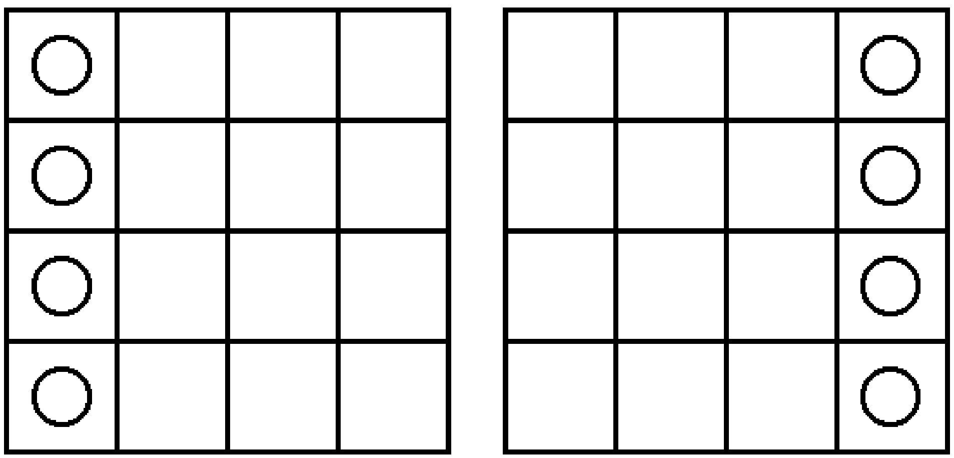
## See what you can learn by thinking about these two puzzles.

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

**Puzzle 1: A warm-up exercise:**

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

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



What is the minumum number of moves required?

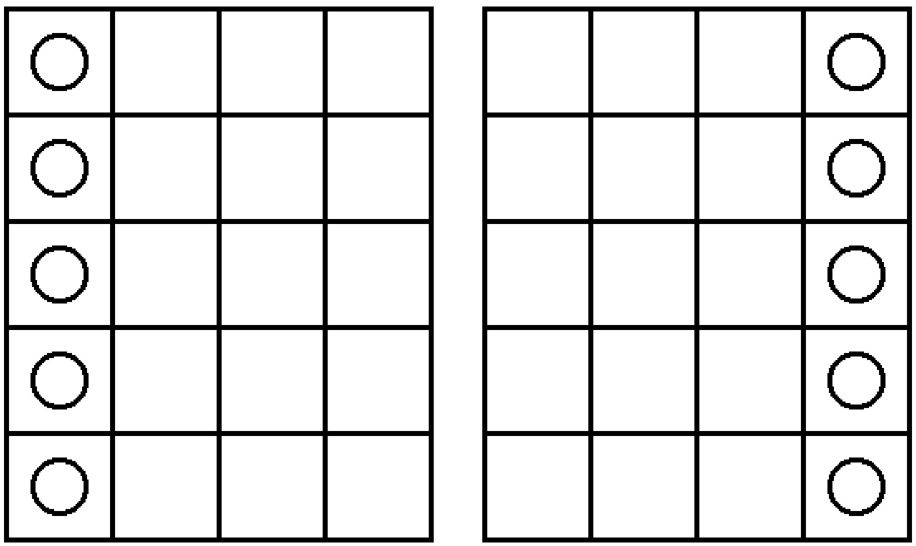
Note that the first question is about what is possible (a solution), and the second is about what is impossible (a solution shorter than ....)

# Gain some first hand experience of toddlerhood II

**Puzzle 2 Now stretch yourself:**

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

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



What is the minumum number of moves required?

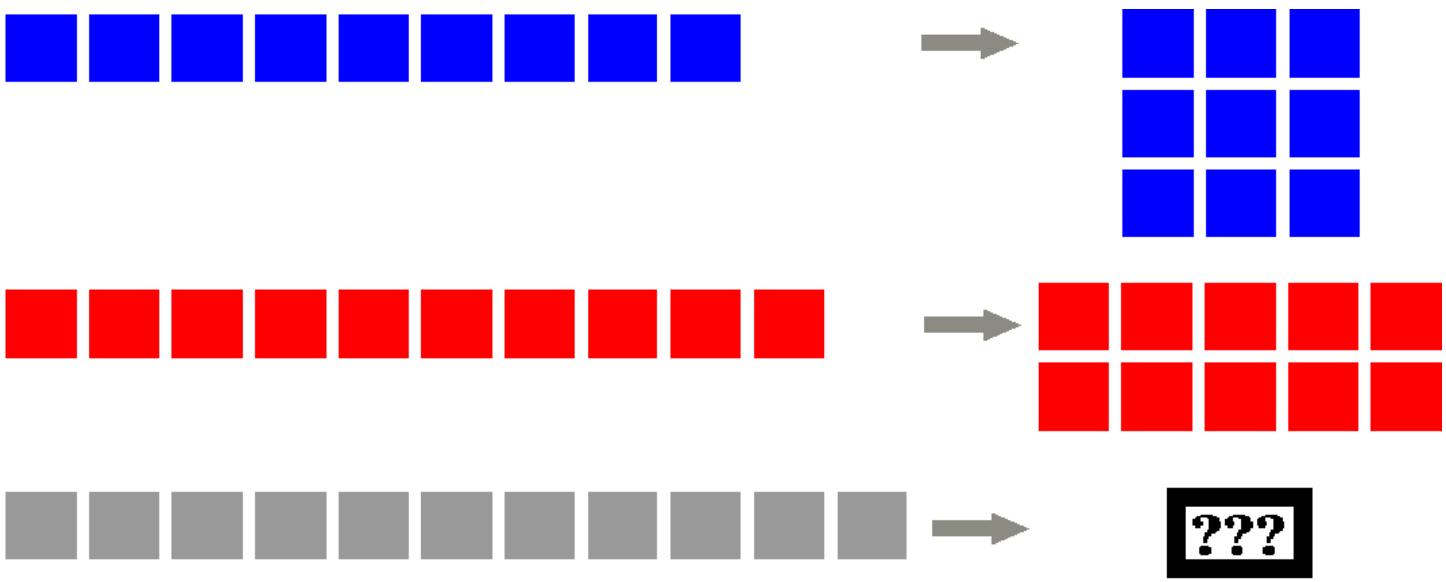
How people work on these puzzles varies according to prior knowledge and experience, and their ability to deploy knowledge they already have.

Even mathematicians sometimes kick-themselves for failing to notice the simple, elegant solution that a child could discover. (Though not all will.)

# The ubiquity of prime numbers

Playing with a collection of cubes of the same size you can try forming patterns with them.

Sometimes you'll fail: Why?




How can a child become convinced that something is impossible?

There are very many more examples.

NB: don't confuse learning about the 'natural numbers' with learning about numerosity of perceived collections – a far shallower competence.

# How can the general concept of a measure applicable to different entities be acquired??

Number of discrete objects?

Length of a straight line?

Length of a curved line? (Arbitrary curves?)

Area of a rectangle?

Area of a shape not made up of rectangles. E.g. a circle?

Area of an arbitrary shape?

Volume of a cubic box?

Volume of a cylinder?

Volume of an arbitrary shape?

How many normal adults understand these concepts?

How many who don't understand them have the ability to understand them?

# Rates of change

Lots more to be said.

How can you come to understand that it is possible for a moving object's speed to be decreasing while its acceleration is increasing?

# Opening and shutting

**Non-epistemic affordances concerned with opening and shutting micro-worlds:**

- In order to shut a door, why do you sometimes need to push it, sometimes to pull it?
- Why do you need a handle to pull the door shut, but not to push it shut?
- How you could use the lid of one coffee tin to open the lid of another which you cannot prise out using your fingers?
- What happens if you shut an open drawer by pressing your hands over the top edge and pushing?

# To be expanded and clarified

These slides are still incomplete.

Additional theory and examples can be found in other presentations and discussions here:

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

`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/AREADME.html`

Provisional references follow

## References


Apperly, I. (2010). *Mindreaders: The Cognitive Basis of "Theory of Mind"*. London: Psychology Press.

Brooks, F. (1996, Mar). The Computer Scientist as Toolsmith. *Communications of the ACM*, *39*(3). Available from http://www.cs.unc.edu/~brooks/Toolsmith-CACM.pdf

Craik, K. (1943). *The nature of explanation*. London, New York: Cambridge University Press.

Dyson, G. B. (1997). *Darwin Among The Machines: The Evolution Of Global Intelligence*. Reading, MA: Addison-Wesley.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.

Lovelace, A. (1842, October). *Notes upon the Memoir by the Translator of "Sketch of The Analytical Engine Invented by Charles Babbage" By L. F. Menabrea.* Available from http://www.fourmilab.ch/babbage/sketch.html

McCarthy, J. (2008). The well-designed child. *Artificial Intelligence*, *172*(18), 2003-2014. Available from http://www-formal.stanford.edu/jmc/child.html

Sauvy, J., & Sauvy, S. (1974). *The Child's Discovery of Space: From hopscotch to mazes – an introduction to intuitive topology*. Harmondsworth: Penguin Education. (Translated from the French by Pam Wells)

Sloman, A. (1978). *The computer revolution in philosophy*. Hassocks, Sussex: Harvester Press (and Humanities Press). Available from http://www.cs.bham.ac.uk/research/cogaff/crp

Sloman, A. (2006, May). *Requirements for a Fully Deliberative Architecture (Or component of an architecture)* (Research Note No. COSY-DP-0604). Birmingham, UK: School of Computer Science, University of Birmingham. Available from http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604

Sloman, A. (2007). *Why symbol-grounding is both impossible and unnecessary, and why theory-tethering is more powerful anyway.* (Research Note No. COSY-PR-0705). Birmingham, UK. Available from http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models

Sloman, A. (2008). The Well-Designed Young Mathematician. *Artificial Intelligence*, *172*(18), 2015–2034. (http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0807)

Sloman, A. (2010, August). How Virtual Machinery Can Bridge the "Explanatory Gap", In Natural and Artificial Systems. In S. Doncieux & et al. (Eds.), *Proceedings SAB 2010, LNAI 6226* (pp. 13–24). Heidelberg: Springer. Available from http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#sab

Sloman, A. (2011a, July). *Evolution of mind as a feat of computer systems engineering: Lessons from decades of development of self-monitoring virtual machinery.* Available from http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1103 (Invited talk at: Pierre Duhem Conference, Nancy, France, 19th July 2011)

Sloman, A. (2011b). What's information, for an organism or intelligent machine? How can a machine or organism mean? In G. Dodig-Crnkovic & M. Burgin (Eds.), *Information and Computation* (pp. 393–438). New Jersey: World Scientific. Available from http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#905

Sloman, A. (2011c, Sep). *What's vision for, and how does it work? From Marr (and earlier) to Gibson and Beyond.* Available from http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk93 (Online tutorial presentation, also at http://www.slideshare.net/asloman/)

Sloman, A., & Chrisley, R. (2003). Virtual machines and consciousness. *Journal of Consciousness Studies*, *10*(4-5), 113–172. Available from http://www.cs.bham.ac.uk/research/projects/cogaff/03.html#200302