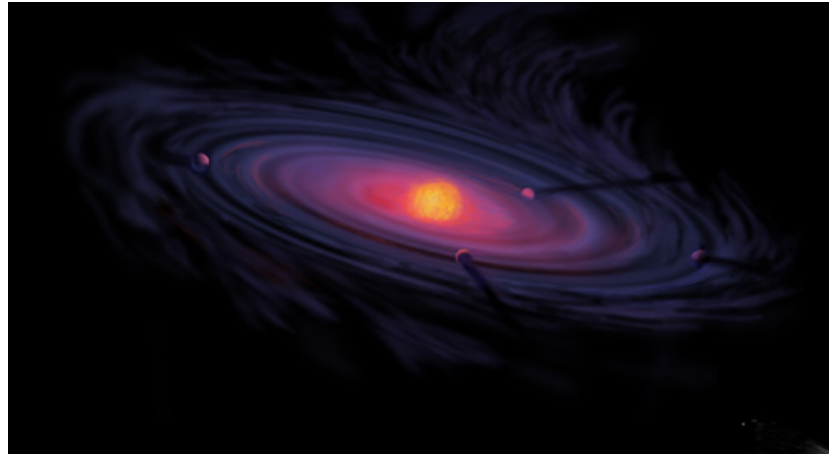


Meta-Morphogenesis theory as background to Cognitive Robotics and Developmental Cognitive Science

How could our minds and the rest of life have come from a cloud of dust?



Artist's concept of a protoplanetary disk (NASA – Wikimedia)

Notes for Dagstuhl Seminar 10-15 Feb 2013

Mechanisms Of Ongoing Development in Cognitive Robotics

Closely related to tutorial at AGI 2012, Oxford December 2012

Video of AGI Tutorial <http://www.youtube.com/watch?v=BNu152kFI74> (Fixed 14 June 2012)

Aaron Sloman

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

School of Computer Science, University of Birmingham.

More Information

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/dagstuhl-13.html>

These slides also serve as post-presentation slides for a talk at
The Bath Royal Literary and Scientific Institution

<http://www.brlsi.org/>

16 Queen Square, Bath, BA1 2HN

On Tuesday Evening 2nd April 2013

Advertised as:

The Meta-Morphogenesis Project

(Partly inspired by Turing's work on morphogenesis)

<http://tinyurl.com/CogMisc/misc/meta-morphogenesis.html>

**How could our minds and the rest of life
have come from a cloud of dust?**

**Including: Microbes, mice, monkeys, mathematics, music, marmite
(along with murder, mrsa, religious bigotry, and other nastiness).**

Original Abstract here

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/bath-brlsi-talk.html>

Many additional presentations related to this are here:

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

Related discussion notes here:

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

High Level Overview

- Since their beginnings, there has been great progress in some areas of AI and Cognitive Science, and abysmal progress in others.
- That's because there are some very deep problems about animal intelligence that have not been solved, and some have not even been noticed by most researchers.
- These include problems connected with human mathematical competences (e.g. in geometry) and problem solving competences in other animals.
- There are also some allegedly hard problems that are actually not so hard – for people who have understood advances in virtual machinery, e.g. problems about the evolution, and implementation of qualia and various ways of being self conscious.
- I'll suggest a (Turing-inspired) strategy for trying to clarify the problems and gain ideas about the solutions produced by evolution that so far surpass anything we have in AI/Robotics.
- The strategy is to attempt to identify and explain **transitions** in the evolutionary history of **biological information-processing** – since microbes – since it is possible that current animals, including humans, still make important use of solutions to old problems, in ways that would not occur to us starting with computers.
- I call this the Meta-Morphogenesis project – partly inspired by reading Turing's paper on Morphogenesis (1952), and partly because mechanisms that produce new mechanisms, can also produce new mechanisms for producing new mechanisms!
- I shall have to select and over-simplify because the topic is far too large for a single talk. I introduce a project, not results, and this is just a sample of the scope of the project.
- **The focus is not physical morphology, or behaviour, but **information-processing**.**

Video Demos Shown

1. BigDog (Boston Dynamics) circa 2010.

Impressive robot with no understanding of what it has done, what it could have done, what it did not do, why it did something, what it may have to do after it has reached a new location, etc., what the consequences of various actions could be, or will be, etc.

2. Betty the New Caledonian Crow

Betty made hooks from straight pieces of wire in order to get food in a bucket out of a tall tube. She used several very different procedures for making the hooks, without any evidence of random exploration of behaviours.

The videos on the Oxford Ecology Lab web site indicate that she knows exactly what she is going to do before she does it.

3. Little Child

A child pushing a broom knows how to move the handle to disengage it from railings and chooses to steer it down a corridor. Before reaching a door on the right he changes the orientation of the broom so that when he gets to the door the broom is pointing into the room and he then gets behind it and pushes it in.

This is one of a collection of videos with comments here

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/vid>
(Search for "broom".)

Too little progress – can a new approach help??

My motivation: to understand **natural** systems (not make new systems)

Making useful artefacts may or may not be a side-effect.

There are lots of AI/Robot systems, but barely up to insect intelligence, **or** very narrow.

Some very powerful reasoning, theorem-proving, planning systems but problems have to be posed in their notation, using their ontology.

Some surpass humans in specialised domains with narrow ontologies and limited states and processes, e.g. playing chess, proving logical or algebraic theorems, checking program correctness.

Others are very powerful narrowly focused AI tools fed on huge amounts of symbolic information – more than any human encounters in a lifetime E.g. IBM's Watson – beats human champions at Jeopardy.

AI systems easily out-perform humans in a large collection of tasks, **BUT:**
Cannot yet match performances of humans, squirrels, crows, etc.

Cannot do very simple(?) reasoning leading up to Euclidean geometry

Can't match the processing of affordances in many animals – not just primates.

Robots are getting better at **online** intelligence (e.g. walking, catching balls...).

Contrast **offline** intelligence (e.g. reasoning about walking, catching balls)

Finding something funny... enjoying mathematics (only humans, so far??)

If we can't make robots that do what animals can do, perhaps we are not using good hypotheses to generate AI systems.

Perhaps we have not understood what needs to be explained?

Perhaps we have not understood the variety of mechanisms possible? **Chemistry?**

The very hard questions are about information

There are hard problems about how physical bodies evolved

e.g. muscular power-weight ratios, various strength-size ratios (bones, tendons, etc), numbers of sensors, power of sensors, numbers of neurons, etc.

But normal bodies are not essential: humans born with seriously deficient bodies can grow up to be **obviously** humans.

E.g. The limbless artist Alison Lapper, people born blind, people born deaf, etc.

(The role of embodiment has recently been much over-rated and misunderstood. (Sloman, 2009b))

The really hard problems are concerned with **information-processing**.

- What information – what is it about? (**Sensorymotor signals vs environment**)
- Where does it come from? (**genes? sensors? learning? culture? ... How?**)
- How is it processed?
- How is it interpreted?
- How are results stored?
- How is the information used?
- What forms of representation are required? What ontologies?
- What mechanisms are required?

What are the different sorts of information processing mechanisms we know about?

Could we be missing something?

How can we find out?

How has biological information processing changed since the earth formed?

Ignore debates about life/non-life boundaries

I don't find it fruitful to argue about where or how to draw the boundary between non-living and living things.

- However it is useful to look for various kinds of discontinuity between extreme cases, e.g. between the extreme of isolated atoms or molecules and the extreme of mammalian life, human and non-human – there are many discontinuities.
- It is impossible for matter to change **continuously** between those extremes, For example, very many intermediate chemical changes are involved, and chemical changes are discontinuous, since it is not possible to add an arbitrarily small portion of an atom to a molecule to produce a larger molecule.
- The Meta-Morphogenesis project is mainly concerned with changes concerned with information processing, as described below.

There are much older traditions – search for changes in physical form or structure, and changes in physical behaviours.

In comparison, changes in information processing are mostly invisible, or very hard to observe, though their consequences can be very visible.
- Finding out what the possible transitions are, including changes in I-P, and what their implications are, is far more important than arguing about where to locate one transition between life and non-life.
- The interesting attempt by Tibor Gánti (2003) to list three minimal conditions for life can be subsumed within this project by discussing his conditions and their implications.

I am very grateful to Jan-Ulrich Kreft for drawing my attention to this excellent book. A useful summary by Gert Korthof can be found here: <http://wasdarwinwrong.com/korthof66.htm>

Gánti mentions information and information-processing but does not explore the varieties of information-processing addressed in this project – e.g. origins of mathematical reasoning.

Problems about natural information-processing

- Where/how did all the information originate?
- Where/how did all the information-processing machinery originate?
- Why is chemistry so important for biological information processing?
- Natural selection cannot be the whole story: something has to be said about what it operates on.
 - E.g. could it have worked on a giant ball in space made of grains of sand held together by gravity?
I suspect not enough variety of structure and process would be possible.
- Natural selection emphasises **competition**, and under-emphasises **interdependence** and **cooperation**: e.g. large animal species depend on massive food pyramids.
 - The role of symbiosis and co-evolution needs to be considered.
- Could something similar to biological evolution have happened in a giant Turing Machine set up millions of years ago (with external power source)?
- Or a large collection of interacting Turing Machines?
 - If not, why not?
 - Note: the success in some application domains of “evolutionary computation” (e.g. genetic algorithms, genetic programming (GA/GP) can mislead some people into thinking they model natural evolution.

Key questions:

What's information?

What roles did information play in evolution?

(Forget about Shannon for now.)

How long have we been processing information?

Jane Austen knew a lot about human information processing, as these snippets from *Pride and Prejudice* (1813) show:

She was a woman of mean understanding, little information, and uncertain temper.

You could not have met with a person more capable of giving you certain information on that head than myself, for I have been connected with his family in a particular manner from my infancy.

She then read the first sentence aloud, which comprised the information of their having just resolved to follow their brother to town directly,...

She resolved to give her the information herself, and therefore charged —, when he returned to — to dinner, to drop no hint of what had passed before any of the family.

Mrs. — about this time reminded — of her promise concerning that gentleman, and required information;

He could tell her nothing new of the wonders of his presentation and knighthood; and his civilities were worn out, like his information.

I was first made acquainted, by —'s accidental information, that —'s attentions to your sister had given rise to a general expectation of their marriage.

As to his real character, had information been in her power, she had never felt a wish of inquiring.

When he was gone, they were certain at least of receiving constant information of what was going on,

But to live in ignorance on such a point was impossible; or at least it was impossible not to try for information.

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

With thanks to Project Gutenberg:

<http://www.gutenberg.org/files/1342/1342-h/1342-h.htm>

Note: Jane Austen does not use the word “information” in Shannon’s sense. Neither do I. Shannon, unlike many of his admirers, was well aware of the difference.

Inspiration from Alan Turing: Morphogenesis

A. M. Turing,

[The Chemical Basis Of Morphogenesis](#),

Phil. Trans. R. Soc. London B 237, pp. 37–72, 1952,

Turing showed how relatively simple processes of reaction and diffusion of two different chemicals in a fluid could give rise to patterns found in living things, e.g. blotches, spots, stripes, spirals.

Can we go beyond explaining physical patterns on plants and animals produced during [individual development](#), to explain how a myriad of forms of information processing in many forms of life can come from chemistry – over [evolutionary time-scales](#)?

The latter needs more complex forms of interaction than reaction and diffusion:

Abilities of chemical structures:

- They can store and release energy
- to form enormously varied physical structures with many different properties (some supporting information processing).
- They allow mixtures of continuous deformation and discrete changes between highly stable states,
- with mixtures of rigidity and flexibility (e.g. foldable DNA).
- A process can trigger others, e.g. catalytic and auto-catalytic processes (Kauffman, 1995)
- Very small scales, very fast, very low energy consumption, ...
- but chemistry can also build much larger scale structures than individual atoms.
- And chemical states and processes can control chemical states and processes (Ganti, 2003).

What would Turing have done if he had lived longer?

Tibor Gánti's important ideas

I did not know about the extremely important work of **Tibor Gánti** until Jan-Ulrich Kreft lent me his book, in 2013: *The Principles of life* (2003)

Although I don't share Gánti's interest in producing a *definition* to correspond to a pre-theoretic notion of "life" I do share his interest in the different kinds of physical/chemical mechanisms that can exist and the implications of the differences.

In particular the book specifies the concept "Chemoton" (Chemical Automaton) which he offers as a specification for a minimal form of life.

Instances of of the Chemoton specification must have three types of functioning that Gánti shows can be supported by mechanisms based on chemical structures and processes suspended in a chemical soup.

A chemoton is a bounded system capable of self-maintenance, proliferation (reproduction), and evolution (changes occurring in the reproductive process) supported by:

1. Genetic material that can undergo template replication, leading to replication of the whole individual.
2. A metabolic system that maintains the functioning system using materials entering through the membrane.
3. A (possibly growing) bounding membrane (semi-permeable).

Gánti shows how chemical processes based on autocatalytic mechanisms can co-operate in the functioning of such a system, including maintenance, growth and reproduction.

He stresses the importance of information and information-processing, but considers only a small subset of roles for information in life, compared with the roles we discuss.

For a longer summary see the enthusiastic review by Gert Korthof:

<http://wasdarwinwrong.com/korthof66.htm>

All organisms are information-processors but the information to be processed, the uses of the information, and the means of processing vary enormously

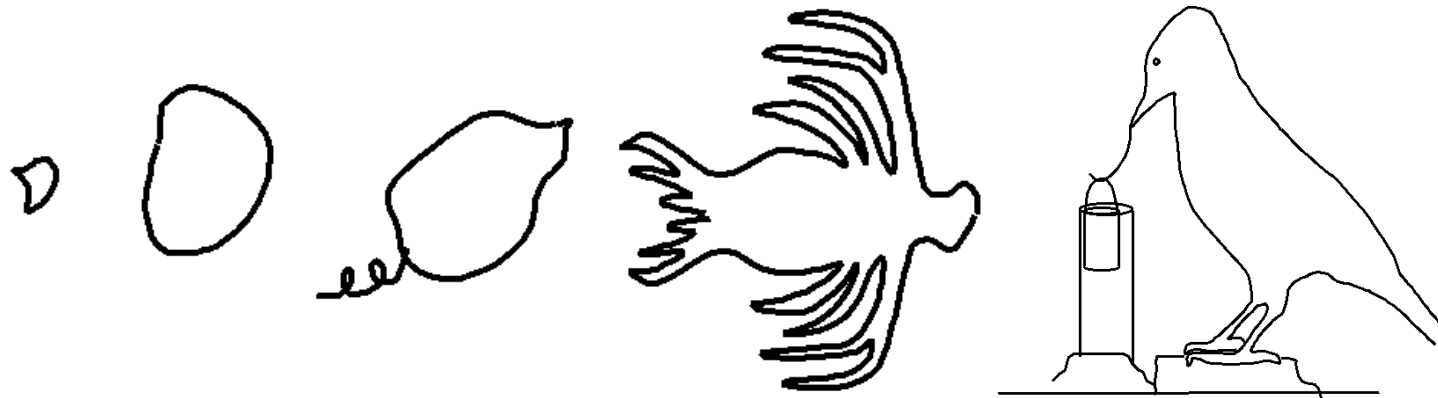
A key fact about (pre-Shannon) information is that it **has uses that depend on the information content**, the most basic use being **control**. (See later.)

Moreover, some uses of information are **internal** and very hard to detect.

Others are externally visible: including use of information about opportunities for behaviour.

Some involve other agents: collaborating, competing, helping, hindering, feeding, protecting ...

Evolutionary and developmental changes in **size, shape, sensing mechanisms, effector mechanisms**, all provide new I-P opportunities, as do changes in **environments, including new prey, predators, symbionts...**

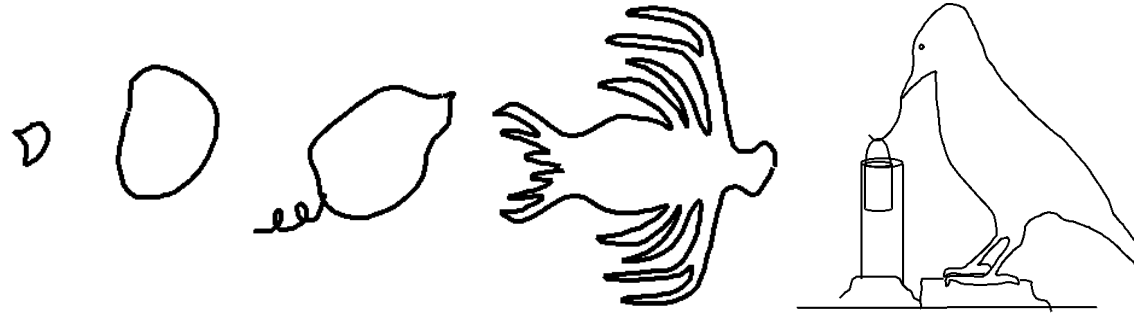


Look up Betty, the hook-making New-Caledonian crow:
Publications, photos, videos at Oxford ecology lab. (Circa 2002-5).

http://users.ox.ac.uk/~kgroup/tools/crow_photos.shtml

<http://users.ox.ac.uk/~kgroup/tools/movies.shtml>

Opportunities, means and obstacles to information-processing, change as environments change



Types of environment with different I-P requirements/possibilities:

- Microbes in a chemical soup with nutrients and other contents [See later slide for alternative origins.]
- Soup with detectable gradients
- Soup plus some stable structures (places with good stuff, bad stuff, obstacles, supports, shelters)
- From soup (sea, lakes, ...) to land, foliage, air, ... new information new means of locomotion.
- Things that have to be manipulated to be eaten (e.g. disassembled).
- Controllable manipulators (e.g. mouths, claws – some providing diverse affordances, and problems)
- Things that need to be built, maintained, and/or repaired (e.g. nests)
- Things that try to eat you
- Food that tries to escape
- Mates with preferences
- Competitors for food and mates
- Collaborators that need, or can supply, information.
- and so on **(Try to extend this list!)**

Identifying I-P transitions can help us understand what biological evolution has achieved, providing clues to mechanisms and architectures, serving changing needs.

Influences of different initial environments

Early organisms attached to solids will have different (but possibly overlapping) information-processing requirements from life starting in chemical soups: could the differences have long term implications?

Options for control of motion (small scale and large scale) for organisms in a chemical soup are different from options for action if located not in soups but in cracks, crevasses, and hollows in solid material, with fixed biomembranes separating them from the rest of their environment.

Both mobile and static individuals may need to control temperature, osmotic transfer, nutrient capture, waste excretion, etc., and differing options for actions imply different opportunities and requirements for information-processing in the control of actions.

The I-P requirements for individuals attached to solids would be very different from the requirements for individuals enclosed in a membrane in a chemical soup.

E.g. initially there would be no possibility for control of motion, avoidance of obstacles, and no opportunity to acquire and use information about the spatial layout of terrain beyond immediate reach of the organism.

Later, growth of flexible, controllable, appendages might provide new I-P needs and opportunities.

Differences in earliest opportunities for motion, and for acquisition and use of information, may have implications for differences in later information-processing mechanisms (and their transitions) – unless all evolutionary routes from the fixed entities to later more complex mobile organisms pass through free-floating membrane-enclosed stages.

This page needs to be re-written and expanded, indicating more clearly some of the differences between kinds of information, and kinds of processing, useful for early pre-biological or biological entities immersed in a soup and those attached to solid structures. We need a better overview of design options and tradeoffs. See (Sloman, 2000)

Changing, evolved uses of information

Organisms (and robots) need information both for **controlling interactions** with things in the environment, and also for **other** purposes, e.g. building explanatory theories, trying to understand the environment (including other information users), teaching others, creating value systems, creating new goals, rejecting goals, inventing stories, jokes and art, ...

Important transitions in I-P functions and capabilities include

- Changes in ontology
 - new components
 - new modes of composition (there are deep unsolved problems about this)
- Changes in forms of representation/forms of encoding (including syntax)
Yes/No categories, Sets of descriptive labels, Partial and total orderings, Measures (1-D, 2-D, N-D),
Use of relations (topological, metrical, structural, causal, functional...), Use of analogical representations
e.g. maps, abstract pictures, Use of varied forms of syntax – logical, pictorial, temporal composition, etc.
- Changes in theory-building and testing capabilities.
- New I-P mechanisms and I-P architectures
algorithms, representation manipulators, storage, inference, computational paradigms, ways of
combining mechanisms and kinds of functionality into larger systems – **originally all chemical!**
- New forms of control (On/Off, continuous, parallel, use of plans, use of tools, collaboration, etc.)
- New types of bootstrapping from birth/hatching or earlier (precocial/altricial tradeoffs).
- Changes in types of environment coped with.
- Changes in information acquired or used in old and new environments.
- New types of communication, collaboration, competition, peaceful co-habitation, etc.
with other information processors (e.g. tickbirds and large mammals).

AI/AGI/CogSci/NeuroSci models are still **very** limited, compared with later organisms.

Note on “information”, “meaning”, “language”

For the purposes of this investigation it is useful to generalise the notion of meaningful, useful, language to include the forms of representation (means of encoding information) used **internally** in many organisms, and increasingly in intelligent machines.

Restricting “language” to refer only to **external** communications is foolish because it

- (a) ignores the important non-communicative **internal** roles of human languages in thinking, reasoning, problem-solving, planning, asking, hypothesising, imagining, deciding, intending, within an individual
- (b) ignores the commonality between (i) requirements for information to be communicated **between** individuals and (ii) requirements for information to be acquired, stored, transformed, combined, and used **within one** individual (requirements include generativity i.e. ability to express novel meanings, (Chomsky, 1965), compositionality, structural variability, ability to refer to nearby and remote things and events, etc.)
- (c) ignores the rich and varied types of information-processing done internally by members of many other species (corvids, orangutans, elephants, hunting mammals, squirrels) and pre-verbal humans – discussed further below.
- (c) restricts questions about **evolution** of language to a narrow, misleading, focus on **external** languages.

Human communicative languages would be impossible without the **internal languages** that evolved earlier in our evolutionary precursors, for perception, learning, etc.

Those precursors are required in internal mechanisms for the use of external languages, e.g. working out sentence structures, encoding of syntactic rules, etc. (Sloman, 1979, 2008a) (Sloman & Chappell, 2007)

The original type of semantic content in organisms seems to have been **control-content**, i.e. specifying what to do. Other kinds, referring to states, events, processes, objects, possibilities, problems, goals, plans, etc. may have grown out of this. Some indications of that process are given in later slides, but there are still many gaps. (Sloman, 1985)

(This is not to be confused with Fodor’s notion of a “Language of Thought” (Fodor, 1975) since internal languages develop, whereas his LoT is supposed to be fixed.)

Aspects of information use

Different roles in use of information can be distinguished.

A user **U** can use an item of information **I** from source **S**, where:

- **S** may be an intelligent agent intending to communicate the content **I** to someone or something,
- **S** may be a physical object or process, providing information **I** about itself e.g. information about visible features of a tree-trunk: its width, curvature, texture, height, etc. – information about the tree's features is **provided by the tree**, without the tree having any intention in the matter: “self-documentation”.

Information **I** acquired by user **U** has a carrier **C**, or a sequence of carriers, some inside **U**, e.g. **C1** (object surface), **C2** (reflected light), **C3** (retinal activation), **C4** (neural transmission), **C5**, **C6**, ... **Cn**, etc. (different brain structures or activation patterns).

Carriers may be printed text, spoken sounds, morse code, retinal activity, a transient or enduring neural structure, a complex molecule or anything else that **U** can map onto meanings. If the information content **I** is complex (e.g. has parts and relationships) then **C** will normally need complexity, i.e. parts and relationships, to express **I**, unless abbreviations are used. (When/how did use of abbreviation evolve?)

If information items **I1**, and **I2**, have carriers **C1** and **C2**, the structures or contents of **C1** and **C2** will differ in ways that **U** interprets as expressing **I1** and **I2** (but there can also be overlap of content).

Some differences are discrete (e.g. where **C1** and **C2** are assembled from a fixed set of parts), others continuous, e.g. continuously changing signal **C** indicates changing size, or distance, or, curvature, etc. in **I**.

Sometimes **context** of use of **C** contributes to information **I** in a systematic way.

Components of **C** may be EITHER arbitrary items juxtaposed to express content **I** in a systematic way, like words in a sentence, OR non-arbitrary tokens (e.g. of varying size, or shape) e.g. parts of pictures or maps.

A **systematic** relationship between **C** and **I** does not imply isomorphism of **C** and **I**: counter-examples include (i) Fregean representations denoting something whose structure is totally unlike the encoding), (ii) **analogical** representations where the semantic relationship is context sensitive, e.g. parts of a 2-D picture representing parts of a 3-D scene, where picture and scene have very different structures (Sloman, 1971).

If the carriers, **C1**, **C2**, ... etc. form a system allowing novel carriers to be constructed to express novel information, we can call the system a **generative** language, possibly **infinitely** generative.

I.T. can limit or muddle thinking about information

The ever increasing role of IT (Information Technology) in our lives has led some people to think that by definition “information” refers to the bit patterns in computers (or possibly their printed forms, e.g. machine code programs, lists of numbers, tables, etc.) which are thought of as meaningless.

An earlier slide demonstrates that Jane Austen, writing just over 200 years ago, knowing nothing about computers or Shannon, was very much at ease with the notion of “information” that I am using, though I’ll show that its scope is wider than she dreamed.

Some engineers ignore that ordinary use of “information” and introduce a dreadfully confused distinction between **data** (e.g. bit patterns, marks on paper, sensor readings) and **information**

The contrast is defined in different ways by different people, often confusing information-bearers and information-sources. (Search for “ ‘data’ vs ‘information’ ”.)

They forget that long before there were digital computers, humans acquired, passed on, distorted, used, believed, disputed, rejected, recorded, combined, filtered (etc.) information from many sources, on many topics. (Like Jane Austen’s characters.) Other animals also use information in that sense.

Acquiring information can expand what you are informed about (but not if it’s redundant).

In that sense “information” and “meaning” are deeply related, and should not be **contrasted** as is sometimes done by people who know only the restricted technical notion of information.

(Note that in English “having meaning” sometimes refers to having import.)

Many English speakers use the word “information” in the sense in which it is very close to “meaning”, and I’ll go on using it in that sense in these slides.

So when I talk about information processing, and uses of information, I am referring to processing of meaning, and uses of meaning – not necessarily expressed in either sentences or bit-patterns.

I wonder whether the excessively narrow interpretation of the word “information” arises mainly in non-native speakers of English, and poorly educated native speakers, who encounter the word only in technical contexts, e.g. in connection with computers, or Shannon’s ideas.

The uses of information

Why do organisms need information? One partial answer:

Nehaniv et al. (2002) attempt to answer this in terms of different **uses** of information

<http://www.alife.org/alife8/proceedings/sub7405.pdf>

[My partial disagreements and queries are expressed in red inserts.]

“ ... we systematically relate information to utility for an organism. Meaningful information is defined here as 1) information in interaction games [Environments play games??] between an organism and its environment or between organisms mediated with respect to their own sensors and actuators and as 2) useful for satisfying homeostatic and other drives, needs, goals or intentions (Nehaniv & Dautenhahn, 1998). In particular, meaningful information need not be linguistically nor even symbolically mediated. [Yes!] It may or may not involve representations [See Note below], but must arise in the dynamics realizing the agent's functioning and interaction in its environment (cf. the notion of 'structural coupling' of (Maturana & Varela, 1992)), supporting adaptive or self-maintaining or reproductive behaviors, goals, or possibly plans. [But see varieties of dynamical system, later.] Under evolution, sensor and actuator channels used in recurring types of interaction games will over generations to some degree be optimized in order to better achieve survival and reproduction, ...” (Page 2) [But evolution is mostly a satisficer, not an optimizer! (Simon, 1969)]

They mention different time-spans for usefulness of information, and change of focus, but (apart from story-telling) seem to ignore the importance of different **spatial** scopes and the possibility of reference to remote places and their occupants, or things that are hidden, or too small to be perceived, or information about abstractions or fictions. (Compare (Sloman, 2011b))

Note: if we take “representation” to refer to the carrier **C** that **embodies** or **encodes** the information, then the suggestion that information “may not involve representations” becomes self-contradictory.

Note that **C**, the carrier or representation, may be a very abstract process pattern (Sloman, 2011a).

N.B., Information can have a role in **control** without being associated with **utility measures** (Sloman, 2009a).

We need to study many examples, to understand the breadth and depth of the problems.

Roles of information in biological and non-biological control

We can distinguish different sorts of functional roles for mechanisms involved in use of information for control, in biological and non-biological systems. A partial list:

- sensing some **state** (e.g. temperature), or **event** (change), or **process** (temperature fluctuations);
- sensory information triggers responses through **direct** physical causation, e.g. mouse-trap, use of wind to rotate a windmill to face the wind, Watt-governor, bending of bi-metallic strip in a thermostat;
- allow **indirect** connection between response selection and response production, e.g. instruction sent to a switch or throttle, or brake;
- select response to a state on the basis of other state information (conditional responses);
- perform analysis or interpretation or integration of sensed (or sensory-motor) information before triggering a response (e.g. **uni-modal vs multi-modal** sensing, **vs** using an **amodal** interpretation);
- allow responses to be produced by different effectors (e.g. grasping with hand vs biting, running away vs flying away), where effectors or mode of response are conditionally selected;
- produce responses by coordinating different sensory-motor subsystems e.g. two hands and mouth used for different functions when peeling bark off a branch;
- integrate responses meeting different sorts of needs (e.g. when both hungry and thirsty, modify route to pass nearer fruit on way to river);
- allow conditions for choosing response to be extended and varied (e.g. using quality or proximity of food);
- choose time of response, or order of responses (e.g. select action but do something more urgent first);
- allow urgency, importance, costs, etc. of different selected responses to be compared in deciding when, whether or how to respond to new opportunity, (Section 6 of (Beaudoin & Sloman, 1993));
- allow a new **type** of response to be created (e.g. by planning a novel complex action);
- altering/extending the ontology used for interpreting sensory information (e.g. states of sensors or motors vs perceivable states of entities in the environment vs unobservable states e.g. magnetism, solubility);
- altering/extending ontology used for specifying actions and goals (strengthen smell vs get close to food);
- performing collaborative actions (e.g. pack hunting, swarming, mating, guarding/feeding infants, etc.);
- extend temporal or spatial gap between information being available and information being used – requires extended storage, new integration, and communication in control systems.
- “peep-hole” vs “multi-window” perception and action ((Sloman, 2000, 2003) also explained below.)

NB: Increasing gaps between acquisition and use

Evolution and individual development can produce bigger and bigger gaps between:

- **acquisition** of information items
- **use** of those information items

Such gaps are of various (overlapping) types:

- **gaps in spatial locations** - information acquired in one place used in another
- **temporal gaps** - information acquired at one time used in another
- **contextual gaps** - information acquired in one context used in another
- **generality gaps** - information about particular or specific cases used to derive or modify information expressing generalisations beyond those cases, e.g. by interpolation
- **abstraction gaps** - information rich in detail used to derive information discarding detail because only more abstract features are required,
e.g. acquiring detailed metrical information about a structure and retaining only topological information (such as information about nesting of locations) derived using details later discarded
- **representational gaps** - information acquired and initially processed in one form of representation can contribute to information expressed, manipulated and used in another form of representation. (E.g. from metrical to topological information, or vice versa.)
Example: in a SLAM mechanism (Simultaneous Localisation And Mapping) detailed sensory information from vision, range-finders, tactile sensors, odometers, etc., can be used to derive information about locations, walls, doors, corridors, etc. preserving **none** of the original sensor information.
http://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping
- **ontological gaps** - information acquired using one ontology contributes to new information using a different ontology (e.g. from physical behaviour to states of mind.)
- and many more...

A challenge to doubters

It's clear what needs to be done to prove that I am wrong about requirements for internal languages, offline processing, and complex architectures, in many non-human animals and in pre-verbal humans.

Namely, demonstrate working robots, [without the mechanisms discussed here](#), that have abilities similar to nest-building abilities of corvids, carcass-dismembering capabilities of hunting mammals, creative problem-solving in squirrels, infant caring capabilities of elephant mothers, play, social interactions, and language learning in pre-verbal humans.

If working models are too difficult now, then at least produce [detailed](#) specifications for designs that in principle future engineers could implement, and provide arguments showing how those designs can produce the required functionality (a type of argument often required before a human engineer's new idea is taken seriously).

The next slide presents an example of squirrel intelligence that arguably cannot be explained by innate reflexes or purely physical influences of the environment on the squirrel's body, or associations learnt from randomly varied actions.

The squirrel had made a number of failed attempts to climb up the door-frame, clearly having worked out a plan of sorts, which at first it could not execute, then later succeeded. (Photograph by Alison Sloman.)

[Many videos of squirrels defeating “squirrel-proof” bird feeders can be found on the internet.](#)

There are several videos on the Oxford ecology web site of Betty, the New Caledonian Crow, making hooks out of straight pieces of wire in order to fish a bucket of food out of a vertical tube She kept on spontaneously trying new methods even though she had already found methods that worked. Nothing in AI so far even comes close: <http://users.ox.ac.uk/~kgroup/tools/movies.shtml> I am **not** saying it's impossible.

Natural intelligence poses many challenges for AI

Researchers have found overlaps between AI and work in developmental psychology and animal cognition.

But there are two directions of influence:

Biologically-Inspired AI (BI-AI)

vs the less noticed alternative:

AI-Inspired Biology (AI-IB)

– the former is often much shallower)

(There have been several workshops and conferences on AIIB in the last few years – not all using that label.)

Explaining squirrel intelligence is a deep challenge
Grey squirrels defeat many “squirrel-proof” bird-feeders – e.g. the squirrel raiding our bird-feeder held by suckers high on a patio door.

Despite the very shallow grip, it managed to climb up the plastic-covered door frame just visible on left, then launch itself sideways across the glass, landing on the tray with nuts – a remarkable piece of creative intelligence – and ballistic launch control.



Squirrels seem to be able to reason about what to do in advance of doing it, even in novel situations, perhaps requiring a primitive form of “theorem proving” about what can work?

Compare Kenneth Craik’s ideas about animals using internal models (Craik, 1943) (“Model” is the wrong concept.)

Robotic understanding of affordances is currently **far** inferior to animal understanding.

See this presentation extending Gibson’s ideas about affordances

<http://tinyurl.com/BhamCog/talks/#gibson> (Gibson, 1979; Sloman, 2011b).

Online and offline information-processing

Recently, as computing devices, sensors, and motors became available with increasing power, while their sizes and costs shrank, there has been steadily increasing work on robots that sense and act on their environment using “online information-processing” not “sense-think-act” cycles.

Contrary to some anti-AI propaganda, and some AI factional propaganda, the need for AI research to include robots interacting with their environment was evident to the early AI researchers, e.g. at Stanford, MIT, and Edinburgh, but the computing resources available in the 1960s-70s were **grossly** inadequate, ruling out most forms of interaction. (See for example http://en.wikipedia.org/wiki/Freddy_II)

If it takes 20 minutes for a robot vision system to identify the rim of a mug, real time interaction is out of the question. Current machines are at least a million times faster, per cpu.

The ability to control real time interaction with objects could be called “**online (or on-line) intelligence**”. (E.g. (Adolph, 2005) shows how online intelligence develops in some young children.)

Other kinds of biological I-P: “**offline (or off-line) information processing**”:

- Information about large scale spatial structures for possible future use (e.g. SLAM “Simultaneous Localisation and Mapping” in robotics),
- Information about what sorts of things are **possible** in a given situation.
- Information about **what would happen if some of those possibilities were realised**
- Causal and mathematical constraints on possibilities: which can lead to scientific and mathematical discoveries, story-creation, art, philosophy, ... **not all in the service of action.** (Sloman, 2006, 2010b, 2011b)

Daniel Wolpert: the main or only function of animal brains is control of movement.

http://www.ted.com/talks/daniel_wolpert_the_real_reason_for_brains.html

That ignores, or disparages, the steadily increasing **decoupling** of information-processing from the immediate, or even remote, environment during evolution of more complex and sophisticated organisms.

Online control without communication

It is often assumed that information is essentially connected with communication, and that information-based control must therefore involve communication. (Norbert Wiener?)

However, use of information in controlling actions need not use communication, at least not **intended** communication.

There can be some internal communication between sub-systems, e.g. if vision is used, since visual processing goes through different stages in different brain mechanisms.

In other cases, e.g. a robot or animal using a **compliant** grasper to manipulate an object, some of the information in the shape of the object grasped is not transmitted along a communication channel, but directly used as part of the sensing process: sensing by deformation of the grasper: the object grasped can control the shape of a grasping hand.

All living things control some of the things that happen in their environment, as well as some of the things that happen inside their own bodies, though there's huge variation between different products of that biological control and huge variation in the means of control (e.g. some with, some without brains).

These forms of interaction do not fit some simple mathematical schema, like communicating by using strings of bits, or using sentences in a formalism defined by a grammar.

(AI researchers noticed that use of compliant manipulators can reduce computational problems of control of manipulation in the 1980s. Mechanical engineers understood this much earlier.)

Some benefits of offline processing

Gibson's ideas about affordances provide a sample of the benefits, but can be extended:

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

Being able to perceive, think about, reason about, make use of, allow for, or try to facilitate or prevent realisation of **possibilities**. (Sloman, 1996)

Being able to do counter-factual reasoning about the past:

What you did not do and why not.

Why someone or something else did not do something.

What could have been done and what the consequences would have been,

This is related to development of meta-semantic competences:

The ability to represent and reason about states and processes going on in other perceivers, thinkers, learners, goal-directed agents, or oneself, including:

- self-awareness:
e.g. "what did I do wrong?" "How can I avoid that error next time?"
- Daydreaming,
- Story telling,
- Joke-making,
- Art,
- Music
- Science
- mathematics. (Expanded below)

See Toddler Theorems <http://tinyurl.com/CogMisc/toddler-theorems.html>

Online vs offline intelligence

This gorilla seems to exemplify offline intelligence: watching something without immediately putting the information gained to use in an action – e.g. perhaps watching the “suspicious” movements of a potential rival male?

Offline intelligence used in connection with conspecifics, predators, or prey often requires use of **meta-semantic** competences: abilities to represent things that represent.

Meta-semantic competence can also be self-directed (e.g. “What was wrong with my planning process?”)

Image courtesy of Wikipedia



<http://de.wikipedia.org/wiki/Gorillas>

Offline processing: the roots of mathematics

Many (but not all) of those who emphasise embodied cognition ignore important aspects of embodied biological information processing using “offline” reasoning processes.

So they cannot explain how humans became mathematicians, scientists, philosophers, creative artists, engineers of the very large or very small...

Some important sorts of learning that were required for development of mathematical knowledge, involve abilities to reason about what is possible or impossible without actually changing anything in the environment, e.g. the discoveries leading up to Euclid’s Elements, (also noticed by Piaget, despite his inadequate theories about information processing)

Such reasoning can be done with eyes shut and no physical actions.

Unfortunately the mathematical education many philosophers and scientists now have had at school does not include experience of doing mathematics, or learning the difference between mathematical necessity and a very high probability.

Some simple examples of geometrical reasoning that seem to be beyond current AI:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html>

A wider variety of examples relevant to very young, even pre-verbal, children, and possibly also some other animals can be found in:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html>

The competences are also relevant to human abilities to silently compose stories, jokes, poems, tunes, etc.

These proto-mathematical competences used in offline intelligence also seem to exist in some other species (e.g. primates, elephants, some nest building birds, ...).

“Offline” reasoning abilities need to be added to machines with “embodied” cognition, or “enactivist” AI. But we need to learn more about their precursors first.

More examples of offline information processing

Added: 4 Apr 2016

A document begun in 2015 and still work in progress, gives many more examples of relationships between offline visual processing and mathematical discovery, using examples of perception of impossibilities.

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

Transitions in motive generation and processing

In simple microbes there could only be triggered reactions

(reflexes that succeed or fail, or hill-climbing reflexes).

Later, mechanisms evolved that trigger behaviours increasingly indirectly.

E.g. by first triggering a goal: something that represents a state to be achieved, avoided, or preserved.

This requires new mechanisms for selecting appropriate actions.

E.g. what to do to get food or to avoid an obstacle can vary, using learnt abilities.

(Goal state specification may have been one of the earliest roles for structured semantic content.)

Goal descriptors allow more flexibility than reflex actions, but more information is needed and the genome must specify context-sensitive mechanisms for selecting actions.

Simple mechanisms could use repeated measurements (e.g. of density of a chemical in the environment) and choose actions that tend to increase or decrease values – hill-climbing.

More complex mechanisms might partition possible actions into categories and explore consequences of each category physically, or theoretically as in AI planners (Craig, 1943).

In later species goal specifications could be more complex (e.g. boolean combinations) or use more sophisticated forms of representation (e.g. predicates, relations, quantifiers, etc.), requiring more and more sophisticated plan generation, plan comparison and goal selection mechanisms.

Initiation and modulation of internal or external processes, whether individual or collaborative, involves motivational mechanisms with many functions,

e.g. proposing, comparing, selecting, or abandoning goals; initiating or modulating actions;

e.g. changing speed, direction, means, tool or order of steps, etc....

Many researchers who know about these mechanisms do not view these transitions as part of a global process of enhancing/enriching forms of biological information processing.

Some of the unobvious required complexity is described in (Beaudoin & Sloman, 1993; Beaudoin, 1994)

Background reading: R.W. White (1959), H. A. Simon: (1967b, 1967a, 1969)

How goals/motives can be generated and selected

Triggering goals provides more flexibility than triggering actions.

- Triggering a goal allows for gaps between initiation of action and selection of details: instead of directly initiating a detailed action sequence (innate or learned) a trigger adds a goal to a motive sub-system, leaving other mechanisms to use additional information (e.g. current context) to select among possible ways of achieving the goal – e.g. choose a route after climbing a tree to get more information.
- It allows more context-sensitive mechanisms using current information to choose between incompatible actions, instead of selecting according to strength or type of the triggering stimulus.
- It allows more flexible ways of dealing with new triggers encountered during action: instead of the new trigger causing or not causing re-direction of action, it can trigger formation of a new goal, and then motive comparison mechanisms can make use of a range of information sources to choose whether to interrupt the current action to pursue the new goal, or postpone the new goal till later, or modify the current action so as to serve both goals, e.g. by modifying a route. (Beaudoin, 1994)

Triggering motives rather than actions can be thought of as triggering a more abstract and general reflex than one that specifies all the detailed muscular contractions.

Motive generating mechanisms could have evolved because they expanded power and generality of information processing in organisms long before evolution produced motives based on **rewards**, or positive and negative **reinforcement**. (Simon, 1967b)

Mechanisms for manipulating scalar or partially ordered rewards would add generality, permitting more general context-sensitive choices between competing motives, instead of requiring all preferences to be learnt separately – but **some** might still need to be! (Sartre)

It's likely that evolution produced more intermediate cases than anyone has noticed.

Architecture-based vs reward-based goal processing

How should goals be triggered and how should they be selected?

It is widely assumed that some sort of reward mechanism is required to generate goals (e.g. needing energy causes an appropriate food seeking goal to be triggered) and to compare and select goals (e.g. act on the goal most likely to produce a high reward).

However, a reflex mechanism can generate a goal **not** because the individual expects that achieving the goal will bring some reward but simply because natural selection “discovered” that having such a goal generator tends to lead to more reproductive success.

When something triggers a mate-seeking goal in a simple organism, the organism need not have any notion of the benefits of mating.

Likewise, in young humans, and the young of many other species, there seem to be motive generators that are triggered to produce goals related to structures and processes in the environment that bring no reward that the individual can expect, but which happen, sometimes to provide information that contributes to later learning.

- So seeing something movable may trigger creation of a goal to move it, and if there are no other goals currently active the child will act on the goal.
- Seeing something move may trigger creation of a goal to repeat, or modify, or reverse, that movement.
- Seeing a gap in a fence or wall may trigger a goal to go through the gap.

A robot that learns a great deal in this way is reported in (Ugur, 2010) – using different terminology.

I call this mechanism “architecture-based motivation” (A-B-M) and the goals “architecture-based goals”.

For more on “Architecture-Based Motivation vs Reward-Based Motivation”, see

<http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#907>

“Intrinsic motivation” is different from architecture-based motivation, as explained below.

Contrast “Intrinsic” Motivation

“Intrinsic motivation” and “architecture-based motivation” are not the same.

There is a vast amount of literature concerned with what is referred to as “intrinsic motivation”, though it is not clear that all authors use that phrase in the same way.

However there is a specific concept of intrinsic motivation that is close to but different from the concept of architecture-based motivation, explained in (Baldassarre, 2011) as follows:

“Intrinsic motivations ... can guide learning of skills and knowledge without having to rely upon learning signals generated by homeostatic need regulation. To this purpose, intrinsic motivations generate learning signals on the basis of the success of the acquisition of skills and knowledge themselves instead of their effects on homeostatic needs. To do this, intrinsic motivations are based on mechanisms that measure the success of the acquisition of skills and knowledge directly within the brain. For example, these mechanisms drive organisms to continue to engage in a certain activity if their competence in achieving some interesting outcomes is improving, or if their capacity to predict, abstract, or recognise percepts is not yet good or is improving: the brain detects all these conditions without involving the visceral body.”

This seems to me to be identical with or very closely related to the “Effectance” goals of (White, 1959), which are selected because they produce a special feeling (of “efficacy”).

Intrinsic motives and effectance goals produce rewards recognised by the individual and explicitly sought by the individual, whereas architecture-based goals when achieved may or may not produce some immediately recognised benefit, but in many cases do not: they merely provide information about the effects of actions which is automatically added to some kind of memory, as in (Ugur, 2010).

Some of the information gained through A-B-M is later used by processes of knowledge reorganisation that derive new complex generative competences, discussed below.

(Karmiloff-Smith, 1992) refers to this as “Representational Redescription”.

The best known example is the transition from pattern-based language use to syntax-based language.

Varieties of motivational mechanisms

We need a better understanding of the diversity of types of motivation and affect found in intelligent systems, operating in different architectural layers, and evolved at different times.

Multi-layer architectures are discussed later.

Varieties of motivational mechanisms and meta-mechanisms include:

- motive generators
- motive comparators
- motive generator generators
- motive comparator generators
- motive generator comparators
- motive comparator comparators
- and so on recursively, within limits (what limits?).

(White, 1959; Simon, 1967a; Wright, Sloman, & Beaudoin, 1996; Sloman, Chrisley, & Scheutz, 2005)

It is often assumed that all selection of goals (states to be achieved, preserved, prevented, etc.) **must** be based on expectation of some reward.

It is also often assumed that alternative rewards **must** all be comparable, so that utility/reward maximisation can be the basis of (rational) decision making.

Both assumptions, though very widely held are totally without justification.

(a) Some animals decide on the basis of evolved or somehow learnt **rules**, even if they do not optimise anything for the decider. E.g. architecture-based motivation (Sloman, 2009a; Ugur, 2010)

(b) Choice mechanisms may use **partial** orderings so that A is preferred to C, and B is preferred to C, but nothing in the system provides a basis for choosing between A and B.

(Later on certain sorts of learning could lead to a preference for A or for B.)

Mathematical competences and biological evolution

Contrary to some theories emphasising the role of embodiment in mathematics, consciousness, intelligence, etc., offline processing is central to the development of intelligence, including use of mathematics, as explained in (Sloman, 2009b).

Most people think that learning mathematics requires being taught by someone who has already learnt mathematics

QUESTION:

Who taught the first mathematicians – before there were teachers?

ANSWER

Biological evolution and the environments we evolved in, working together.

(Sloman, 2010b)

That's the process educators are contributing to.

The bootstrapping is still continuing – and we have opportunities to push it forward.

building on spontaneous discoveries of toddler theorems:

<http://tinyurl.com/CogMisc/toddler-theorems.html>

But the opportunities are not being taken.

It seems that most schools no longer teach mathematics properly, so that students do not experience the activity of discovering and proving theorems; so many researchers in cognitive science, philosophy, and robotics, do not understand the importance of off-line (i.e. temporarily disembodied) intelligence for many of our advances in mathematics, science and engineering.

And they don't understand the difference between mathematical necessity and high probability, nor the need to be able to reason about possibilities in order to think about probabilities.

Changes in information transfer across generations

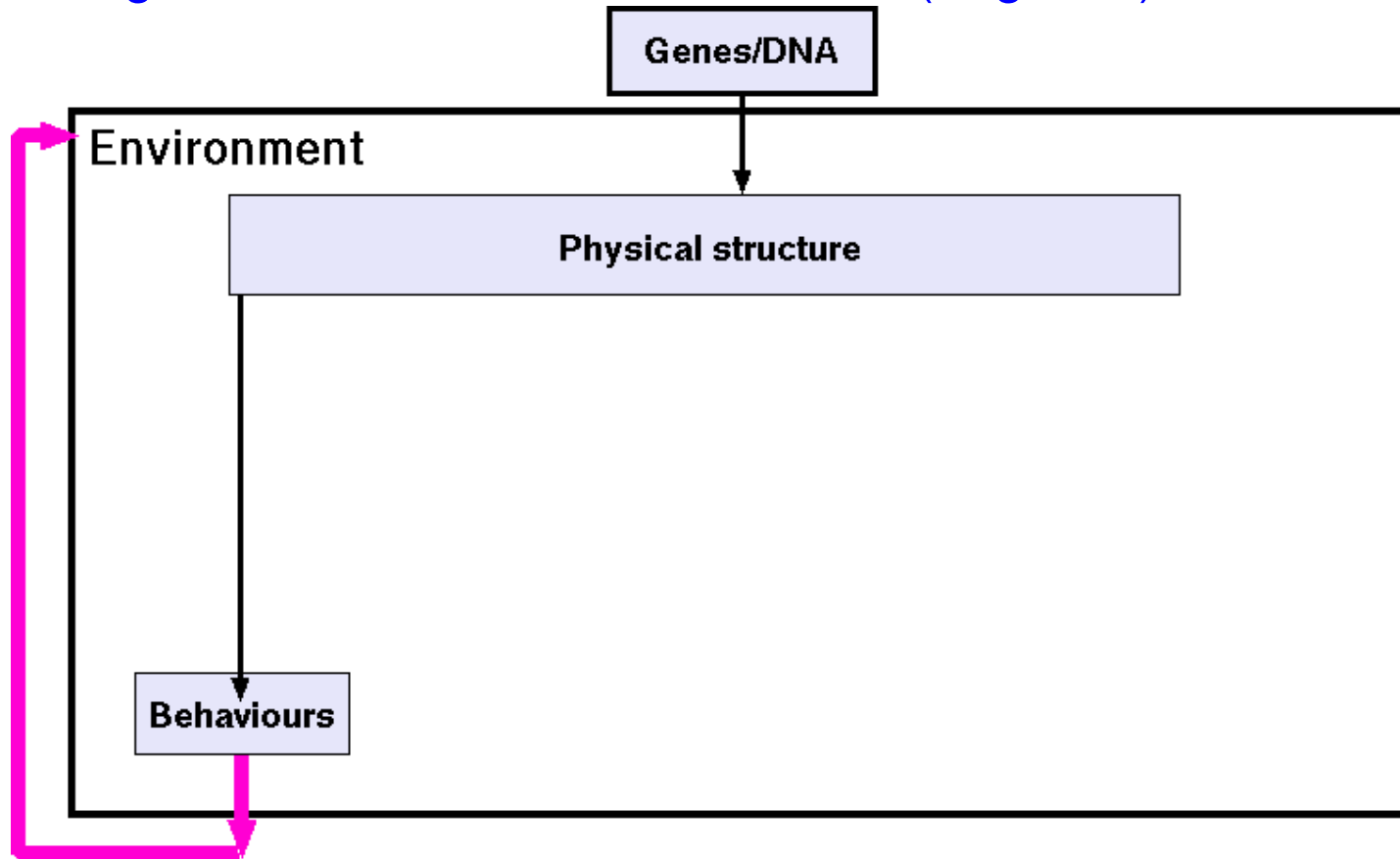
Another class of evolutionary transitions:

- changes in information transferred across generations,
- changes in how it is transferred
(genetic information, in the broadest sense)
- precocial/altricial strategies, e.g.
- content transferred vs learning mechanisms transferred
- generic schemas to be instantiated transferred
- meta-schemas (to be instantiated to generic schemas) transferred,,,
- illustrated by Chappell-slovan-miall diagrams depicting “epigenetic layering” in the following slides, based on (Chappell & Sloman, 2007)

May be compared/contrasted with Waddington’s “Epigenetic Landscape”, depicted below.

Individual developmental trajectories I

Routes from genome to behaviour : the direct (original?) model.

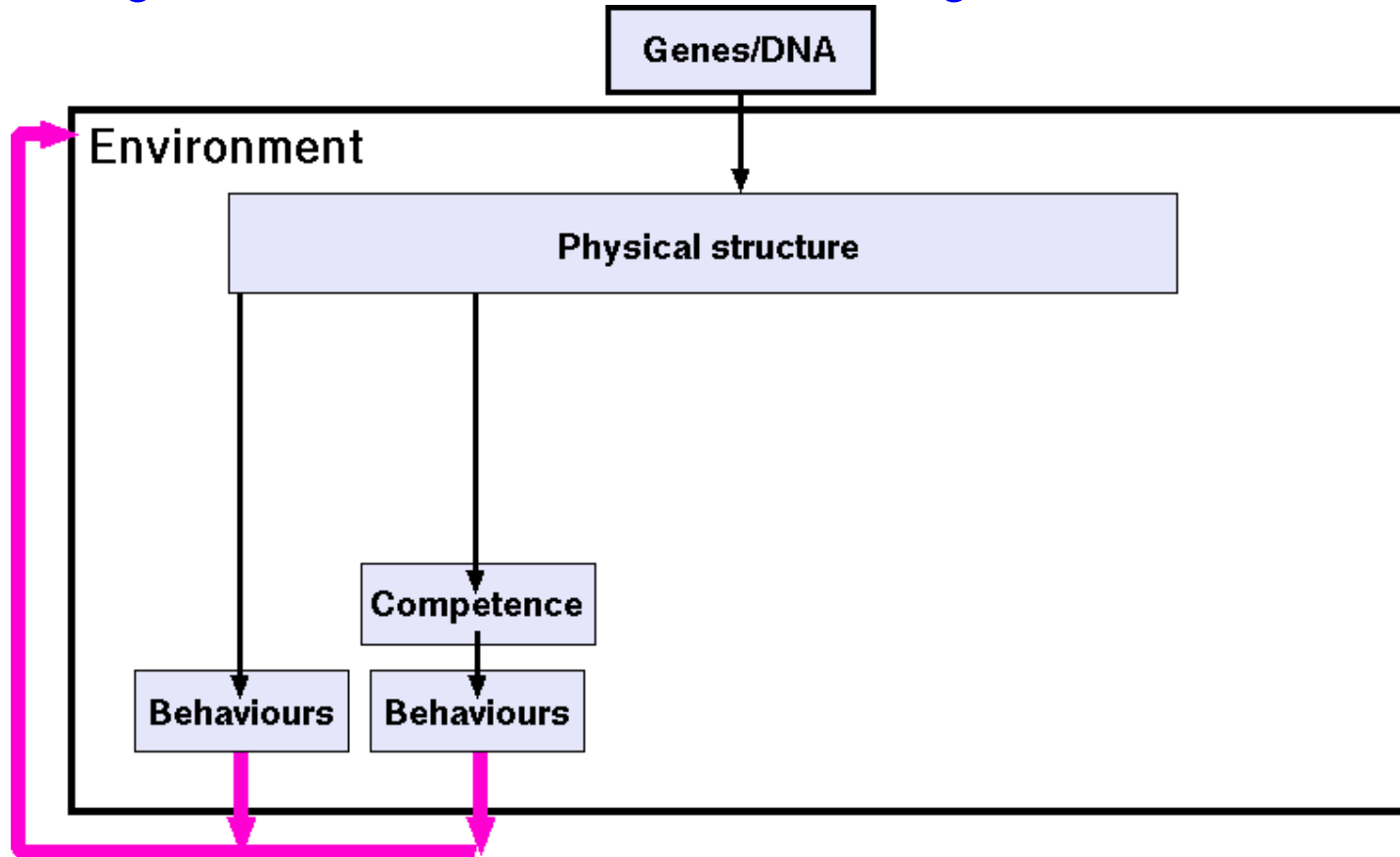


The vast majority of organisms (including micro-organisms) are like this. Many don't live long enough to learn much – they have to make do with innate reflexes. Other organisms have more “inside the box”.

Some of the more complex learning/development possibilities can be represented by alternative routes, as shown in the next few slides.

Individual developmental trajectories II

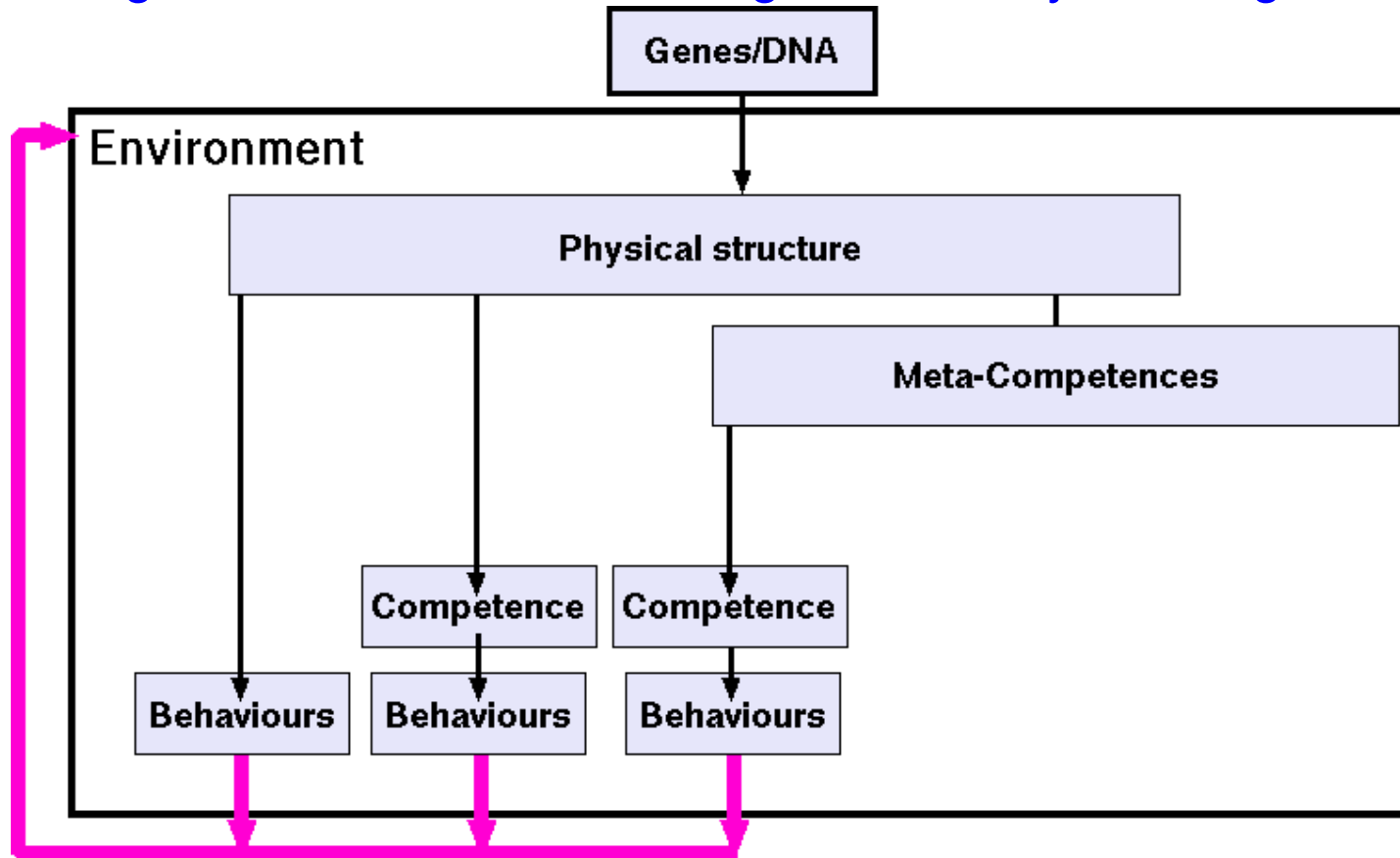
Routes from genome to behaviour : the two-stage model.



Some more complex organisms, instead of having only rigid (reflex) behaviours, also have competences that allow them to respond in fairly flexible ways to the environment: adapting behaviours to contexts: reflexes trigger motives that produce behaviours instead of triggering behaviours directly. Requires innate, or trainable, mechanisms for selecting means to achieve the goals.

Individual developmental trajectories III

Routes from genome to behaviour : stages added by learning.

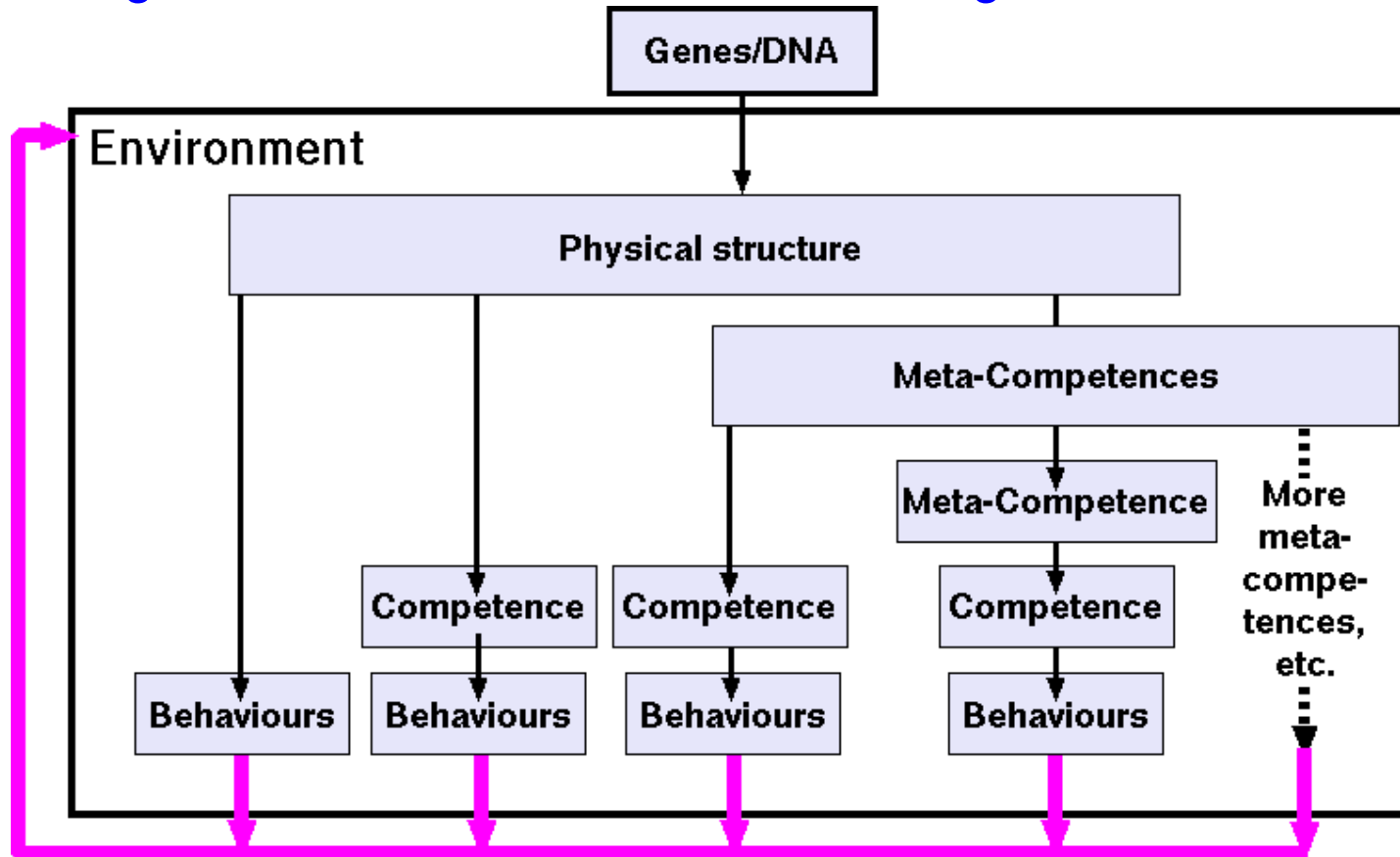


Genetically determined **meta-competences** allow individuals to respond to the environment by producing new types of competence, based on learnt features of the environment, increasing flexibility and generality.

This requires extra learning mechanisms to allow appropriate competences to be generated and to allow those competences to select and execute behaviours.

Individual developmental trajectories IV

Routes from genome to behaviour : the multi-stage model.



Some can also develop new meta-competences, on the basis of meta-meta competences.

Humans seem to be able to go on developing meta-meta-competences until late in life.

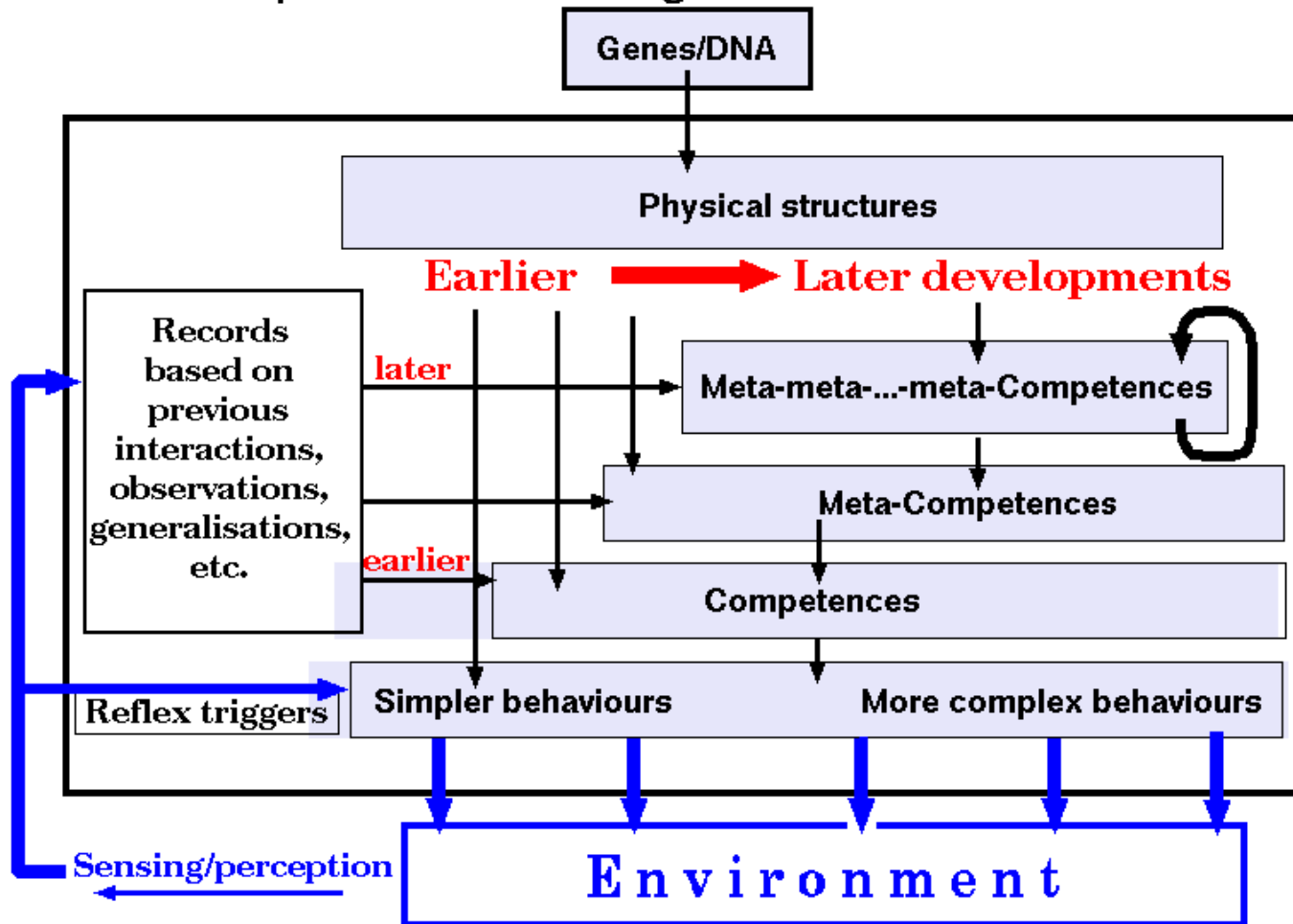
E.g. learning to learn from reading, from complex experiments, etc.

Individual developmental trajectories V

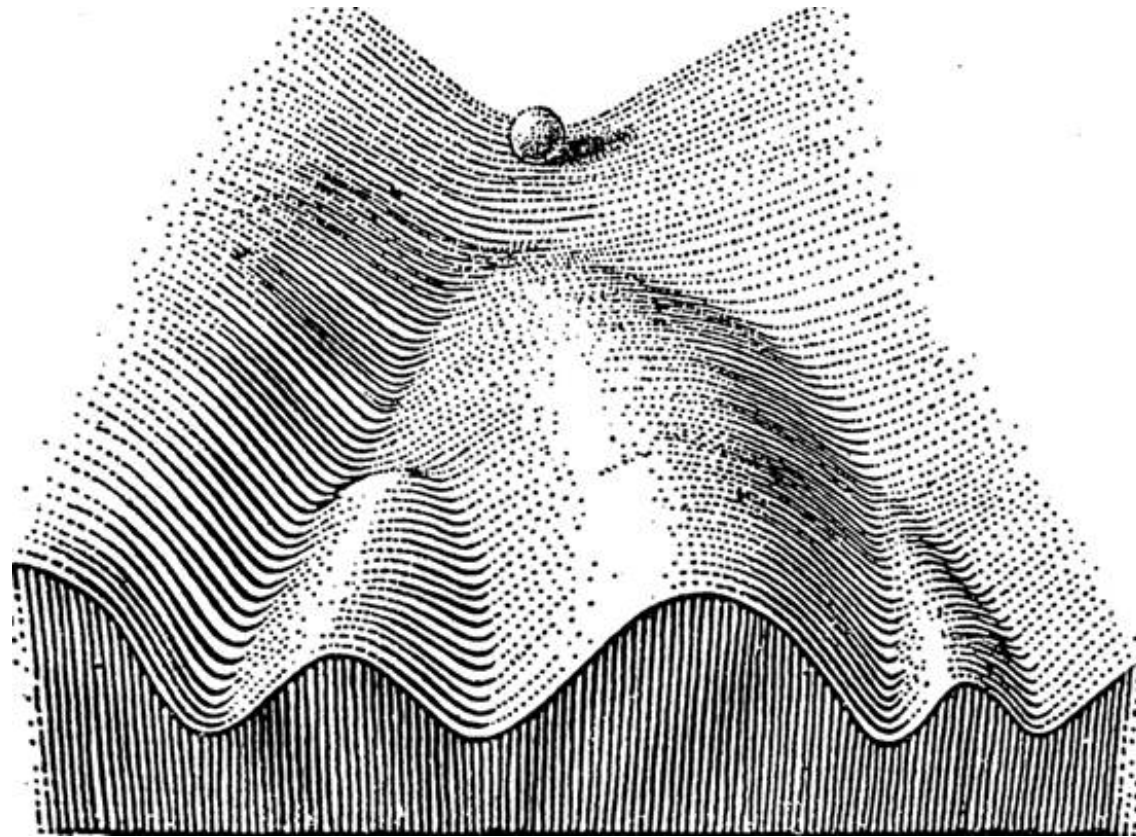
(Added 4 Apr 2016) [More on routes from genome to behaviour](#) :

This version adds more detail regarding the different stages in development of competences of various sorts.

Multiple routes from genome to behaviours



Compare Waddington's Epigenetic landscape



(Waddington, 1957)

Not nearly rich enough – I suspect he would have agreed.

Warning: diagrams don't constitute a theory

The diagrams above do not express a precise theory.

They are merely intended to suggest varieties of “multi-layered” interactions between the genome and the environment (which may or may not contain other intelligent systems, e.g. prey, predators, rivals, offspring, etc.) during development and learning.

The above ideas imply that one of the evolutionary transitions required for such mechanisms to work was delayed development of some brain mechanisms, e.g. meta-cognitive mechanisms whose operation depends on inspecting results of substantial prior learning and then developing new forms – e.g. replacing collections of empirical summaries with a **generative** system.

This seems to be an example of what Karmiloff-Smith calls “Representational Redescription”, in her 1992 book *Beyond Modularity*, discussed here:

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

Additional examples that seem to arise in the development of young children are referred to as “Toddler Theorems” here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html>

The idea that delayed brain development (e.g. in frontal lobes) was important for humans and some other “altricial” species was also suggested in lectures by Manfred Spitzer at a workshop in 2002.

Later slides will introduce architectures with different information processing layers that evolved at different stages. Some of the higher level layers may be grown using the late-developing mechanisms.

Meta-competences and abstraction

The ability to produce meta-competences that can be instantiated in many different ways can be compared with some of the developments in programming languages and programming strategies in the last 60 years.

In particular, it looks as if evolution was often able to discover that something evolved was a special case of something more general that could be instantiated in different ways.

Evolving the more general mechanism requires something like mathematical abstraction.

And the ability to use new abstractions requires something like the inheritance mechanisms in programming languages with polymorphism: especially parametric polymorphism.

E.g. give me a list of entities of any sort and a test for which of any two entities should come first, then I can give you a polymorphic procedure that will use your test to sort an arbitrary list of items, of any type.

But I can't tell you what lists sorted by my procedure will look like, since that will depend on which "before" test is used.

(A sort of "reverse Baldwin" effect.)

Are there information-processing mechanisms that "naturally" support this process of polymorphic abstraction?

Perhaps information encoded chemically naturally allows fixed sub-structures to be replaced by generic templates that can be instantiated differently?

Genetic Programming (GP) techniques, which operate on tree structures are closer to this than Genetic Algorithms (GAs), which operate on linear strings.

<http://www.genetic-programming.org/>

http://en.wikipedia.org/wiki/Genetic_programming

<http://cswww.essex.ac.uk/staff/poli/gp-field-guide/>

This is also a recurring feature of mathematical discovery.

Compare "Representational redescription" done by individuals. Maybe evolution can do it too?

Back to online vs offline intelligence

The next few slides introduce ideas often associated with the emphasis on embodied cognition, enactivism, and the like.

In particular the idea that humans and other animals form dynamical systems closely coupled with the environment.

I'll try, using some very abstract and metaphorical diagrams, to show how the ideas about dynamical system can be generalised to give dynamical systems a useful role without being tightly coupled to the environment via sensory motor systems.

In particular, some previously constructed dynamical systems can be in a dormant state ready to be rapidly assembled into a complex, possibly novel, dynamical system to deal with a new problem, e.g. interpreting a new complex visual scene rapidly, or forming a new explanatory hypothesis.

Based partly on the informal experiments on vision reported in (Sloman, 2008b)

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk88>

These ideas are all still in a very early stage of development.

Nothing has been implemented!

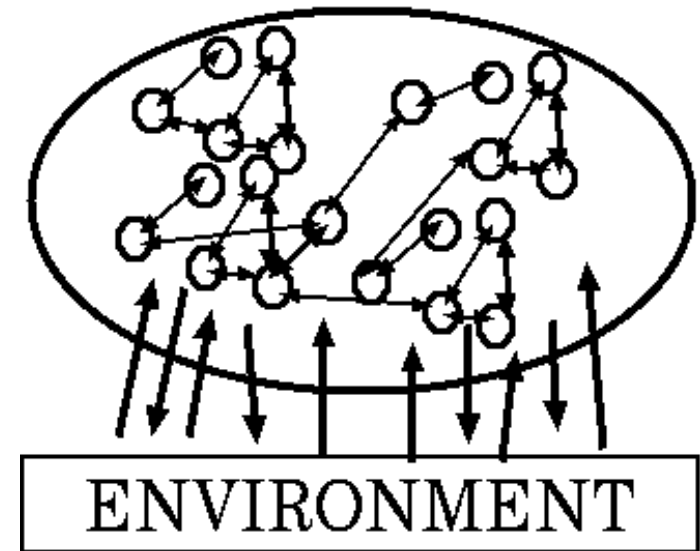
Dynamical systems I (Online intelligence)

Many researchers who emphasise the importance of embodiment, and online intelligence, also emphasise dynamical systems —
With constantly changing internal states.

Especially dynamical systems closely coupled with the environment, e.g. (Berthoz, 2000) where the nature of the coupling depends partly on the agent's morphology and sensory motor systems.

The figure depicts a relatively simple architecture closely coupled with states and processes in the environment.

The small circles crudely represent possible states (of the whole system) and the small arrows possible transitions between those states (not necessarily all reversible). States are often thought of as vectors of scalar values of input signals, output signals and various internal changing measures.



But evolution “discovered” the need for, and produced designs for, many more types of dynamical system with very different properties, used in many different sorts of information-processing architecture. (Sloman, 2006)

Figures below show multi-layered, multi-functional, dynamically constructed, or dynamically activated (when needed), dynamical systems, some decoupled from the environment.

Dynamical systems II (including offline intelligence)

More complex dynamical systems may have large numbers of sub-systems, many dormant most of the time but able to be re-activated as needed, sometimes spawning new ones, used over different time-scales.

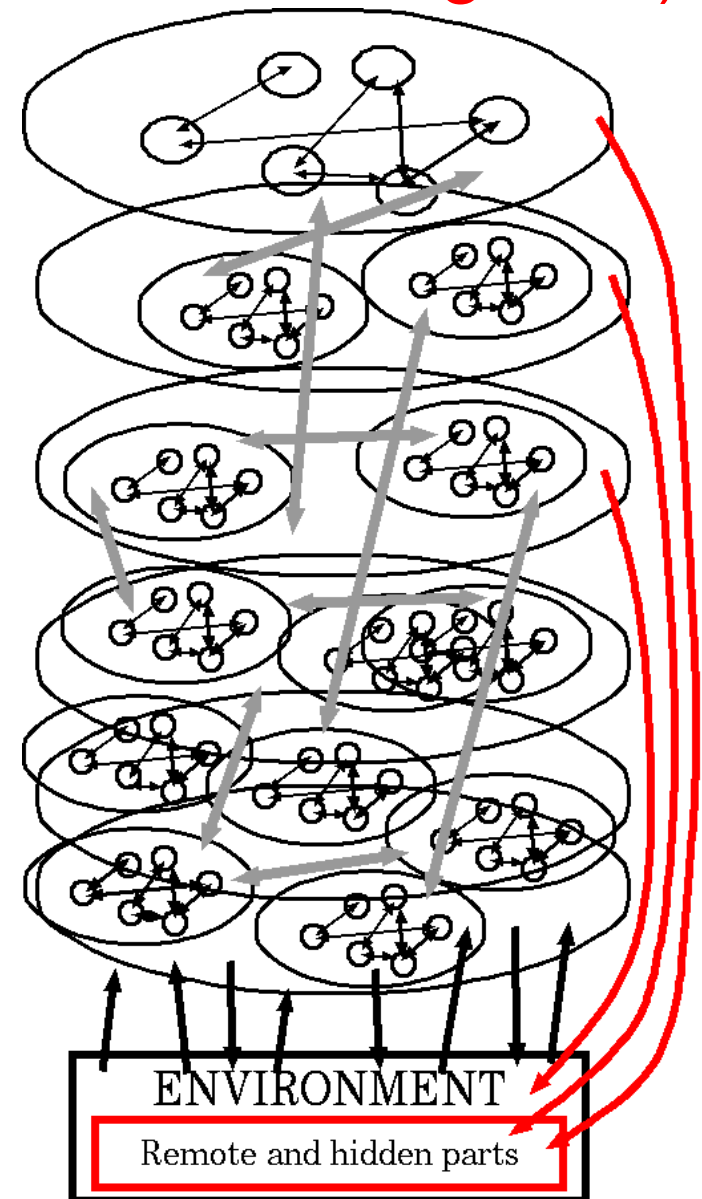
E.g. Perceiving, motive-formation, adaptation, learning, acting, self-monitoring, socializing, cooperating..., can all involve information processed at multiple levels of abstraction.

Hypothetical reasoning: Science, mathematics, engineering, philosophy...

Some abstract processes may run decoupled from the environment – but sometimes produce results that are stored in case they are useful...

Which species can do what? – What intermediate stages are possible:

- in evolution?
- in individual development?



Do not believe symbol-grounding theory: use theory-tethering instead.

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models>
(The “tethering” label was proposed, in discussion, by Jackie Chappell).

Dynamical systems III

A point that is obvious to mathematicians, but usually ignored by roboticists:

Sometimes, the forms of reasoning and representation required can challenge the brain's working memory capabilities.

As a result some portions of the external environment are used as extensions of internal forms of representation.

E.g. using diagrams or calculations on paper, or using an abacus for doing arithmetic, or seeing the internals of a machine as explaining its powers. Discussed in (Sloman, 1971, 1978)[Chapter 7]

More examples are in:

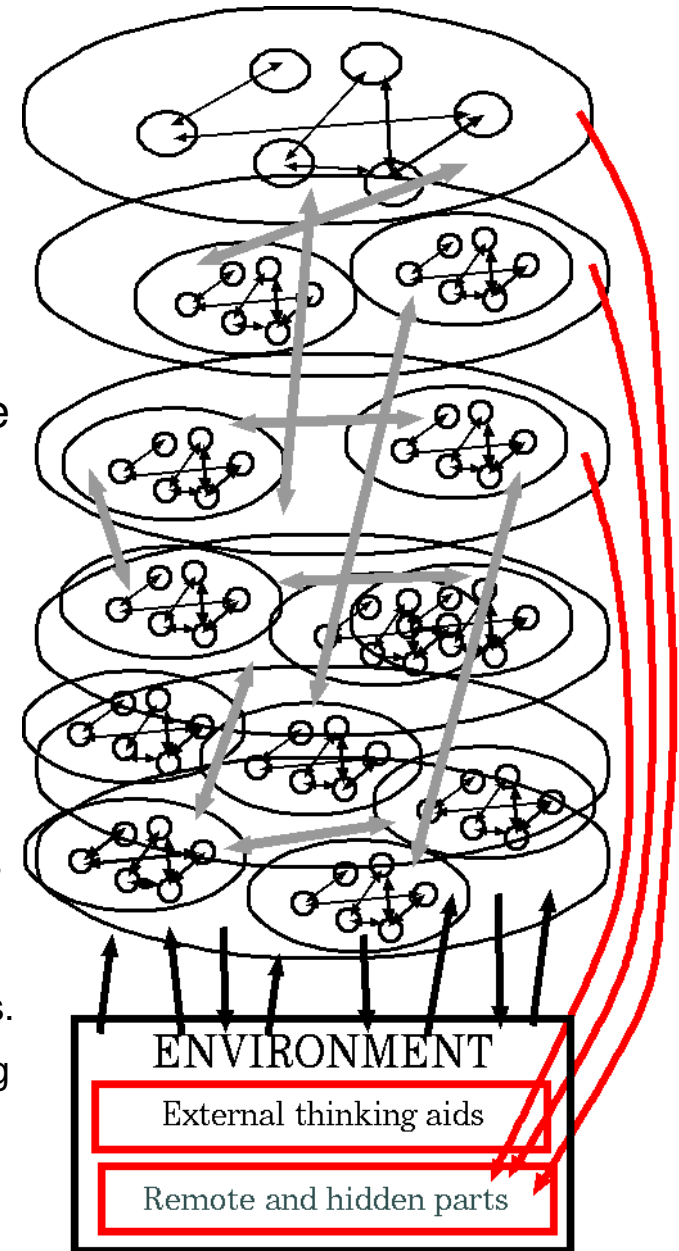
<http://tinyurl.com/CogMisc/triangle-theorem.html>

Mediated remote reference

P. Strawson's discussion (Strawson, 1959) of reference to particulars that are remote (in space and/or time) also mentions the reference-facilitating role of intermediate objects in the environment, including other humans, and their causal and referential relationships.

How something can refer to, or express information about, something else is an old philosophical problem.

One old philosophical answer, "concept empiricism", says that anything that refers must be derived from patterns in sensory and/or motor signals – a theory recently labelled "symbol grounding", still widely accepted although it is wildly wrong: see the next slide.



Alternatives to symbol-grounding

The building blocks of sentence-meaning, used in constructing propositions (true, false, incoherent), questions, motives, plans, theories, arguments, etc., are called “concepts”.

Concepts can also be used in constructing pictures, models, diagrams, maps and other non-verbal information structures, but the modes of representation and composition are different. (Sloman, 1971)

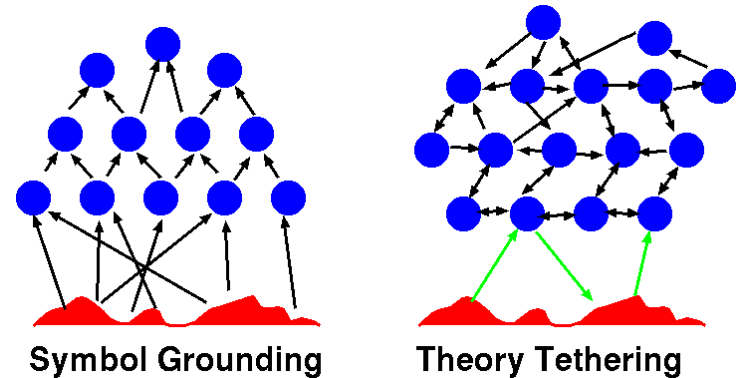
(Compare concepts expressed in sign languages used by deaf communicators.)

A set of available concepts, plus modes of composition, determines an **ontology**.

“Concept Empiricism” is the ancient philosophical view that every concept must either be derived (e.g. by some kind of abstraction process) from experience of instances, or else defined in terms of such concepts (e.g. using “and”, “or”, “not” and other logical operators). This would make science impossible!

Information users need an ontology

Many roboticists, AI and cognitive theorists assume (like some old philosophers) that concept empiricism **must** be true, though following (Harnad, 1990) they call it “Symbol Grounding” theory. Immanuel Kant (1781) demonstrated the incoherence of concept empiricism, and 20th century philosophers of science showed that it could not account for deep explanatory concepts e.g. “charge”, “voltage”, “current”, “neutrino”, “gene”, ... (Sloman, 2007b).



A much better theory is that references to entities outside the perceiver use **concepts embedded in powerful predictive and explanatory theories**, where the concepts mostly get their meaning from their roles in the theories, not by being “grounded”. (The structure of the theory limits possible models.)

The theory (not individual symbols) is then linked to observables and measurables, via perceptual and other mechanisms (e.g. using special, changeable, measuring technology or methods of testing). (Carnap, 1947) This can be called “Theory tethering” (label due to Jackie Chappell (Chappell & Sloman, 2007)).

Theory tethering is more powerful than symbol grounding (Sloman, 2007b). **Robots don't need symbol grounding!**

Dynamical systems IV

The dynamical systems diagrams in previous slides may be regarded as schematic architecture diagrams –

indicating that some organisms have layer upon layer of interacting dynamical systems, running in parallel, some of them dormant at any time,

with some sub-systems closely coupled with the environment through sensory and motor mechanisms others uncoupled and more “free running”.

In (Sloman, 1978, Chapter 9),(Sloman, 2008b) it is argued that some facts about human visual perception, suggest that human minds/brains (especially visual subsystems) can very rapidly construct highly complex, possibly temporary, special purpose dynamical systems with some details in registration with the optic array, some parts of which persist over time, while new systems are constructed for new scenes.

Example: the speed with which you can take in many (though not all) features a complex scene in an unfamiliar city at various levels of abstraction, when turning a corner into a new busy road, with some of the information persisting even if the scene changes – e.g. remembering the briefly seen car parked opposite even though a large van now obscures it.

[Compare animals moving rapidly through dense foliage.](#)

Human abilities to perceive and reason about [affordances](#) of many kinds suggest that the visual processes include meta-cognitive mechanisms that explore possible changes in the situation, and deduce constraints and consequences of realising the possibilities.

((Sloman, 2011b))

Illustrated with examples of (pre-Euclidean?) reasoning about triangles, in

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html>)

These slides present alternative ways of viewing information-processing: every complex system has multiple views.

Ontologies for vision I

The next few slides present examples of visual information processing that provide pointers to some of the sub-ontologies and forms of representation used in human vision.

Ideally there should be comparisons with vision in other animals – not easy to do.

A useful approach (suggested to me many years ago by Max Clowes) is to consider what changes in interpretations of ambiguous figures – as in next slide.

Other questions include:

What ontology is required to express what's common to two views of the same object?

In what processes can the perceived structures take part?

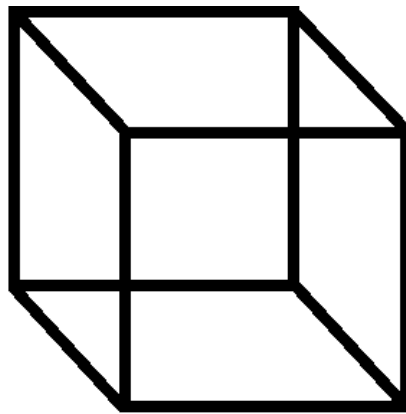
What causal relationships are perceived?

(Sloman, 1982, 1986, 1989, 1996, 2001)

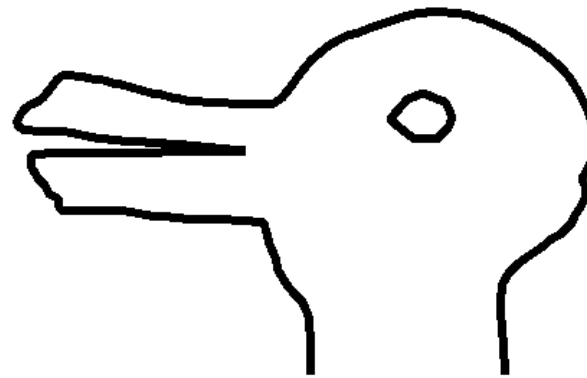
Ontologies for vision II

Contrary to the assumptions of many vision researchers, ontologies used in human, animal or robot perception may be very abstract and hard to identify.

Ambiguous figures can reveal some different ontologies used in interpreting the same visual input (physical stimulation): [What changes when these figures flip?](#)



Necker Cube



Duck-rabbit

In one picture, only geometrical relations change. In the other no geometrical relations change, though there are important non-geometric changes.

What are the implications:

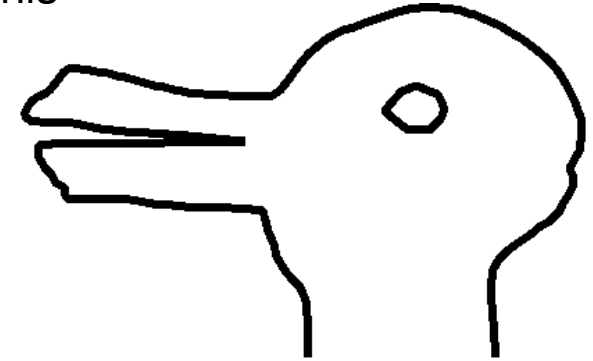
- for ontologies used?
- for forms of representation?
- for architectures?
- for practical uses of the information?

More on “Seeing as”

The Duck-Rabbit (Compare the Necker cube)

There’s a 2-D pattern as with the cube, but you probably see this one as an animal with particular physiological features.

If you stare long enough it should flip between two animals with different features, even though there is no change in either the 2-D or the 3-D geometry of what you see (unlike the Necker cube).



E.g. a part of the image can be seen as ears, or a bill:
but both alternatives occupy the same 3-D space.

An even more abstract change probably occurs: you see the animal facing to your left or facing to your right depending whether you see it as a duck or as a rabbit.

How can a machine be made to see something as facing left, or as facing right?

The machine would have to have an **ontology** that includes things having states that involve facing, moving, seeing, etc.

It might relate the direction in which something is facing to what it can see, and which way it is likely to move.

This requires having the ability to think about what **could** happen even if it is not **actually** happening now.

How would such possibilities, and constraints on possibilities be represented in the machine (or in an animal’s mind).

How can we find out how animals (including humans) represent such information?

Will trying to build a working model help? (Often? Rarely? Never?)

Ontologies for vision III

What sort of ontology is required to express the differences that you see in these two faces? (**Stare at each in turn - what's different?**)

They are not merely physical differences – you probably also see differences that presuppose that you are looking at an information-processor, e.g. something like a human with human-like information-processing capabilities, and resulting states.

Do the eyes in the two faces look the same?

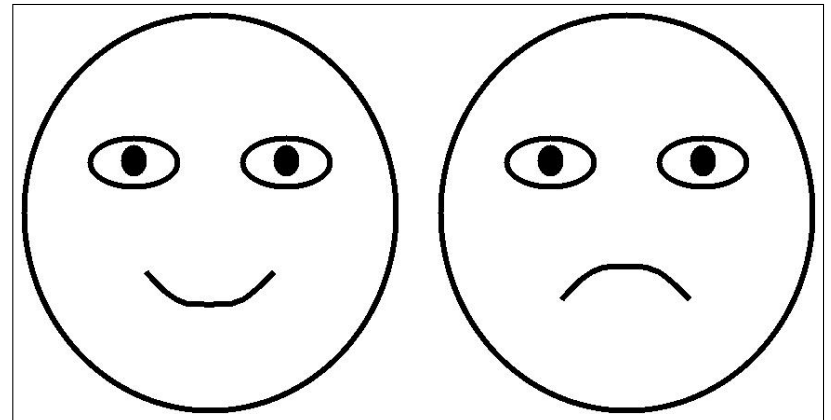
(Most people I've asked see differences in the eyes!)

Geometrically the two pairs of eyes are identical.

What sort of ontology do you need in order to express the difference in the eyes?

What are the implications:

- for ontologies used?
- for forms of representation?
- for architectures?



As indicated in summarising the Popeye program below, human visual perception (and vision in human-like robots) requires a multi-layered architecture, including some layers that evolved much later than others, overlapping with biologically newer, more central, “cognitive” layers. (Sloman, 1989)

Compare requirements for hearing or reading verbal communications.

Ontologies for vision IV

Add more examples illustrating the varieties of sub-ontologies used in human vision.

E.g.

perception of processes

perception of possibilities and constraints (Sloman, 1996)

perception of causal relationships (Michotte, 1962)

perception of “effort” in movements. (G. Johansson, 1973)

perception of social interactions (Heider & Simmel, 1944.)

There's lots more to be said about the functions of human and animal vision, the mechanisms used, the ontologies used, the forms of representation used, and lots of questions to be asked about the transitions in evolution, in individual development and learning, in cultural influences,

Later slides using pictures of impossible objects undermine the belief that seeing 3-d structures involves building 3-d models.

Toy Example: The Popeye visual I-P architecture

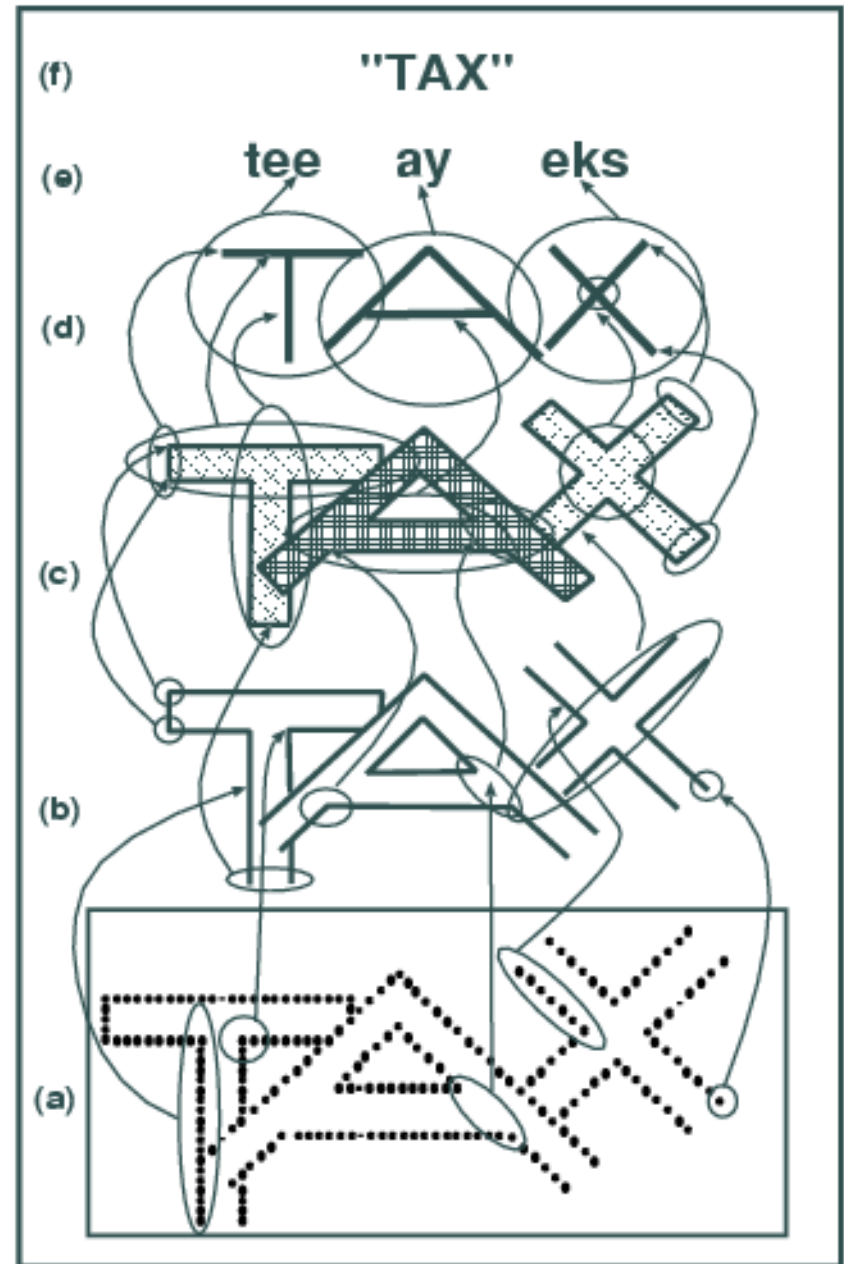
The POPEYE vision system developed at Sussex University was summarised in Ch.9 of (Sloman, 1978)

It interpreted artificial “dotty” test pictures (e.g. (a) here) depicting a word made of laminar capital letters with straight strokes, drawn in a low resolution binary array with problems caused by overlap and positive and negative noise.

It degraded gracefully and often recognised a word before completing lower level processing. The use of multiple levels of interpretation, with information flowing in parallel: bottom-up, top-down, middle-out, and horizontally allowed strongly supported fragments at any level to help disambiguate other fragments at the same level or at higher or lower levels.

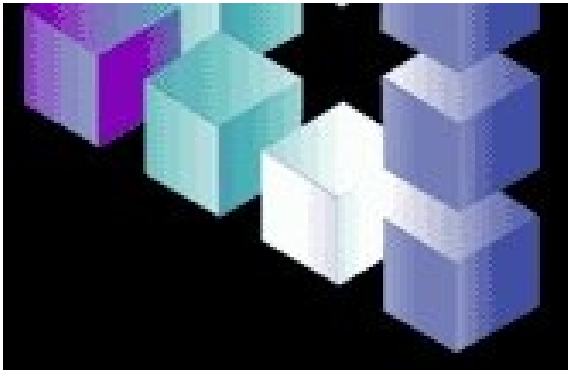
Although **enormously** simplified, this illustrates some of the requirements for a multi-level perceptual system with dynamically generated and modified contents using ontologies referring to hypothesised entities of various sorts.

Further development would require reference to 3-D structures, processes, hidden entities, affordances, other possibilities, and causal interactions increasingly remote from sensory contents.



3-D structures – and possible actions

Why visual representations are not 3-D models



This is part of a picture by Swedish artist, Oscar Reutersvaard.

It is easy to imagine 3-D processes occurring in this scene, e.g. blocks being moved, or thin objects being moved through the spaces between blocks, etc.

2-D and 3-D Qualia: more of Reutersvaard

The previous picture is repeated on the upper right.

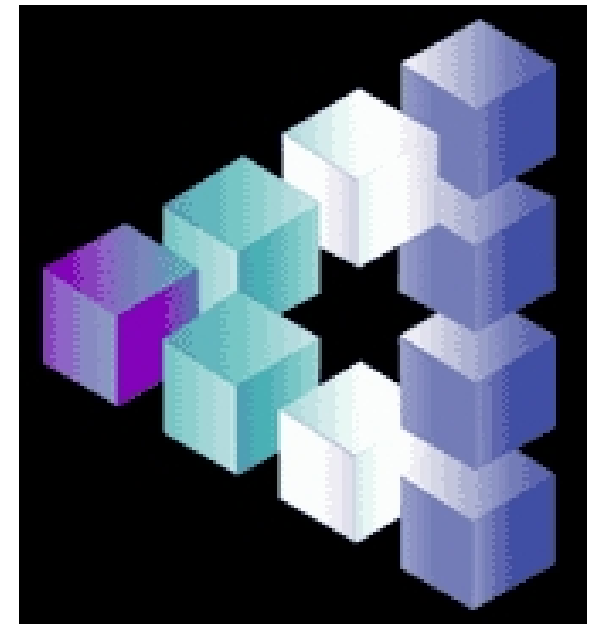
But in the context of the larger picture below, it is no longer possible to build a model of the scene depicted, for the whole scene is impossible, and no model is possible.

Yet we see a whole scene, and any portion remaining after removing two adjacent blocks is consistent.

That implies that our perception of 3-D structure here does not use a model, but [a form of representation that can express something inconsistent](#) – as sentences can.

That could be a collection of descriptions of relationships between parts of the scene.

Although it is not possible to build a 3-D model that is actually inconsistent (as opposed to appearing inconsistent), it is possible to build a collection of descriptions of relationships that forms an inconsistent totality.



Types of architecture: The CogAff Schema

Newer more complex mechanisms evolved after older simpler ones, and made use of them, forming a layered architecture.

This applies to perception and action subsystems as well as more central processing.

By dividing sub-mechanisms into “Perception”, “Central Processing” and “Actions”, and also according to whether they are “Purely reactive” or including some “Deliberative capabilities” (thinking about past, remote places, and possible futures) or including some “Meta-management” (self-monitoring, or self-modulating) mechanisms, we get a 3x3 grid of possible components (some boxes may be empty in some instances):

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

For more on this see the Cognition and Affect project papers:

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

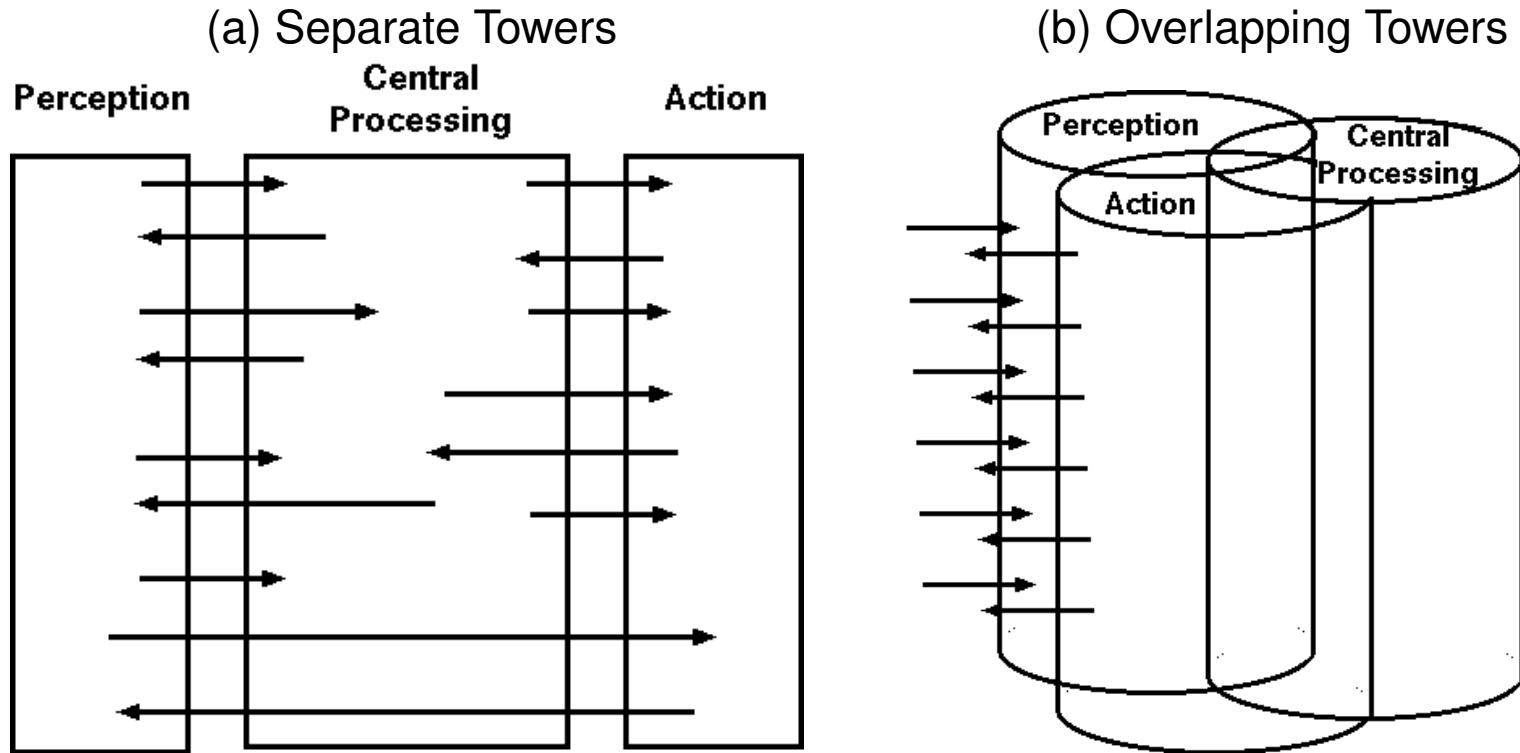
NB: this is not an architecture. I call it the “**CogAff Architecture Schema**” because it can be instantiated in many different ways, producing different I-P architectures meeting different sets of requirements.

The 3x3 structure is just a convenient introduction to a more complex collection of possible forms of architecture, possibly with more layers of sophistication, and more internal and sensory motor divisions.

Ways of slicing information-processing architectures

Thanks to Nils Nilsson (Nilsson, 1998): Towers and layers

Triple Towers (two ways of thinking about them):



Patterns of information flow may be different in different cases.

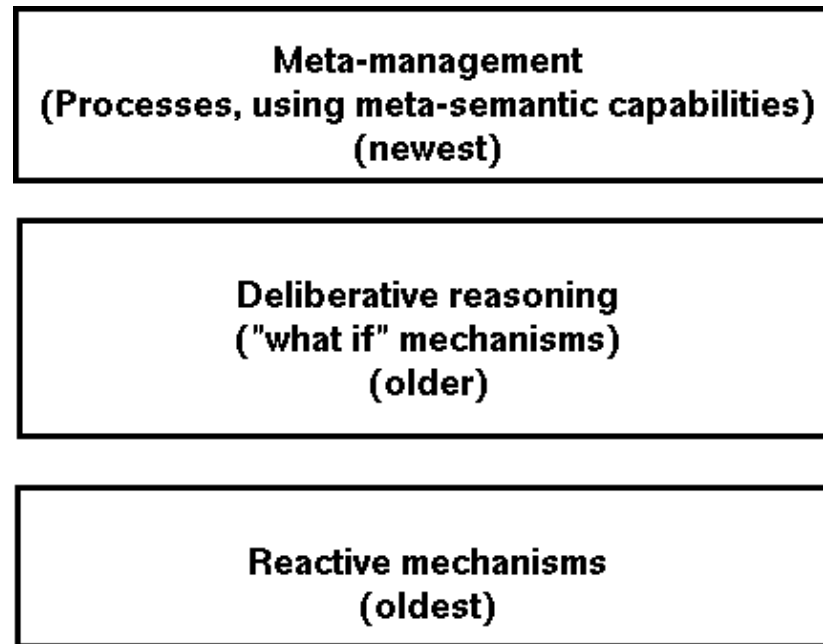
J.J. Gibson and others have emphasised the need for the right hand view of perception: perception is generally active, not passive reception.

E.g. human visual perception would be very different without control of vergence, saccades, head movements, etc. Haptic perception often depends crucially on moving a hand on the sensed object.

Three towers are not enough for the variety of functions. More on this below...

Ways of slicing information-processing architectures

Triple Layers (oldest and simplest(?) layers at the bottom):



There may be many intermediate cases, and some sub-systems need not fit neatly into one layer or one tower, or one layer-tower intersection box (e.g. alarm mechanisms).

(Compare Minsky's 6 layers – (Minsky, 2006) adding more horizontal slices.)

Many researchers in AI/Cognitive science focus on only one layer, or possibly two, often differentiated not by function but by mechanism, e.g. symbolic vs neural.

It is important to make full use of analyses of **information processing functions (uses, requirements)** for comparing design alternatives.

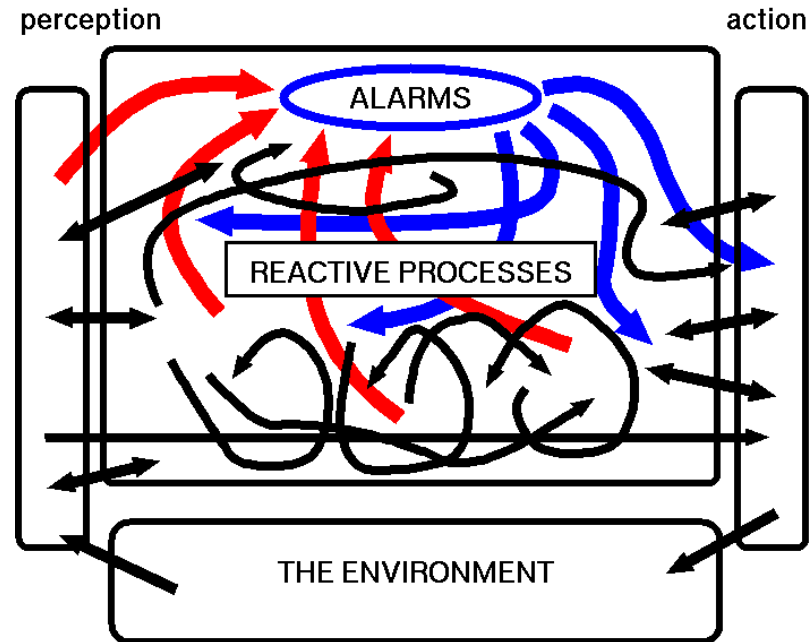
There are usually trade-offs: no single best design for any environment.

Special case of CogAff Schema: only reactive control layers

A purely reactive architecture lacks the deliberative and meta-management layers.

But it can still be very complex with many concurrently active (reactive) subsystems.

It seems likely that all microbes and insects (and most other invertebrates?) are like this:



Purely reactive mechanisms (no consideration of “what if” or “why not”), plus “alarm” mechanisms that detect dangers and opportunities that require immediate rapid re-direction of processing:

If you touch a woodlouse crawling along it will rapidly curl up: a defensive alarm mechanism grabs control?

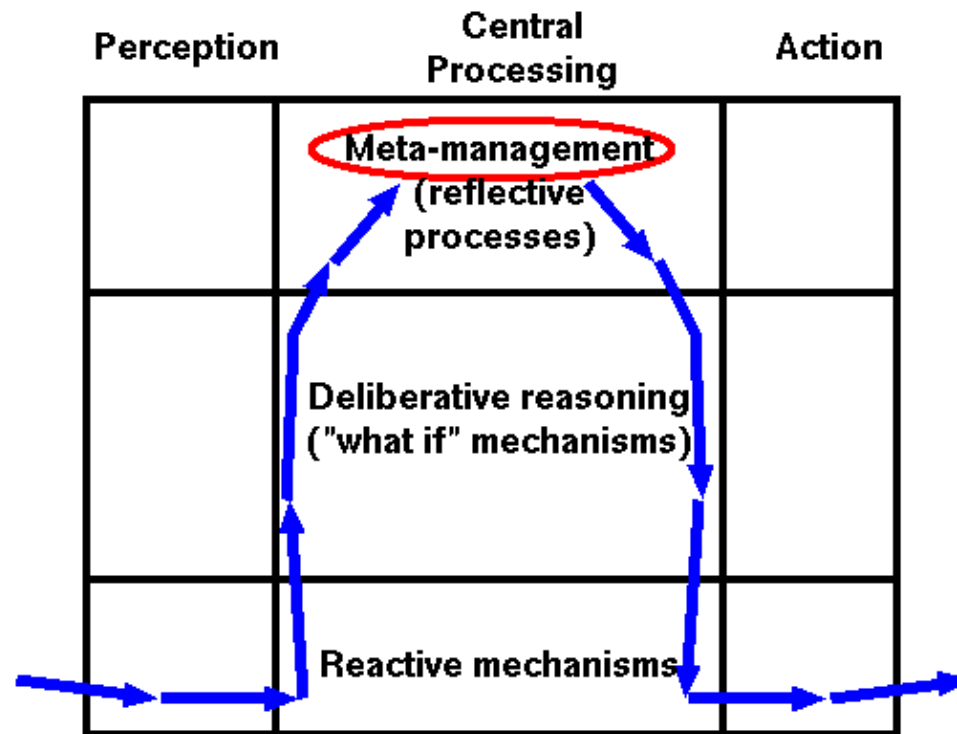
A fly that detects the approaching swatter will immediately abandon finishing its lunch and buzz off.

Some forms of adaptation or learning may modify the reactive processes, e.g. by strengthening or weakening connections, as a result of positive or negative reinforcement.

Another special case of the CogAff Schema

The Omega architecture

E.g. “contention scheduling” systems. (Cooper & Shallice, 2000)



“Omega”-type architectures, like many others, use “peephole” perception and action, restricting input/output processes to very low level sensing and acting signals.

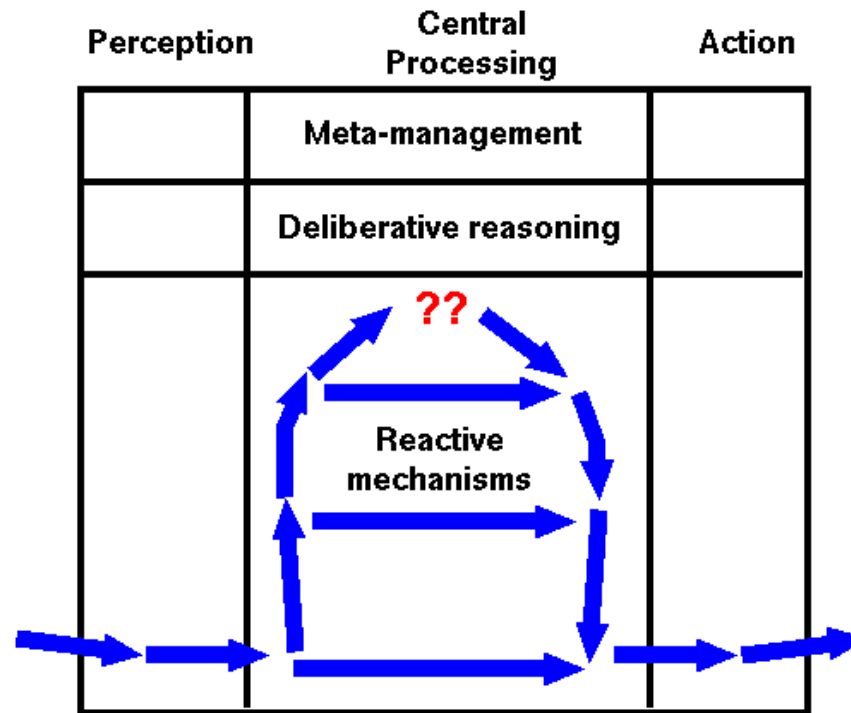
We can get some clues as to what is missing from a design by studying the changing perceptual experiences caused by ambiguous figures, as shown previously.

E.g. support for the ability to process information about possibilities and affordances may require special mechanisms for hypothetical reasoning.

Brooks-like subsumption

Another special sub-case of the CogAff architecture schema:
Brooks-type subsumption architecture (purely reactive)

(Brooks, 1986). <http://people.csail.mit.edu/brooks/papers/AIM-864.pdf>



If necessary, non-reactive layers could be added to the subsumption designs – e.g. with deliberative capabilities of various types.

See:

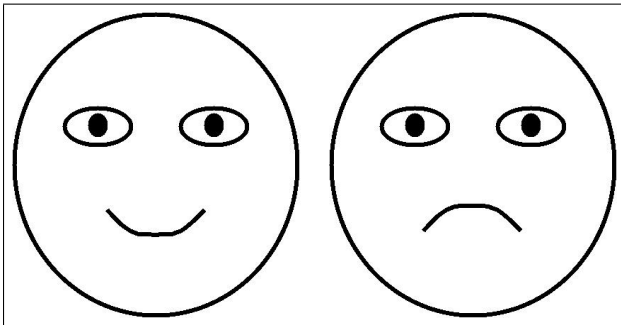
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604>,

“Requirements for a Fully Deliberative Architecture (Or component of an architecture)”

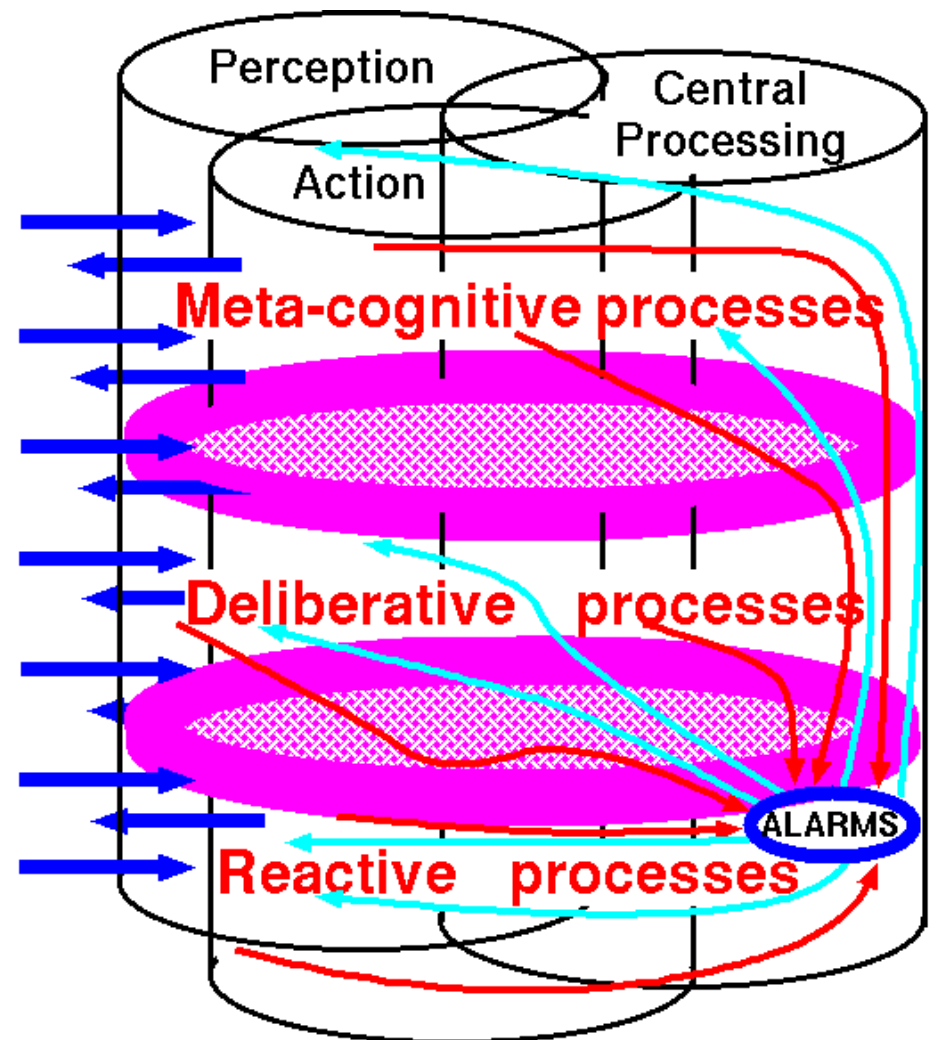
Another view of the CogAff architecture schema

This is similar to the previous “overlapping towers” depiction, but indicates (crudely) the division into evolutionary layers, comprising: reactive, deliberative, and meta-semantic/meta-cognitive processes, and possibly also an “alarm” subsystem.

The “alarm” mechanism can have both input signals from and output signals to any other subsystem that might get into a state requiring urgent action (e.g. fighting, fleeing, feeding, freezing, or mating).



The appearance of emotional states in the two faces above indicates the need for perceptual and action mechanisms to be closely connected with the meta-cognitive system, and not only with mechanisms for interacting with the physical environment.



(Thanks to Dean Petters for suggesting the diagram)

Why should a system have only one architecture?

Current computing systems have multiple architectural layers: mostly virtual machinery.

Perhaps biological evolution also produced layers of virtual machinery?

They don't all seem to be there at birth.

Should there be innate specifications for the architectures grown?

Or should it be a combined genetic and epigenetic process, as illustrated previously (Chappell & Sloman, 2007)

New layers may be grown partly under the influence of

– the problems posed by the environment

and

– the opportunities provided by the environment

as well as the genetically specified bootstrapping competences.

So an individual may need different information-processing architectures at different times.

I think this is consistent with the ideas in (Karmiloff-Smith, 1992)

Discussed in <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/beyond-modularity.html>

Contrast I-P requirements for: larva → pupa → butterfly stages of development.

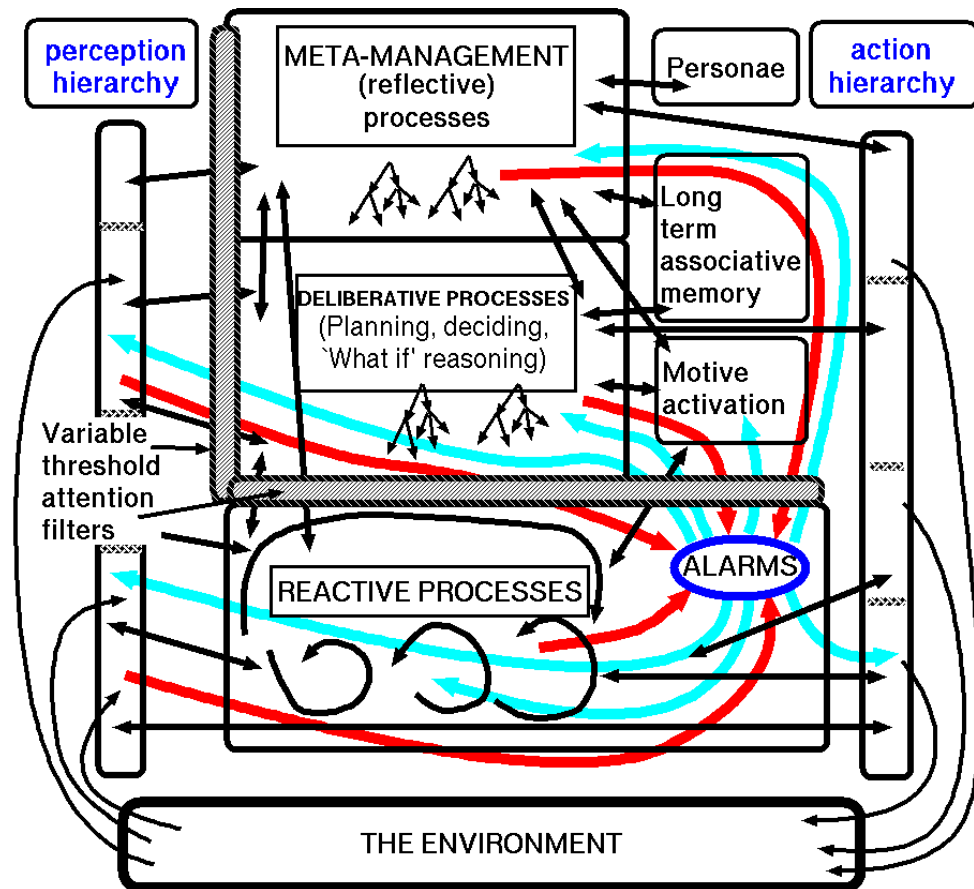
(All must be specified in the butterfly genome.)

H-Cogaff: example multi-layered architecture (much more work on this sort of thing is needed)

This is another special case of the CogAff schema. (Wright et al., 1996) (Sloman & colleagues, 2010)

Such an architecture

- has to somehow be specified in the genome, at least partly schematically
- has to grow itself, possibly over many years in some species
- includes major components that are not physical mechanisms, but instead use virtual machinery, and therefore cannot be easily detected by physical observations, or physical measuring devices (e.g. brain scanners)



The conjectured H-Cogaff (Human-Cogaff) architecture

See the web site: <http://www.cs.bham.ac.uk/research/cogaff/>

New multi-disciplinary approaches are needed for understanding the requirements, constructing explanatory designs, testing theories, working out practical implications, e.g. for psychology, for psychotherapy, for education, for social policy, and for building intelligent machines.

Meta-morphogenesis:

For any biological changes B1, B2, etc., and for any environmental changes E1, E2, etc., there can be influences of the forms

- E changes B (environmental change influences genome)
- B changes E (evolved organisms influence environments)
- E_i changes E_j (one environmental change influences another)
- B_i changes B_j (new biological features lead to further biological changes)
- Combinations of E_i, B_i, B_j, ... cause changes in B, E etc., etc.

Meta-Morphogenesis (MM):

Things that cause changes can produce new things that cause changes.

Old phenomena may be produced in new ways

e.g. information acquired and ways of acquiring and using information can change.

Often new mechanisms can produce new biological phenomena

e.g. organisms that can discover what they have learnt.

organisms that make and use mathematical discoveries.

In particular, most 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.

NOTE: Real and artificial evolution

The biological processes are very different from artificial evolutionary computation (GA, GP, etc.) with a fixed evaluation function, often used to solve engineering problems.

One reason:

Evaluation in natural evolution keeps changing, as environments change.

Moreover, there is not a single scalar evaluation function: different designs have different costs and benefits and evolution allows exploration of good designs in parallel.

(Evolution is a satisficer, not an optimizer – H.A. Simon)

Non-trivial, non-definitional ontological extension

Evolution, development, learning and social/cultural changes can all produce extensions to the ontologies used by organisms.

The variety of types of entity, property, relationship, process, structure, causal interaction that can be referred to by information-processing mechanisms in organisms can be extended by

- changes in the genome,
- developmental processes,
- individual learning,
- processes involving collections of interacting individuals (social/cultural processes).

“Trivial” ontological extensions introduce new terms that are **definable** in terms of old ones

E.g. A **pentagon** is a **polygon with five sides**.

“Non-trivial” extensions introduce new terms not explicitly definable in terms of older ones.

E.g. Newton’s concept of mass, concept of a gene, a child’s concepts of the kinds of **stuff** of which objects are composed, etc. and concepts used in a robot doing SLAM (Simultaneous Localisation And Mapping), which refer beyond the concepts used in sensors.

New terms are typically **implicitly, partially, defined by the theories that use them**.

So theories and measurement techniques 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. the child needs an **exosomatic** ontology referring to things in the environment.

(This is impossible according to concept empiricism and “symbol grounding” theory.)

When/How did use of exosomatic ontologies evolve?

Ontologies, evolved, or just learnt?

The first exosomatic ontologies were probably produced by evolution,

E.g. organisms that construct primitive maps recording spatial occupancy (i.e. not recording input/output signal patterns), as a result of exploring an environment.

Later, mechanisms developed enabling individuals to create their own ontologies triggered by interacting with the environment.

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)

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

The process may involve stimulating learners to create their own explanatory theories. (Conscious or unconscious scaffolding.)

NB: Cultural transmission may not be possible by explicit definition.

NB: Not all culturally developed ontologies are good ones (e.g. religious ontologies can be positively harmful – mind-binding).

See also:

The Computer Revolution in Philosophy, 1978. Now online, revised, especially chapter 2.

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

We learn more from trajectories than from cases

CONJECTURE: AI researchers (and others) can learn more if, instead of simply trying to build or study special cases, they study both evolutionary and developmental precursors, and alternative evolutionary trajectories based on common mechanisms.

Compare “abnormal” learning/development trajectories:

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

What do infants, toddlers, and other animals actually learn?

How do they learn it?

(what information processing mechanisms are required?)

Young primates (including humans) spend a great deal of infancy learning about kinds of matter by manipulating things

skin, hair, lips, teeth, tongue, other parts of bodies – their own and others – different sorts of food, many other things in the environment... rigid, impenetrable, bendable, stretchable, squishy, sticky, smooth, ...

Does anyone know the ontology for kinds of stuff, and kinds of process involving kinds of stuff, developed by a six month old infant, or a crow, orangutan, elephant, or octopus?

Very few roboticists have attended to this: they mostly just assume that everything a robot has to interact with is going to be rigid, or possibly hinged, like doors and door handles, or able to slide if pushed or pulled, like things on a table, or drawers.

How many soccer playing robots understand differences between a ball rolling, skidding, moving through air, or being continuously pushed?

If we study things at the wrong level of generality we'll fail to understand what's going on in the examples we investigate.

(However work is beginning on robots that fold, squeeze, prod, etc. non-rigid objects.)

Trajectories in design space and niche space

For particular organisms, or classes of organisms (species) the threats and opportunities in the environment – the sets of requirements – can change over time, producing changes in the niches.

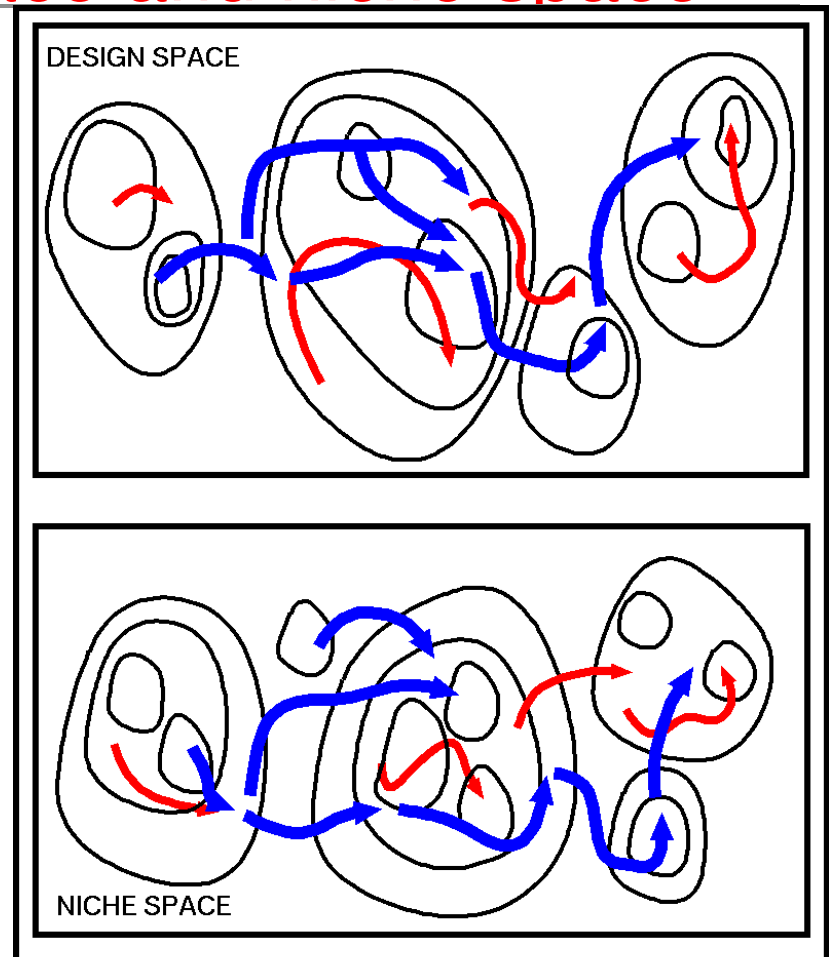
The lower half of the figure shows how such changes can constitute **trajectories** in design space for an individual or species.

The upper half of the figure shows how the designs, and therefore the competences and the behaviours of individuals can change over time, some of the changes being physical including changes in size and strength, while other changes involve the information acquired and the competences developed by the individual. These changes correspond to individual trajectories (i-trajectories) in design space.

In addition to changes in individuals represented by individual trajectories (i-trajectories) there will, over longer time scales, be changes in designs of species, namely evolutionary trajectories (e-trajectories).

Both i-trajectories and e-trajectories may have discontinuities, as explained previously, whereas the gaps in e-trajectories (species changes) will typically be larger.

There are likely relationships between e-trajectories and i-trajectories not shown, e.g. many short i-trajectories roughly aligned with longer e-trajectories. (I don't intend to claim that i-trajectories exactly follow e-trajectories.)

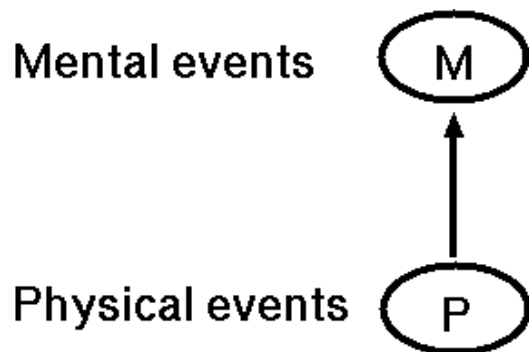


The 20th C Philosophical breakthrough: Virtual machinery

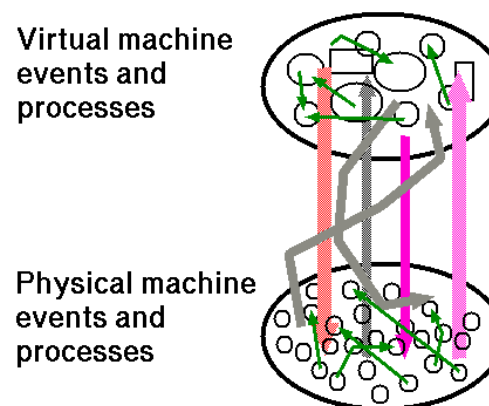
Brief introduction to the philosophical significance of the technology of virtual machinery (not virtual reality) developed since around 1950:

Processes and events in running virtual machines can be causes and effects, despite being implemented in deterministic physical mechanisms.

Bad model:



Good model:



The erroneous picture on the left implies that there is only **one-way** causation from physical to mental (epiphenomenalism).

The upper layer on the right does not indicate state-transitions, but an architecture composed of multiple concurrently active virtual machines (not necessarily synchronised).

As the picture on right indicates, we need to think about running virtual machinery that co-exists with, **and influences**, underlying physical machinery, **which it helps to control**, even though the virtual machinery is all **fully implemented in the physical machinery**.

I.e. running concurrent software can cause changes in physical hardware, just as increases in poverty can cause increases in crimes involving movement of stolen objects.

For more on [virtual machine functionalism](#) and [virtual machine supervenience](#) see (Sloman, 2010a) and <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html>

The danger of Ptolemaic theories

We are in serious danger of concentrating on what's observable and measurable.

But that can lead to Ptolemaic theories.

Those are concerned with explaining relatively easily available observations.

But may miss deep and powerful generalisations and the underlying mechanisms.

The need for comparative studies: controlled and observational

We need a comparative mindset (often looking backwards, forwards and sideways) because we can't really understand how a complex mechanism works without understanding what difference various changes would make:

- removing bits
- adding bits,
- replacing bits with alternative sub-mechanisms
- altering connections (information flow)
- using a very different information-processing architecture.

Evolution explores and compares alternatives, but without understanding them.

Many researchers on vision, language, learning, problem solving, planning, emotion, etc. try to understand only fragments of the whole picture, a bit like time-travelling medieval scholars trying to work out how what we type in a computer program can produce the observed behavioural competences of the machine, without bothering to learn anything about how to design, debug, or extend such machines.

We can build on the work done by evolution and try to understand.

Some notes on the need to relate research on autism to collection of questions not about autism illustrate this point:

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

Autistic Information Processing

Steps toward a generative theory of information-processing abnormalities.

Further Issues That Need To Be Addressed

There's lots more to be said about transitions in evolution, development, learning, culture, etc., leading to the following (some discussed further in online papers):

- ability to think about **particular** objects, agents, places, times, events, processes, etc.
- using a perceptual ontology that includes **kinds of stuff** (skin, hair, water, meat, mud, ...)
<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#babystuff>
- alternatives to use of global metrical coordinate systems for spatial structures and relationships: e.g. use **networks of partial orderings** (Contrast (I. Johansson, 2008))
- **alternatives to the use of probability distributions** to cope with noise and uncertainty, e.g. **use more abstract representations** that capture only what is certain about structures, relationships and processes, e.g. X definitely getting closer to Y, though exact speed and trajectory are not known:
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0702>
Example: Finding the right ontology for predicting affordance changes with certainty.
- perception of and reasoning about **processes of many kinds**
including not just spatial change, but also causal and epistemic changes, and intentional actions.
- perception of and reasoning about **multi-strand relationships**
(not only whole objects have perceivable changing relationships, but also parts and fragments of those objects, and parts of parts, etc.)
- perception of **multi-strand processes**
(processes in which there are multiple concurrent sub-processes, e.g. changing relationships.)
- alternatives to the mistaken view that symbols need **“grounding”** (concept empiricism):
<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models>
- **learning** to perceive many things
how changes in ontologies, forms of representation, architectures, and also changes in **uses** of visual information can occur during evolution, development, learning and cultural change.
- Many aspects of social/cultural evolution, learning, interaction, ...

A sample of some particularly hard problems

These are some of the things I have been thinking about over many years and still feel that we have made only a little progress compared with what needs to be done.

- What are the forms of representation, reasoning, inference, control, used in organisms, including those with and those without brains.
- What is the role of chemistry in animal mentality and how much of that role could be replicated in systems based on digital computers: are there aspects that cannot be? (What about effects of neurotransmitters, hormones, drugs, addictive substances?)
- What is it about the physics and chemistry of our universe that allows ever more sophisticated information processing systems to emerge spontaneously in some physical configurations? Is Kauffman's answer too general?
- What were the biological mechanisms that allowed our ancestors to start making mathematical discoveries leading to Euclidean geometry, understanding the natural numbers, the real continuum, etc., before there were mathematics teachers presenting such things? In particular, how does human/animal **spatial** reasoning work? (Sloman, 1971)
- What sorts of mechanisms allow causation in virtual machinery to have features experienced by humans, e.g. conflicts of desire that remain after decisions have been made? <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk86>
- What are the boot-strapping mechanisms for I-P in babies/toddlers? (Chappell & Sloman, 2007)
- What is the architecture of a human/animal visual system, how does it grow, and how does it work so fast? <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk88>
- How can qualia go unnoticed? <http://tinyurl.com/CogMisc/unconscious-seeing.html>
- What is finding something funny, and what mechanisms make that possible?

Philosophy is one of the relevant disciplines

For example all these are relevant to the Meta-Morphogenesis project:

- Conceptual analysis
- Metaphysics
- Kinds of meaning – semantics, ontologies
- Philosophy of mind
- Philosophy of science
- Philosophy of mathematics
- What are affective states and processes and what makes them possible?
- Logical geography vs logical topography (Sloman, 2007a)
- What is a form of representation?
Donald Peterson, Ed., *Forms of representation: an interdisciplinary theme for cognitive science*
Intellect Books, 1996,
- What forms of representation can a genome use?
- Kinds of machinery : e.g. physical, virtual
- Kinds of causation: e.g. in virtual machinery
- What sorts of things can biological evolution do (in principle)?
- How did ethical capabilities arise in evolution and why is that irrelevant to answering ethical questions?
- What is finding something funny?

Why was my prediction in (Sloman, 1978), about how soon philosophers would regard knowledge of AI as essential for doing philosophy, so wrong?

(One problem is that I share McCarthy's and Minsky's very broad vision of AI – unlike some AI researchers.) I mistakenly thought it would take about 10-15 years for philosophers to notice. Many still have not.

The end – but really just the beginning

- Turing's work showed how a precisely defined class of fairly simple machines performing **discrete** operations on an unbounded linear tape could accomplish a very surprisingly large and varied collection of logical and mathematical tasks.
- Later work in computer science and engineering showed how such machines, connected to sensors and motors via analog/digital interfaces, could control a huge and very varied collection of machinery acting in natural and artificial environments – with interfaces to physical objects and machines, and (indirectly) to human minds.
- But long before that, biological evolution had already **spontaneously** produced information-processing machinery performing an even richer collection of control functions in myriad organisms, also using a limited class of basic machinery – but not as limited as Turing machines: namely chemical machines that can be built using approximately 112 chemical elements, themselves built from more fundamental(?) components <http://www.chemicalelements.com/>
- We still lack a theory of the information-processing capabilities of chemical machines comparable to our theory of Turing machines (and their equivalents), though a possible route to such a theory is a deep and broad survey of types of information-processing such chemical machines can do, including ways in which their interactions can produce new, more complex, instances, including new kinds of virtual machinery: at least we'll then have a better idea of what the missing theory needs to be able to explain.
- A crucial difference: **Turing machines require something very different (A/D and D/A converters) to provide interfaces to the environment, whereas chemical computers can directly interface with a physical/chemical environment.**
- In these slides, part of the Meta-Morphogenesis project, I've tried to begin the task of producing a survey of transitions in information-processing since pre-biotic molecules, and analysing **requirements** for the still missing explanatory theory. Please join in.
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html>

The next few slides are an optional appendix

(Followed by references.)

The power of special purpose computers I

Suppose a red rope and a blue rope lie on a plane surface, as indicated in the picture, and you are asked to decide which is shorter

You could use a ruler to measure small approximately straight portions of each rope, and then sum all the measurements, or take a digital photograph, and run a line-finder then sum edge-feature lengths.

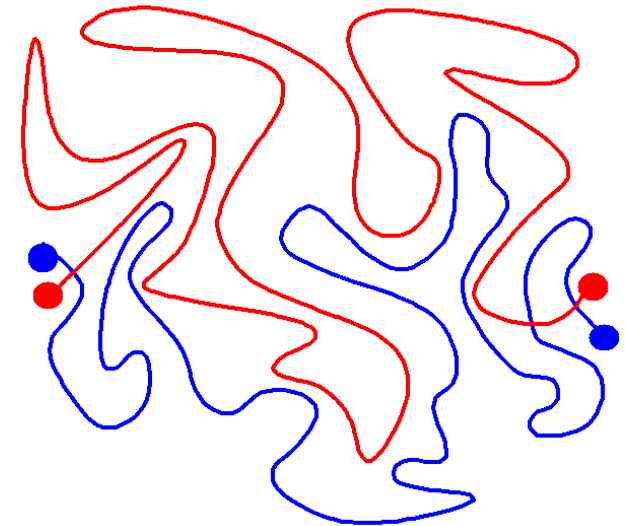
Before such techniques existed, a fast way to find the answer could be used. If the ropes can be moved: let them do the calculation:

Hold a red and blue end together and hold the other red and blue ends together and then move the two pairs of ends apart.

The first rope to become straight will be the shorter.

(What assumptions make that necessarily true?)

The measuring process produces more information (e.g. the difference in length) – but irrelevant to the task.



Alternative mechanisms are possible (with different preconditions):

- Instead of flexible ropes, tubes with the same internal diameter throughout, could have water pumped into one end of each tube at the same rate, to determine which tube first emits water at the far end.
- Or do one at a time, and measure the times taken for water to emerge.
- If there are chemical structures that use catalytic mechanisms to transmit a change at a fixed speed along a linear molecular structure, then the change could be initiated at one end of each chain, with detectors at the opposite ends determining which one altered first, triggering a decision.

All of these mechanisms depend on propagation of causation through space over time, and the availability of mechanisms to detect, compare and use the resulting states.

(Fast local processing may need even faster global signalling of results, e.g. to block rival paths.)

If brains could use such mechanisms, that might provide abilities to solve mathematical problems, but without mathematical understanding of the problems and the solutions.

The power of special purpose computers II

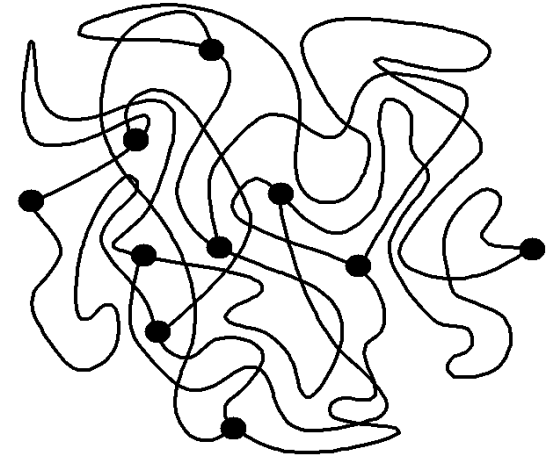
Now consider a more general map-based problem: there are towns indicated by black circles, connected by curved roads shown as black lines, and where there are no towns at crossings, i.e. where lines cross without a circle, there are no road junctions, only bridges, or tunnels, precluding turning from one road into another.

How can we find the shortest route between two towns?

One way: measure all the lengths of the road segments linking towns then search for the shortest route, using a standard AI planning algorithm.

A familiar alternative would be to make a string model of all the routes, labelling each string to indicate its road, with labelled knots for towns, then gently pull apart two knots to find the shortest route between them.

In this case, physical mechanisms acting in parallel compute the shortest route, as long as no tangled loops or knots form.



Alternative methods could use tubes and fluid pumped in at a start node, flowing at a uniform speed along all routes. As soon as fluid reaches a target node, tracing back would show the shortest route.

Using a network of linked polymers as a map, chemical mechanisms could propagate signals outwards from a start node along all routes. A detector at the target node receiving such a signal could stop all propagation, and initiate another process to record the route used. The time to find the shortest route would depend only on the route, not the size of the network, which might cause AI search algorithms to explode, if there are millions of slightly longer routes, or routes that don't reach the target.

It's possible to simulate this process on a computer, but the parallel (breadth-first) search has a cost.

Do molecular mechanisms provide a (huge) space of special purpose information-processors for solving problems whose complexity would defeat, or bog-down, Turing machines, or modern computers?

I don't know: but **if** it is possible, evolution may have discovered such mechanisms, for use within animals.

(Compare control of global behaviours by external stigmergy in social insects and microbes.)

The power of special purpose computers III

Previous slides do not demonstrate that special purpose molecular computations are **actually** used in organisms to perform tasks that have high complexity for Turing machines or computer based search algorithms.

They merely raise the question:

Organisms use the huge and varied information-processing potential in chemistry, for processes of growth, digestion, metabolism, energy transport and conversion, tissue repair, control of infection, reproduction, etc. – do they also use it for many of the capabilities that we have been trying to give computers by giving them AI algorithms?

As more and more is learnt about the power of chemical information processing e.g. from studies of reproduction, cell growth, brain construction, reproductive malfunctions, mechanisms by which microbe invaders interact with their hosts, mechanisms by which hosts respond to infection, the growth of cancer cells, the effects of drugs, and other studies,

we may find that functions that were previously assumed to be performed by neural networks are, in some cases, performed in a different way.

(Possibly refuting calculations of computer power required to replicate brain functions.)

There's a growing literature on chemical computation, e.g. (Berry & Boudol, 1990; Kovac, 2006; Hanczyc & Ikegami, 2010)

DISCLAIMER

I am not saying that the existence of problem-solving mechanisms of the sorts described would **suffice** to mimic human reasoning about spatial problems.

E.g. many humans both **discover and use solution methods or mechanisms** and also **understand why they work**, for example as a result of proving mathematical theorems. Compare (Jamnik, Bundy, & Green, 1999)

Understanding requires meta-cognitive mechanisms operating on the problem-solving systems, using abstraction and reasoning, illustrated in connection with simple theorems in Euclidean geometry here:

<http://tinyurl.com/CogMisc/triangle-theorem.html>

Abramsky on What is a process?

In his panel contribution at the Manchester Turing conference in June 2012, Samson Abramsky talked about “The Big Questions in Computation, Intelligence and Life”. A video of his presentation is here

<http://videlectures.net/turing100-abramsky-big-questions/>

He identifies as a major open question how to give a general enough characterisation of the notion of a “process” – comparable in power and generality to the answers given in the last century to the question “What is computation?”

My slides are not an attempt to answer the question “What is a process?” but groundwork clarifying [requirements for a satisfactory general answer](#).

A general answer needs to include not only processes of the sorts previously studied, e.g. in physics, chemistry, mathematics and computer science, but also processes in biology on many scales from molecular interactions within organisms to changes in ecosystems where the ecosystems nowadays also include effects of processes that depend on human societies, such as economic, political, cultural, and even military processes.

Processes involving use of information, not only about [what does or doesn't exist](#), but also information about [things to achieve](#), and information about [how to maintain, achieve or prevent things](#), have increasingly dominated changes on this planet.

(That's a temporary phase, however, since eventually some physical process will shut all this down, whether it's the death of our sun or major planetary collision, or inter-galactic event, or...).

Biological evolution has changed our planet so that one process P1 can be changed by information about another P2, in a way that depends not only on the content of the information and the capabilities of P1, but also P1's goals and other information, well as constraints on P1. When many processes interact ... some of them may involve brains.

Useful Background Reading (A tiny sample)

Paul Davies (Physicist)

The Fifth Miracle: The Search for the Origin and Meaning of Life, Simon & Schuster, 1999.

Not a collection of religious mush, as you might expect from the title. Explicitly addresses the general question: “Where does the biological information come from?”, but doesn’t give an answer.

Mark Pallen (Biologist, University of Birmingham)

The Rough Guide to Evolution

Rough Guides (2009) See also: <http://roughguidetoevolution.blogspot.co.uk/>

A useful overview of Darwin’s ideas, some of the reactions to them, and influences. Apart from genetic information, he doesn’t seem to think it important to raise questions about information processing.

[But that is essential to fully refute intelligent design and creationist challenges, and answer scientific and philosophical questions about “The explanatory gap” \(Huxley\)](#)

Richard Dawkins (Dawkins, 1076, 1982)

Although Dawkins does not talk much about information and information processing that’s what he is writing about. A trawl through his publications would probably be an excellent source of examples of transitions in information processing.

Tibor Ganti, The Principles of life (translated OUP, 2003)

Margaret A. Boden (Boden, 2006)

Mind As Machine: A history of Cognitive Science (Vols 1–2), OUP, 2006

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/boden-mindasmachine.html>

Especially Chapters 2 and 15. Look up in index: Brian Goodwin, Stuart Kauffman, John Maynard Smith, Darcy Thompson, Waddington,

Arnold Trehub’s 1991 book, *The Cognitive Brain*, MIT Press,

Is not as well known as it should be. It assembles many deep issues ignored by others, and proposes theories spanning chemistry of neurones to high level cognition. Now online here: <http://www.people.umass.edu/trehub/>

Keith D. Farnsworth, John Nelson & Carlos Gershenson

Living is information processing; from molecules to global systems.

(Farnsworth, Nelson, & Gershenson, 2012) (I am not sure they are using the right concept of information.)

My draft notes on transitions in information processing in evolution, development, learning, social interactions, symbiosis, etc. are here: (with over 80 transition types/subtypes by Jan 2012):

<http://tinyurl.com/CogMisc/evolution-info-transitions.html>

Gert Korthof's web site

This amazing web site is full of treasure, though I have so far sampled only a tiny subset.

It touches on the topic of information in various places, though I have not yet found an unambiguous reference to the idea of organisms making use of information about the environment, their current state, their needs, other organisms, possible goals, possible actions, etc.

<http://wasdarwinwrong.com/>

Top level: "Towards The Third Evolutionary Synthesis"

<http://wasdarwinwrong.com/>

Introduction to the Evolution literature

<http://wasdarwinwrong.com/korthof59.htm>

Independent origin and the facts of life

There's lots more.

Related online materials

- [1] 'What's information, for an organism or intelligent machine? How can a machine or organism mean?' In **Information and Computation** Eds G. Dodig-Crnkovic and M. Burgin, pp. 393–438, 2011.
<http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#905>
- [2] http://en.wikipedia.org/wiki/R/K_selection_theory R/K evolutionary tradeoff.
- [3] Requirements for a Fully Deliberative Architecture (Or component of an architecture)
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604>
- [4] If learning maths requires a teacher, where did the first teachers come from?
<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk79>
- [5] Annette Karmiloff-Smith,
Beyond Modularity: A Developmental Perspective on Cognitive Science, MIT Press, 1992,
(Discussed in <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/beyond-modularity.html>)
- [6] Evolution of mind as a feat of computer systems engineering:
Lessons from decades of development of self-monitoring virtual machinery.
<http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1103>
Virtual Machinery and Evolution of Mind (Part 3)
Meta-Morphogenesis: Evolution of Information-Processing Machinery
<http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1106d>
- [7] There are strong connections with Marvin Minsky's work, which focuses on the special case of a **human** architecture, E.g. *The Emotion Machine*. <http://web.media.mit.edu/~minsky/>
- [8] John McCarthy's 'The Well-Designed Child' (McCarthy, 2008), overlaps with this work in several respects, but, like Minsky, focuses mainly on the human case.
- [9] There is much work on architectures in AI and Cognitive Science, and many different architectures are proposed either because people ignore previous work or because different researchers focus on different subsets of requirements.
A partial survey of architectures is available at <http://bicasociety.org/cogarch/>
- [10] My work with colleagues at Birmingham on requirements not just for one architecture, but for a space of biological architectures of many kinds (the CogAff project) can be found here <http://tinyurl.com/BhamCog/#overview>
- [11] Draft overview of the Meta-Morphogenesis project
<http://tinyurl.com/CogMisc/meta-morphogenesis.html>
For related PDF presentations (including this one) see <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/>
Some also in flash format on Slideshare: <http://www.slideshare.net/asloman/presentations>

Bibliography

References

- Adolph, K. (2005). Learning to learn in the development of action. In J. Lockman, J. Reiser, & C. A. Nelson (Eds.), *Action as an organizer of perception and cognition during learning and development: Minnesota Symposium on Child Development* (Vol. 33, pp. 91–122). Mahwah, NJ: Erlbaum. Available from <http://www.psych.nyu.edu/adolph/PDFs/MinnSymp2005.pdf>
- Baldassarre, G. (2011). What are intrinsic motivations? a biological perspective. In *Development and learning (icdl), 2011 ieee international conference on* (Vol. 2, p. 1-8).
- Beaudoin, L. (1994). *Goal processing in autonomous agents*. Unpublished doctoral dissertation, School of Computer Science, The University of Birmingham, Birmingham, UK. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#38>
- Beaudoin, L., & Sloman, A. (1993). A study of motive processing and attention. In A. Sloman, D. Hogg, G. Humphreys, D. Partridge, & A. Ramsay (Eds.), *Prospects for artificial intelligence* (pp. 229–238). Amsterdam: IOS Press. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#16>
- Berry, G., & Boudol, G. (1990). The chemical abstract machine. In *Proceedings of the 17th acm sigplan-sigact symposium on principles of programming languages* (pp. 81–94). New York, NY, USA: ACM. Available from <http://doi.acm.org/10.1145/96709.96717>
- Berthoz, A. (2000). *The brain's sense of movement*. London, UK: Harvard University Press.
- Boden, M. A. (2006). *Mind As Machine: A history of Cognitive Science (Vols 1–2)*. Oxford: Oxford University Press.
- Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, *RA-2*, 14–23. Available from <http://people.csail.mit.edu/brooks/papers/AIM-864.pdf> (1)
- Carnap, R. (1947). *Meaning and necessity: a study in semantics and modal logic*. Chicago: Chicago University Press.
- Chappell, J., & Sloman, A. (2007). Natural and artificial meta-configured altricial information-processing systems. *International Journal of Unconventional Computing*, *3*(3), 211–239. (<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609>)
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Cooper, R., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, *17*(4), 297–338.
- Craik, K. (1943). *The nature of explanation*. London, New York: Cambridge University Press.
- Dawkins, R. (1976). *The Selfish Gene*. OUP.
- Dawkins, R. (1982). *The extended phenotype: The long reach of the gene*. Oxford, New York: Oxford University Press.
- Farnsworth, K., Nelson, J., & Gershenson, C. (2012). Living is information processing; from molecules to global systems. *CoRR*, *abs/1210.5908*. Available from <http://arxiv.org/abs/1210.5908>
- Fodor, J. (1975). *The Language of Thought*. Cambridge: Harvard University Press.
- Ganti, T. (2003). *The Principles of Life* (E. S. Szathmáry & J. Griesemer, Eds.). New York: OUP. (Translation of the 1971 Hungarian edition, with notes)
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Hanczyc, M. M., & Ikegami, T. (2010). Chemical Basis for Minimal Cognition. *Artificial Life*, *16*, 233–243.
- Harnad, S. (1990). The Symbol Grounding Problem. *Physica D*, *42*, 335–346.
- Heider, F., & Simmel, M. (1944.). An experimental study of apparent behaviour. *American Journal of Psychology*, *57*, 243–259.
- Jablonka, E., & Lamb, M. J. (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge MA: MIT Press.
- Jamnik, M., Bundy, A., & Green, I. (1999). On automating diagrammatic proofs of arithmetic arguments. *Journal of Logic, Language and Information*, *8*(3), 297–321.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *14*, 201–211.
- Johansson, I. (2008). Functions and Shapes in the Light of the International System of Units. *Int Ontology Metaphysics*, *9*, 93–117. Available from DOI10.1007/s12133-008-0025-z
- Kant, I. (1781). *Critique of pure reason*. London: Macmillan. (Translated (1929) by Norman Kemp Smith)

- Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.
- Kauffman, S. (1995). *At home in the universe: The search for laws of complexity*. London: Penguin Books.
- Kovac, L. (2006, June). Life, chemistry, and cognition: Conceiving life as knowledge embodied in sentient chemical systems may provide novel insights into the nature of cognition. *EMBO Report*, 7(6), 562–566. Available from <http://dx.do.org/10.1038/sj.embor.7400717>
- Maturana, H. R., & Varela, F. J. (1992). *The Tree of Knowledge: the Biological Roots of Human Understanding (Rev. ed.)*. Boston, MA: Shambala Publications, Inc.
- McCarthy, J. (2008). The well-designed child. *Artificial Intelligence*, 172(18), 2003–2014. Available from <http://www-formal.stanford.edu/jmc/child.html>
- Michotte, A. (1962). *The perception of causality*. Andover, MA: Methuen.
- Minsky, M. L. (2006). *The Emotion Machine*. New York: Pantheon.
- Nehaniv, C. L., & Dautenhahn, K. (1998). Embodiment and memories – algebras of time and history for autobiographic agents. In R. Trappl (Ed.), *Cybernetics and Systems '98, volume 2* (pp. 651–656). Austrian Soc. for Cybernetic Studies.
- Nehaniv, C. L., Polani, D., Dautenhahn, K., R. te Boekhorst, & L. Cañamero. (2002). Meaningful Information, Sensor Evolution, and the Temporal Horizon of Embodied Organisms. In R. Standish, M. A. Bedau, & H. A. Abbass (Eds.), *ALife VIII, Proc. 8th Int. Conf. on Artificial Life* (pp. 345–349). Cambridge, MA: MIT Press.
- Nilsson, N. (1998). *Artificial intelligence: A new synthesis*. San Francisco: Morgan Kaufmann.
- Simon, H. A. (1967a). Motivational and emotional controls of cognition. In H. A. Simon (Ed.), *reprinted in models of thought* (pp. 29–38). Newhaven, CT: Yale University Press.
- Simon, H. A. (1967b). Rational Choice and the Structure of the Environment. In H. A. Simon (Ed.), *Models of Thought* (pp. 20–29). Newhaven, CT: Yale University Press.
- Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press. ((Second edition 1981))
- Sloman, A. (1971). Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence. In *Proc 2nd ijcai* (pp. 209–226). London: William Kaufmann. (<http://www.cs.bham.ac.uk/research/cogaff/04.html#200407>)
- 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. (1979). The primacy of non-communicative language. In M. MacCafferty & K. Gray (Eds.), *The analysis of Meaning: Informatics 5 Proceedings ASLIB/BCS Conference, Oxford, March 1979* (pp. 1–15). London: Aslib. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#43>
- Sloman, A. (1982). Image interpretation: The way ahead? In O. Braddick & A. Sleigh. (Eds.), *Physical and Biological Processing of Images (Proceedings of an international symposium organised by The Rank Prize Funds, London, 1982.)* (pp. 380–401). Berlin: Springer-Verlag. (<http://www.cs.bham.ac.uk/research/projects/cogaff/06.html#0604>)
- Sloman, A. (1985). What enables a machine to understand? In *Proc 9th ijcai* (pp. 995–1001). Los Angeles: IJCAI. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#4>
- Sloman, A. (1986). *What Are The Purposes Of Vision?* Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/12.html#1207> (Presented at: Fyssen Foundation Vision Workshop Versailles France, March 1986, Organiser: M. Imbert)
- Sloman, A. (1989). On designing a visual system (towards a gibsonian computational model of vision). *Journal of Experimental and Theoretical AI*, 1(4), 289–337. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#7>
- Sloman, A. (1996). Actual possibilities. In L. Aiello & S. Shapiro (Eds.), *Principles of knowledge representation and reasoning: Proc. 5th int. conf. (KR '96)* (pp. 627–638). Boston, MA: Morgan Kaufmann Publishers. Available from <http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15>
- Sloman, A. (2000). Models of models of mind. In M. Lee (Ed.), *Proceedings of Symposium on How to Design a Functioning Mind, AISB'00* (pp. 1–9). Birmingham: AISB. (<http://www.cs.bham.ac.uk/research/projects/cogaff/00-02.html#56>)
- Sloman, A. (2001). Evolvable biologically plausible visual architectures. In T. Cootes & C. Taylor (Eds.), *Proceedings of British Machine Vision Conference* (pp. 313–322). Manchester: BMVA.
- Sloman, A. (2003). *The Cognition and Affect Project: Architectures, Architecture-Schemas, And The New Science of Mind*. (Tech. Rep.). Birmingham, UK:

- School of Computer Science, University of Birmingham. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/03.html#200307> ((Revised August 2008))
- 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. (2007a). *Two Notions Contrasted: 'Logical Geography' and 'Logical Topography' (Variations on a theme by Gilbert Ryle: The logical topography of 'Logical Geography')*. (Tech. Rep. No. COSY-DP-0703). Birmingham, UK: School of Computer Science, University of Birmingham. (<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0703>)
- Sloman, A. (2007b). *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. (2008a). *Evolution of minds and languages. What evolved first and develops first in children: Languages for communicating, or languages for thinking (Generalised Languages: GLs)?* (Research Note No. COSY-PR-0702). Birmingham, UK. Available from <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0702>
- Sloman, A. (2008b). *A Multi-picture Challenge for Theories of Vision* (Research Note No. COSY-PR-0801). Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk88>
- Sloman, A. (2009a). Architecture-Based Motivation vs Reward-Based Motivation. *Newsletter on Philosophy and Computers*, 09(1), 10–13. Available from <http://www.apaonline.org/publications/newsletters/v09n1-Computers.06.aspx>
- Sloman, A. (2009b). Some Requirements for Human-like Robots: Why the recent over-emphasis on embodiment has held up progress. In B. Sendhoff, E. Koerner, O. Sporns, H. Ritter, & K. Doya (Eds.), *Creating Brain-like Intelligence* (pp. 248–277). Berlin: Springer-Verlag. Available from <http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0804>
- Sloman, A. (2010a, 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. (2010b). If Learning Maths Requires a Teacher, Where did the First Teachers Come From? In A. Pease, M. Guhe, & A. Smaill (Eds.), *Proc. Int. Symp. on Mathematical Practice and Cognition, AISB 2010 Convention* (pp. 30–39). De Montfort University, Leicester. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#1001>
- Sloman, A. (2011a). 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. (2011b, 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., & Chappell, J. (2007). Computational Cognitive Epigenetics (Commentary on (Jablonka & Lamb, 2005)). *Behavioral and Brain Sciences*, 30(4), 375–6. (<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0703>)
- Sloman, A., Chrisley, R., & Scheutz, M. (2005). The architectural basis of affective states and processes. In M. Arbib & J.-M. Fellous (Eds.), *Who Needs Emotions?: The Brain Meets the Robot* (pp. 203–244). New York: Oxford University Press. (<http://www.cs.bham.ac.uk/research/cogaff/03.html#200305>)
- Sloman, A., & colleagues. (2010). *Origins and Overview of The Cognition and Affect (CogAff) Project*. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/#overview> (This web site is updated from time to time)
- Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. London: Methuen.
- Turing, A. M. (1952). The Chemical Basis Of Morphogenesis. *Phil. Trans. R. Soc. London B* 237, 237, 37–72.
- Ugur, E. (2010). *A Developmental Framework for Learning Affordances*. Unpublished doctoral dissertation, The Graduate School of Natural and Applied Sciences, Middle East Technical University, Ankara, Turkey. Available from <http://www.cns.atr.jp/~emre/papers/PhDThesis.pdf>
- Waddington, C. H. (1957). *The Strategy of the Genes*. MacMillan.
- White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, 66, 297–333. (5)

Wright, I., Sloman, A., & Beaudoin, L. (1996). Towards a design-based analysis of emotional episodes. *Philosophy Psychiatry and Psychology*, 3(2), 101–126. Available from <http://www.cs.bham.ac.uk/research/projects/cogaff/96-99.html#2>

And lots more ...