How can a planet (or a dust cloud) produce Microbes, Minds, Mathematics, Music, Marmite

(along with murder, religious bigotry, and other nastiness).

An introduction to the Meta-Morphogenesis project

(Partly inspired by Turing's work on morphogenesis)

More Information:
`http://tinyurl.com/CogMisc/misc/meta-morphogenesis.html`
`http://tinyurl.com/CogMisc/misc/evolution-info-transitions.html`

Aaron Sloman
`http://www.cs.bham.ac.uk/~axs/`
School of Computer Science, University of Birmingham.
(Philosopher in a Computer Science department)

Related Presentations (PDF):
`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/` This one is `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk107`
Also to be added to Slideshare (in flash format): `http://www.slideshare.net/asloman`

Related discussions (mostly html):
`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/`

Contents

1 Preliminaries

Thanks – AGI-2012

The Fifth Conference on Artificial General Intelligence was held at St Anne's College Oxford, in December 2012. (For details, see `http://agi-conf.org/2012/`)

I would like to thank the organisers, especially Joscha Bach and Ben Goertzel, for accepting my offer of a tutorial on Meta-Morphogenesis, and for very useful post-conference comments on my slides.

I promised, during the tutorial that I would later make my slides available, after expansion and reorganisation. This took much longer than expected because of the vast amount of material that's relevant, the diversity of topics, the unfamiliarity of the approach, the diversity of the potential audience, and the amount of literature I felt I had to read in order to evaluate and summarise related published material by others.

There's far too much for one person to cover – and the quality is very mixed!

Some time after beginning work on these slides I came to realise that the ideas of Stuart Kauffman, e.g. in (Kauffman, 1995), are closely related – mentioned again below.

I have not covered all the material I had in mind, nor provided as much concrete detail as I had hoped, nor presented it as clearly as I had hoped, especially explaining the limitations of other approaches whose benefits I hope this work subsumes (or will eventually!)

This is an interim version, last modified on April 11, 2013.

There will later also be an appendix on the same web site, with additional materials in the same format.

I shall go on making changes to these slides, and expanding the associated online web pages, especially this web page collecting examples of transitions in biological information-processing.

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

which will later have to be split into several parts. These slides include a subset of that material, re-organised.

Quotes from Pride and Prejudice by Jane Austen

Jane Austen knew a lot about human information processing as these snippets show:

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

Catherine and Lydia had information for them of a different sort.

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."

This information made Elizabeth smile, as she thought of poor Miss Bingley.

This information, however, startled Mrs. Bennet ...

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 Mr. Collins, when he returned to Longbourn to dinner, to drop no hint of what had passed before any of the family.

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

Elizabeth loved absurdities, but she had known Sir William's too long. 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 Sir William Lucas's accidental information, that Bingley'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.

and at last she was referred for the truth of every particular to Colonel Fitzwilliam himself-from whom she had previously received the information of his near concern in all his cousin's affairs,

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

Mr. Bennet had been to Epsom and Clapham, before his arrival, but without gaining any satisfactory information;

Elizabeth was at no loss to understand from whence this deference to her authority proceeded; but it was not in her power to give any information of so satisfactory a nature as the compliment deserved.

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

but to her own more extensive information, he was the person to whom the whole family were indebted

Darcy was delighted with their engagement; his friend had given him the earliest information of it.

The joy which Miss Darcy expressed on receiving similar information, was as sincere as her brother's in sending it.

With thanks to Project Gutenberg:
`http://www.gutenberg.org/files/1342/1342-h/1342-h.htm`

Related video presentations

At the AGI conference Adam Ford interviewed Margaret Boden and me separately on topics related to the conference.

In January 2013 he made both interviews available online on Youtube, with some editing, including some additional video effects – not seriously intrusive.

Links to the videos

- Margaret Boden interview: `http://www.youtube.com/watch?v=5dEXIOiAsaw`
- Aaron Sloman interview: `http://www.youtube.com/watch?v=iuH8dC7Snno`

I was very pleased to discover that Adam had respected my request not to add background music and noises which so many radio and TV producers insist on adding to spoken presentations, and which ruin them for the large and growing numbers of people with age-related hearing disability (presbycusis): for such people adding background music can affect spoken words in the same way as turning a telescope out of focus can affect viewed text: the words can be seen but not made out.

There are many more interviews by Adam Ford here
`http://www.youtube.com/user/TheRationalFuture`

2 High level overview

What this is about

- 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, even 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 cannot interface to their environment without using something very different (e.g. A/D and D/A converters) whereas chemical computers can directly interface with a physical/chemical environment.
- In these slides (part of the Meta-Morphogenesis project) I try to specify 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`

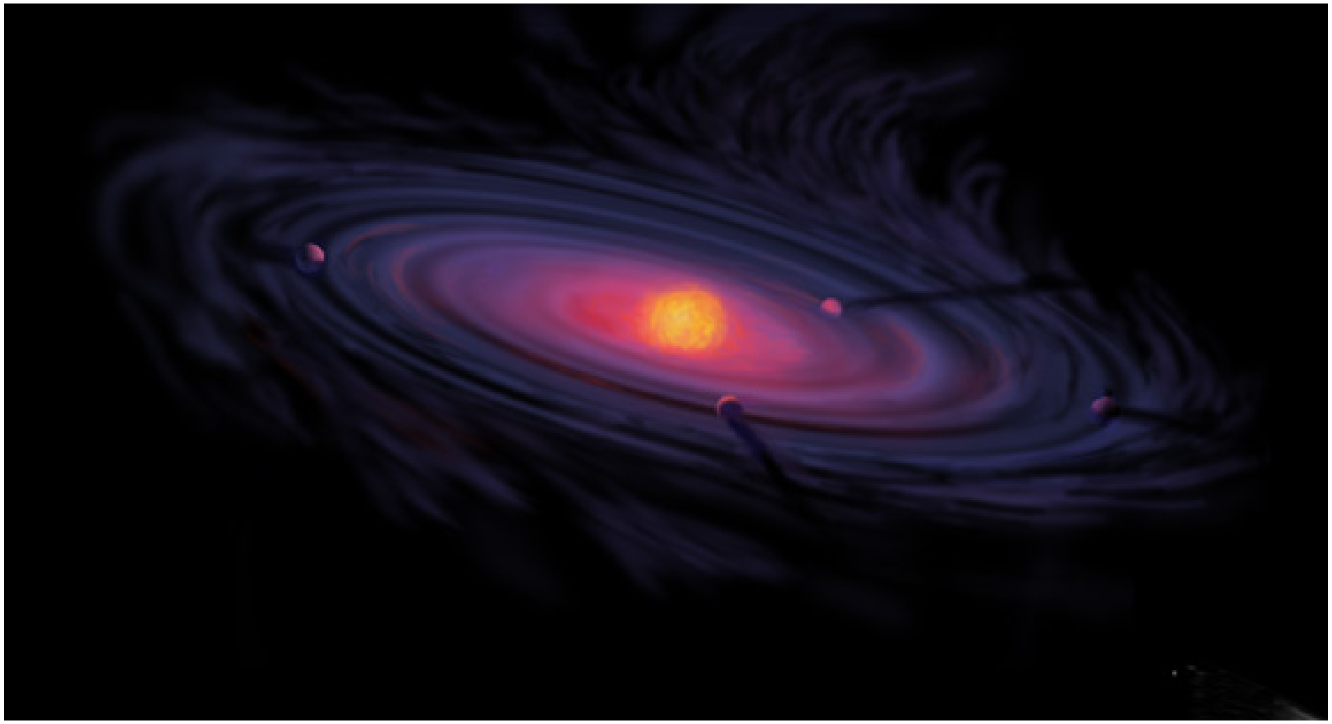
The main themes: all concerned with information

- All living things use information: of various kinds, from different sources, represented/encoded in different ways, and used for different functions or purposes. (Like Jane Austen's characters.)
- Acquiring, manipulating and using information requires information-processing mechanisms.
- Evolution transfers information across generations, using mechanisms of generation and selection.
- Learning and development use and extend available information (of various types) in individuals.
- The notion of "Information" used here is not Shannon's concept, but the much older idea of information content that refers to something that may or may not exist (e.g. my youngest daughter). (Sloman, 2011a)
- The mechanisms of evolution, development and learning change as a result of evolution, development and learning: hence the use of "Meta-" in "Meta-Morphogenesis".
- The changes and meta-changes are driven by a host of different factors, including changing physical environments, changes in other species, and new opportunities arising from previous changes.
- Not all organisms use information in the same way: the sources differ, the contents differ, the forms of analysis, synthesis, interpretation, storage, derivation, retrieval, and uses of information differ.
- Most of those aspects of biological information processing, and their changes, are invisible and cannot be reliably discovered either by observation of behaviour, or by study of brains and other mechanisms: like all deep science (and good philosophy) it needs a highly creative process of, observation, guesswork and hypothesis testing, using our steadily increasing abilities to design and build working systems.
- Very many organisms, including the earliest organisms and proto-organisms, microbes and plants, do all their information-processing without having any brains, using chemical and other mechanisms instead.
- Even organisms that do have brains make essential use of chemical information processing, perhaps in undiscovered ways, with some possible implications I'll discuss later.
- This is a huge project, likely to take many decades, and requiring collaboration between scientists, philosophers, educators and engineers, especially robot designers.
- I have tried to be bold and challenging – and have probably made some serious mistakes.
  Critics, helpers, collaborators welcome!

How did it all start?

3 Origins and inspiration

Wayback, when dusty



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

As the dust and gas condenses, planets, asteroids, and larger dust particles, etc. form.

Large numbers of tiny disconnected particles condense, through gravitational attraction and chance impact, into larger cohesive structures – some planet-sized.

The processes are disrupted by many collisions – e.g. formation of asteroids and moons.

It seems that the laws of physics and chemistry allow undirected processes to generate increasingly large and varied chemical structures: then new things can emerge, ...

Under what conditions?

(Which initial physical configurations suffice? Which possible worlds?)

Inspiration from Alan Turing: Morphogenesis

In his 1952 paper on morphogenesis (Turing, 1952), Turing tried to explain how sub-microscopic molecular patterns could produce visible patterns such as stripes and spots on a developed organism.

We can generalise his question by asking: how can current life forms, and their activities, including human mental processes be produced by initially lifeless matter?

Many have tried to address the origins of life by seeking evidence for changes in **physical structures** and **physical/chemical behaviours** produced by natural selection acting on products of a vast amount of random chemical "chatter".

There's a deeper question, which Turing might have asked, had he lived on:

What changes in information-processing (I-P) mechanisms would have been required, and how did those changes occur?

I suspect he would not have claimed that the causal processes involved could be replicated on a single physically implemented Turing machine – at least for reasons of tractability, and possibly for deeper reasons. I'll return to that below.

[Caveat: According to (Numerico, 2010), Turing's unpublished papers suggest that he was more interested in how brains could be grown – a physical process.

I recently learnt about Turing's unpublished work on morphogenesis summarised in (Swinton, 2004). It did not explicitly address information processing. So perhaps I've got Turing wrong. Perhaps not!??**]**

I'll try to show that the biologically important notion of information-processing (which I'll abbreviate as "I-P") using the pre-Shannon concept of semantic information, whose key feature is referring to something, is central to understanding what life is and how life forms with different capabilities evolve, including different information-processing capabilities.

(Unfortunately "semantic information" is now often taken by engineers to refer to formalised information.)

Inspiration from Turing: Meta-computation

In 1936, Turing constructed a problem that could not be solved by a Turing machine, namely the problem of specifying, for any program runnable on a TM, whether that program terminates or not.

(Other researchers, e.g. Gödel, had done similar things, though not in relation to Turing machines.)

The problem formulation presupposes that it is possible in principle for a program to inspect a program and discover something about it, where the inspection could be either

- dynamic, i.e. discovering some feature of the program's behaviour while it runs (e.g. to find out how much memory it uses when running, or how many program steps it requires to reach some decision, for given inputs), or
- static, i.e. examining the instructions, without running them, in order to prove that they will or will not terminate when run with given inputs – as some programs should do, others not (e.g. operating systems).

The original reasons for investigating such possibilities were concerned with theoretical problems about the nature and limits of computation or mathematical or logical reasoning.

More recently, practical problems that arise in computing systems have led to development of many varieties of meta-computation: programs that investigate properties of other programs or of themselves, either by statically analysing their properties (e.g. finding that there are no loops, and no recursive procedure calls in the program, and concluding that the program must terminate), or dynamically observing the program, e.g. to check on the resources it uses, or to ensure that it doesn't violate any access restrictions.

All this suggests that evolution might also have chanced upon useful forms of information-processing that allow organisms to inspect (statically or dynamically) some of their information-processing activities, e.g. thinking about what went wrong in a planning process if executing the plan failed to produce the expected result. (Sussman, 1975)

"Meta-cognition" can cover a variety of biological forms of meta-computation – how wide that variety is requires substantial research. I'll present examples later.

Inspiration from Turing: Meta-Morphogenesis

New biological mechanisms (including chemical mechanisms) able to produce new forms of **I-P** have evolved – new forms of reproduction, learning, development, cultural change, and “unnatural” selection mechanisms such as mate-selection and animal breeding.

So mechanisms of morphogenesis can produce new mechanisms of morphogenesis, thereby altering ways in which environments influence development of new **I-P** systems.

So mechanisms of morphogenesis can also lead to new mechanisms – producing “Meta-Morphogenesis”.

All newer forms of I-P in organisms rest on the earlier, molecular forms, though some (e.g. logic) are also re-implementable on other machinery, e.g. UTMs or modern computers.

A subset of forms of human arithmetical and logical information-processing have been replicated on modern computers, inspired by products of biological evolution.

But not all have, so far: e.g. `http://tinyurl.com/CogMisc/toddler-theorems.html#primes`

Mathematical and logical reasoning (both products of evolution), influenced human engineering design capabilities (also products of evolution), and thereby produced new (biological) reasoning mechanisms.

Previously evolved competences, and environmental influences, can combine to produce new forms of information-processing, including new mechanisms of evolution.

These ideas are not new, but are mostly ignored, e.g. by AI researchers seeking a particular type of I-P rather than aiming to investigate and compare different types.

The meta-morphogenesis project aims to identify significant changes in biological I-P, and the processes and mechanisms that drove them, including environmental factors.
One of the payoffs should be better understanding of some of the more recent, more powerful, products of evolution and its environments, including evolved forms of learning.

This could challenge attempts to base AGI on a uniform knowledge-free learning strategy.

Inspiration from Turing: Meta(virtual)-Machinery

Turing showed that there is a type of Turing machine, a Universal Turing Machine (UTM) whose symbols and transition rules are so powerful that any other Turing Machine, universal or not, can be modelled on a UTM. (1936)

So, any form of computation that can be run on any Turing machine can also be run on a UTM, though often with a large cost in speed and length of tape required.

- This was an early intimation of the possibility of producing virtual machines that run on other machines, including virtual machines running on other virtual machines: this is not just a theoretical possibility but an essential pre-requisite for the development of a huge variety of very different information-processing systems all using the same types of physical computer.
- However, each Turing machine, universal or not, has to be set up with an initial (real or virtual) tape configuration and then started off, running until its task is complete, or possibly running forever if the program contains a bug or is intended to generate an unending set of results, e.g. all the prime numbers.
- Biological I-P systems, using chemical or chemical and neural information processing, don't just run in a pre-programmed way: what they do depends on their (possibly changing) environments, and their actions don't merely change their internal memory (like the tape on a Turing machine), they also use information to control what they do externally, and the environment can, in turn, produce new information that alters the biological machine. (This blurs the internal/external distinction (Sloman, 1978, Chapter 6))
- The molecular information-processing mechanisms that allow the earliest forms of pre-biotic structures and organisms to grow, survive and replicate, later provide a basis for evolution and development of additional mechanisms that support I-P of new kinds, including systems that modify their own powers under the influence of the environment, e.g. a physical source of information about food, or a teacher.
- I am not saying that molecular computations exceed, or even share, Turing universality, though they may have space, time, or energy benefits in restricted, biologically important, domains.

  E.g. some of them may be able to solve problems that are solvable in principle on a TM, but intractable.

  I'll present illustrative examples later.

4 Languages of many kinds – used externally and internally

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 bad because it

(a) ignores the important non-communicative 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 (Chomsky, 1965), compositionality, structural variability, ability to refer to nearby and remote things and events, ability to express novel meanings, etc.)

(c) ignores the rich and varied types of information-processing done internally by animals 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, 2008b) (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.)

IT limits thinking about information

A short note of clarification on the previous slide.

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.

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.)

Acquiring information can expand what you are informed about (but not if it’s “old hat”).

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 (though 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 also talking about processing of meaning, and uses of meaning.

> 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.

An earlier slide provides conclusive evidence 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.

5 Environments can be referred to

What’s an environment?

- Things I have been reading while writing these slides, including the work of Turing, Kauffman, various discussions of information and biology, and disputes about limits of natural selection, almost all give the environment in which an organism lives, acts, grows, reproduces, gets eaten, etc. a marginal role.
- The main emphasis seems to be on what can happen **inside** an organism, provided that it has whatever is needed from outside to keep the processes going (e.g. food and sources of energy).
- In some ways this mirrors the main concern in much of theoretical computer science: namely trying to understand what sorts of things can happen inside a computer, or perhaps a network of linked computers, when presented with some problem to solve, or some state to achieve or maintain.
- Turing did allow that something outside the computer could play a role, namely his oracle: but as I understand it the oracle did nothing more than answer questions – e.g. it contained no rivals, or prey!
- My interest in information-processing (computation) from the start has always been concerned with how an organism or machine can cope with a complex, richly structured, multi-faceted, multi-scale spatio-temporal environment with which it has ongoing interactions. E.g. (Sloman, 1971), (Sloman, 1978, Ch 6)
- So, which problems arise, and whether a problem is solved or not, and whether an old solution remains a solution, and whether some new problem has arisen that should be given highest priority, can depend on things going on in the environment – which can contain other information processors.
  (Biology = chemistry + ethology + information-processing???)
- From this viewpoint, the work of Turing, more recent computer scientists, and people who try to represent biological processes as genetic algorithms or genetic programming with a fixed evaluation function, or who study models of computation in closed systems (even chemical computations) all omit a major feature of biological information processing: rich interactions with an extended, changing universe.
- Much biological information processing is about metabolism, growth, repair, immune responses, reproduction, etc. but not all is: the information processed, and therefore the internal languages used, can refer beyond the skin and forwards and backwards in time, and even to unobservable entities in the environment.

How to prove me wrong

It's clear what needs to be done to prove that I am wrong about the requirement for internal languages in many non-human animals and in pre-verbal humans.

Namely, demonstrate working robots 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, but without the mechanisms discussed here.

If working models are too difficult, then at least produce detailed specifications for designs that in principle future engineers could implement, with arguments showing how those designs will 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 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 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`

6 Aspects of natural intelligence

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) (Later I’ll criticise “model”.)

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` (J. J. Gibson, 1979; Sloman, 2011b).

7 Progressions in information processing

Sub-cellular chemistry

All biological information processing emerged from, and still makes use of, chemical processing – for instance chemical information-processing controls construction of brains; and brains, once built, rely crucially on chemistry, though computer models of neural networks mostly ignore this.

Our understanding is still at an early stage, and could turn out to be seriously incomplete.

Online (growing) collections of videos and lectures about sub-cellular chemical processes, vividly illustrate the remarkable functions of complex molecules in living organisms, such as reproduction, transport, storage and use of energy, anti-viral defences, etc. (E.g. `http://secret-universe.co.uk/`)

The role of **information** in these mechanisms, e.g. information used to control, rather than information that passively refers, is not always brought out clearly in such videos.

In part that is because it is much easier to depict motions produced by physical forces than to depict changes in abstract, dispositional states, with causal powers that are usually dormant.

E.g. changes in dispositions can occur when a state moves closer to or further from a threshold: the changes may be very subtle and difficult to detect, even though reaching the threshold produces major, more easily detectable, changes.

We should not expect to see good visualizations of thinking processes in complex virtual machinery running on physical brains – e.g. because thinking is very different from image manipulation. (See later.)

Later I'll try to collect examples of useful online visualisations and add pointers here.
Suggestions welcome!

We also need to understand how states and processes in the environment relate to internal information processing – as sources of the information, as determinants of and constraints on what the information could be used for, and as referents for some of the information, whether represented chemically or neurally, or in some other way.

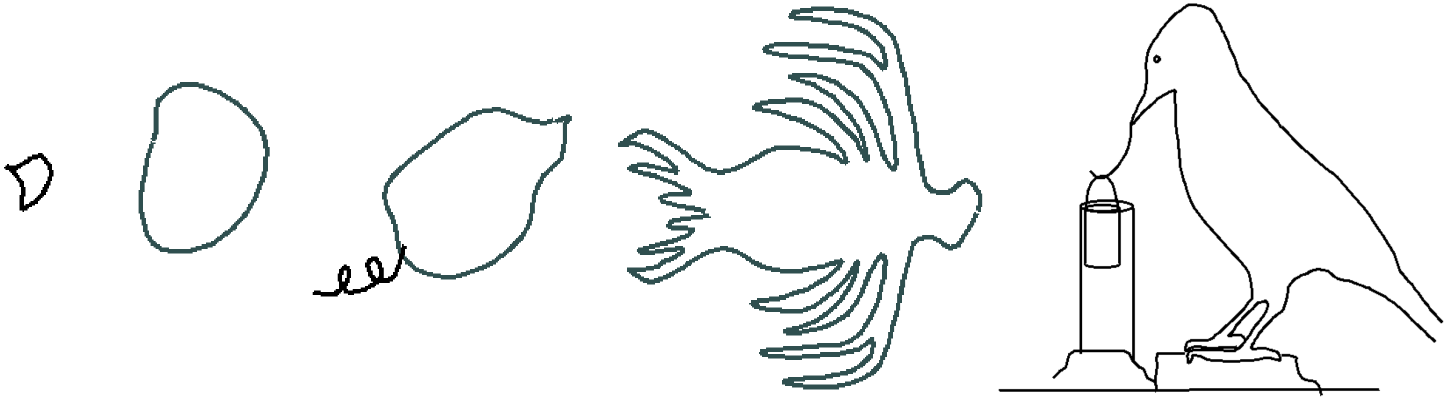
All organisms are information-processors but the information to be processed, the uses of the information, and the means of processing vary enormously

The key feature of (pre-Shannon) information is that **information has uses.** Moreover, some uses of information are internal and very hard to detect.

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

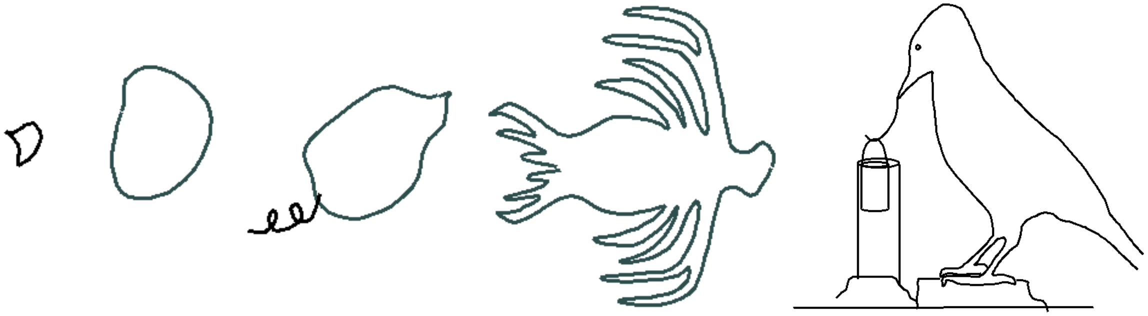
Some involve other agents: collaborating, competing, helping, hindering...

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 disclaimer about soups in later slide.]
- 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, ...
- 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 before reading on.)**

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

8 Alternatives to soups for starters

What if life did not start in a soup?

Chemical soups are not the only possible initial environment for life forms.

Alternative hypotheses are physically/chemically possible.

E.g. Nasif Nahle (Nahle, 2004) claims earliest life forms could not have developed in a soup, because of problems of controlling osmotic pressure, though they could have started in dust grains covered in ice, giving shelter against solar and cosmic radiation.

> *"The dust grains (fractals) acted like the biomolecules protective 'eggshells' against the ionizing solar radiation. Thus, the chemical changes allowed the synthesis of more complex carbohydrate, proteins and lipids molecules, which reached to the structure of quasi-stable and highly-lasting membranes in the shape of microspheres. Nevertheless, those membranes persisted on being ephemeral by the intensity of the cosmic radiation that could destroy them. Many microspheres that were enclosed by membranes subsisted in that hostile atmosphere thanks to that they were into solid dust grains covered by water ice."*
> `http://www.biocab.org/Abiogenesis_Synopsis.html`
>
> The final section "Summary of the origin of living beings in the universe" of the longer, more detailed, paper, proposes that at a later stage water vapor condenses "forming heavy drops that precipitate on the planetary soils dragging the grains of dust with and without microspheres with them", allowing a subset to persist, then later amalgamate through electrochemical affinity, fusing and forming vesicles with continuous membranes, that rest on the humid soils or in the bottom of shallow or subterranean ponds, where holes of soils, full of chemical substances, are covered by the microenvironments chemically similar to the fluids within modern cells. `http://www.biocab.org/Abiogenesis.html`
>
> I am not able to evaluate this theory, but for now it suffices to illustrate some possible stages in the transition from clouds of dust to early life on this planet.

This sort of theory has implications for the earliest requirements for information processing in a non-soup environment.

Compare the above with the hypothesis that early 'biological' molecules were formed by metal-based catalysis on the crystalline surfaces of minerals. "In principle, an elaborate system of molecular synthesis and breakdown (metabolism) could have existed on these surfaces long before the first cells arose."
Alberts B, Johnson A, Lewis J, et al. New York: Garland Science; 2002. `http://www.ncbi.nlm.nih.gov/books/NBK26876/`
From: The RNA World and the Origins of Life, in *Molecular Biology of the Cell* 4th edition.

The influences of different 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)

Two types of evolutionary trajectory

Consider a mobile (swimming) organism in a chemical soup that contains some fixed structures, e.g. rock surfaces in different places with different nutrients and different noxious materials.

Suppose the organism can acquire information about the spatial locations of the different substances, and use that information to propel itself (by some sort of swimming motion, or use of jets) to the location that meets its current most urgent need, while avoiding proximity to the dangerous surfaces.

Now consider two evolutionary trajectories leading up to that collection of competences:

1. A trajectory starting with floating blobs of chemicals gradually evolving the sensory motor morphologies and information-processing capabilities required to learn about and flourish in marine environments with large scale relatively static structures.
2. A trajectory starting with blobs of chemicals attached to solid objects and gradually developing the ability to control movement on larger and larger surfaces with different properties, and then later to swim between such objects as well.

Are both trajectories feasible as evolutionary routes from passive blobs to active marine explorer/foragers?

I suspect not, but will not argue the case here: for now, I merely wish to illustrate a type of question generated by the Meta-Morphogenesis project.

9 Use of information in control

Biological control by chemical information

Over centuries, human designers have explored increasingly complex ways of using information to control machinery – but evolution's explorations have gone on much longer in a much wider variety of "laboratories", with a much wider variety of materials, and a vastly richer collection of "discoveries"

(probably many more than human scientists have become aware of).

- Information can have many control roles – the simplest being to switch on or off some deployment of energy: discrete control.
- More complex cases involve continuous control - of speed, direction, rate of flow, etc.
- Structural control: control of assembly, disassembly, of physical/chemical structures
- Many forms of Meta-control: modulating control function of something else – many varieties are being discovered in studies of genetics, epigenetics, brain functions, etc.
- Some of that variety comes from the variety of underlying physical mechanisms used to implement the control functions – e.g. chemical control functions become more varied as molecules get larger.
- The basic forms of chemical control require spatial proximity, but can be very rich: multiple relationships between parts of complex molecules change in coordination, in parallel, partly as a result of changing spatial relationships and partly as a result of creation, removal or modification of chemical bonds.

  Examples: Catalytic control; or use of DNA to control assembly of new molecules.
- Chemical control can also operate at a distance if molecules or ions assembled in one location are transported to other locations where they have their effects – as is the case with hormonal control. Neurotransmitters are similar but control the operation of neurons, often over short distances.
- Pheromones can cross boundaries of whole organisms, and act across varying spatial and temporal gaps.

Control over changing time and space gaps

Before brains evolved, and even now in the simplest organisms, most control has to be chemical - molecules determining what should happen to other molecules.

There must be interesting stories to be told about how various forms of chemical control on very small time-scales and spatial-scales were superseded by, or enhanced by, other forms of control including use of chemical messages, electrical and electromagnetic (e.g. light-using) mechanisms, and extending over larger spatial regions and time intervals.

Part of that story will need to be clarification of the transition from physical/chemical causation to control.

A key feature of the transition is separation between what provides the **energy** for some process, and what provides the **information** that is used to select among alternative processes using the same energy source, or possibly used to continuously modulate the use of energy, e.g. steering something.

For example, car drivers originally caused changes of direction by physically changing orientation of wheels; but power steering technology now allows drivers use steering wheels to provide information about required changes of direction, while the engine provides the energy.

Are there similar transitions in evolution of biological control mechanisms?
(To be expanded)

10 Changing control functions of information

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 whatever **embodies** or **encodes** the information then the suggestion that information "may not involve representations" becomes self-contradictory.
However, some representations may be very abstract process patterns. (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 states of entities in the environment);
- 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.)

Aspects of information use

A deep theory needs to disentangle different aspects of use of information.

An active item of information **I** may have a source, **S**, which could be

- an intentional agent e.g. when information is in a communication, or
- merely a physical cause, e.g. when the information is about visible features of this tree, e.g. its width, curvature, texture, height, etc. – Here the information about the tree's appearance is provided by the tree, without the tree having any intention in the matter: information is made available, without being communicated. (I call this "self-documentation".)

Information **I** acquired by a user **U** has a carrier, or encoding, **C**, that's part of **U** or something **U** can access.

Typically **I** will be complex and C will need complexity, i.e. parts and relationships, to express **I**.

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.

In some cases all differences will be discrete (e.g. where C1 and C2 are assembled from a fixed set of parts), in others continuous, e.g. as a changing signal C indicates changing size, or distance, or, curvature, etc. in I.

Changes in carrier C will correspond to changes in information I in a systematic way, but there are very many different types of systematicity.

E.g. components of C may be arbitrary tokens whose juxtaposition expresses the content I in a systematic way, like words in a sentence, or parts in a picture, but non arbitrary tokens (e.g. of varying size) can also be used.

Systematicity in the relationship between C and I does not imply that C has to be isomorphic with I: counter examples include both Fregean representations (f(a, b, c) can denote something whose structure is totally unlike that encoding), and 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.

Similar points apply to inactive information, stored for possible future uses.

11 Many uses of information

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 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 organisms.

NB: Increasing gaps between acquisition and use

I shall repeatedly illustrate ways in which both evolution and individual development can produce bigger and bigger gaps between

- the **acquisition** of information items
- the **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
- 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., is 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...

12 Varieties of motive generation and motive processing

Transitions in motive generation and processing

The initiation and modulation of processes, whether internal or external, individual or shared, whether collaborative or cooperative, involves forms of motivation with many functions, e.g. proposing goals, comparing goals, selecting goals, abandoning goals, initiating action, modulating action (changing speed, direction, or means, or tool or order of steps, etc...).

Later I'll return to the topic of transitions in motive processing across evolutionary and developmental changes – starting with simple and direct reflexes or reflex-like responses, and adding more and more intervening complexity (and sometimes delay) between whatever triggers or influences actions and performance (or suppression) of those actions.

There are also important social and cultural influences in some cases.

All this is obvious to many researchers who do not necessarily see these transitions as part of a global process of enhancing/enriching forms of biological information processing.

First I'll turn to another class of transitions: changes in information transferred across generations, and changes in how it is transferred – genetic information, in the broadest sense.

Evolution enriches starting states for new individuals

[Place-holder: to be expanded]

Many assume that a new-born learner has to deal with an undifferentiated, or at least unstructured, stream (or "firehose") of sensory signals, working out how to impose structure by seeking patterns in the "blooming, buzzing, confusion":

Hypothesised by William James – discussed by John Hawks:

`http://johnhawks.net/weblog/topics/minds/baby/james-blooming-buzzing-baby-2010.html`

(He seems to be unaware of Immanuel Kant's argument (1781) that the learner cannot learn from totally unstructured experiences. Compare recent work by Ben Kuipers and Joseph Modayil [REF])

The existence of many innate competences, in many species, demonstrates that on this planet evolution can provide its products with some powerful innate assumptions, either about very specific environments in the case of the majority of organisms, or about general ways in which environments (on this planet) can vary and can be distinguished, so that learning about them can be speeded up (e.g. using ideas about space, time and causation to help drive the process) Compare (McCarthy, 2008), and Kant's ideas about apriori knowledge (1781).

Biologists report (a) many examples of precocial species that start with very specific assumptions about the environment, e.g. new chicks, and also (b) many examples of altricial species that learn more slowly but more deeply. At least one species has members that are scientists, engineers, musicians, etc.

An alternative to the nature-nurture dichotomy allows interleaved data-driven learning and partly nature-driven reorganisation to bootstrap ever more complex types of learning in a developmental process of meta-morphogenesis – sketched later. (Chappell & Sloman, 2007)

Ideas about "representational redescription" in (Karmiloff-Smith, 1992) and McCarthy's discussion of a "well-designed" child, help banish simple binary oppositions (McCarthy, 2008).

Special features of control of physical processes

In the earliest organisms where most control concerned chemical interactions, options may have been mostly discrete – but things changed.

As organisms became larger,

- with articulated parts interacting with other physical objects translating, rotating, bending, squeezing, stretching, twisting, levering, etc.,

increasingly many physical processes generated were continuous

- including multi-strand processes in which different properties and relationships change, or start existing or cease existing, in parallel (Sloman, 2008a).

Many AI systems operate on information and decisions that ignore physical continuity, assuming a discretised world: "go to door, open door, go through door, turn left, go to desk, ... etc."

But robot designers have to ensure that continuous movements are generated and appropriately controlled, within certain task specific tolerances.

This can "hide" the continuous processes from the robot's cognitive mechanisms.

However there are many tasks where continuous control is part of being intelligent, even if some transitions are discrete.

Examples include

- using visual servoing to control picking up an object by its edge, using finger and thumb,
- moving an armchair through a doorway which is too narrow for it simply to be pushed through, by rotating the chair in 3-D

among many others discussed on my web pages E.g. (Sloman, Chappell, & CoSy PlayMate team, 2006)

Control of complex, changing, physical processes

Rigid (unconditional) and flexible (conditional “gappy”) control.

Examples of rigid control based on information “encoded” in physical shapes: (a) transfer of water from a source to a destination by means of an unchanging channel or water pipe, (b) a helter skelter (tornado slide) plus gravity producing spiral downward motion of a marble released at the top.
Contrast with (c) A person walking along a road deciding which way to go at junctions when they are encountered, or (d) someone building a house and postponing choice of some materials.

Examples of types of flexible control – a system with:

- a range of possible actions available at all, or at certain, times or locations, where actions available for selection may be performed immediately and locally or at some future time and/or distant location.
- mechanisms for switching (continuously or intermittently) between actions, or speeds, manner, etc.
- enabling/disabling mechanisms for acquiring information (e.g. looking, listening, feeling,...) used to select or modify actions (usually while the source of energy for the actions is unchanged)
- contrast a simple conditional schedule that specifies which actions should be selected under which circumstances, with more complex cases where the schedule is itself controllable and modifiable by “higher level” mechanisms, e.g. a planner, or a system of values and goals that can change.
- use of changing sources of information to use as a basis for selecting between alternatives (e.g. elders).

Physical structures provide very many flexible controls, using matter to manipulate matter: including controlled sub-microscopic matter-manipulation in organisms.

Familiar examples include use of physical surfaces to guide grasping (sensing and responding to sensory information are tightly integrated), and using visual cues to follow a path (loose, indirect, connection).

When you use a hand to scoop sand or mud to add to a sand-castle, or pull a sweater over your head, or bend a branch to break it – there are continually changing forms of interaction and control: different subsets of sensory and motor neurones are mapped onto constantly changing external structures in a constantly changing way. How?

13 Non-human organisms should also be studied, when attempting to understand human minds.

Not only humans ...

Many other animals perform complex actions to achieve complex results with huge variations in the details of the perceptual information and control information in different situations. Examples:

- an elephant using its trunk to get leaves and other food into its mouth,
- an orangutan high in a tree, holding her baby while making a safe bed of leaves and branches with her free hand, for the night (making a different bed every night),
- an octopus squeezing itself through apertures of different shapes and sizes,
- a nest-building bird weaving a new twig into place in a partly built nest
- an ant pushing a larger insect carcass towards its nest, and on meeting an obstacle going round to pull the carcass over the obstacle. [REF?]

Current work on robotics that I have seen isn't even moving in the right direction to replicate such competences: the forms of representation, the ontologies, the learning mechanisms, and the forms of control mostly seem to be wrong – in part because designers feel compelled to make use of particular sorts of mathematics, e.g. with state spaces composed of, or derived from, vectors of scalar values in sensory and motor signals, instead of more abstract features of the environment, and attempting to use probabilities to deal with uncertainties rather than abstractions and partial orderings.
Perhaps we can find out what needs to be done by looking very closely at many evolutionary and developmental precursors?

(E.g. many species share the ability to drink water, and mammals share the ability to suck milk at birth Perhaps abilities to consume liquids, in a common ancestor, were precursors to many manipulative capabilities that evolved later for manipulating non-liquids – requiring more precision?).

Does the chemical basis of biological information-processing play a crucial role in controlling actions in all these cases, or did evolved virtual machinery make it irrelevant?

14 Understanding possibilities and necessities is initially more important than understanding probabilities.

Why I don’t want to use Bayes and probabilities

Possibilities and impossibilities are more fundamental than probabilities.

There are many more sets of possibilities than we can assign probabilities to
– Including structures and processes involving multiple relationships
– changing in parallel in most processes

There are deep forms of learning and development that involves extending what’s possible

In some cases that leads to new discoveries of impossibilities.

We can do a great deal without having probabilities.

In some cases that uses a partial ordering of likelihoods, but very very partial.

Compare Piaget’s last two books: on Possibility and Necessity.

But I don’t like the details of his theories.
Karmiloff-Smith(1992) is better.

Show some videos from here
`http://www.cs.bham.ac.uk/research/projects/cogaff/movies/vid`

15 The need to capture the diversity of types of example, and the forms of formation-processing they require, including online and offline processing – leading to mathematical know-how.

Need for examples – lots of examples

The changing uses of information were listed above in a very general, abstract fashion.

However, the details are what matter.

Addressing that means finding and analyzing examples – many, many examples.

Too often, researchers take one example, or a tiny subset of examples (e.g. board games or tile-worlds), as definitively illustrating all possibilities, and as a result they fail to understand the diversity of problems and never notice the inadequacy of their proposed solutions and competing solutions.

> One of the worst cases is trying to understand minds, learning, consciousness, motivations, emotions, evolution, etc., by focusing only on humans.
>
> We can compare that with trying to develop chemistry by studying one complex molecule, e.g. haemoglobin.
>
> This has been true of philosophers for a long time.
>
> Narrow ranges of types of example are also often used unwittingly by factions in psychology, AI/cognitive science to defend their theories from attack.

Towards a remedy:

A small start, on a huge project, to collect and organise examples, is in:
`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/evolution-info-transitions.html`

The ideas presented in these slides help to reorganise those examples, which in turn reveal the need for changes in these slides.

Additional types of transition in biological information-processing follow – starting with the transition from online to offline intelligence.

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`)

The ability to control real time interaction with an environment could be called **“on-line intelligence”**. (E.g. (Adolph, 2005) shows how on-line intelligence can develop in young children.)

There’s a huge amount of recent research on how to get robots to walk, run, dance, pick things up, go smoothly through doorways, move so as to avoid obstacles, keep their balance, etc.
An impressive example is BigDog: `http://www.bostondynamics.com/robot_bigdog.html`

Those are important, non-trivial research goals, possibly relevant to explaining important subsets of the behaviour of many animals, including insects, fishes, birds, grazing mammals, etc. **But not the only goals.**

Other kinds of biological I-P: “offline information processing”.

This includes collecting information about large scale spatial structures for possible future use (e.g. SLAM “Simultaneous Localisation and Mapping” in robotics), but also acquiring information about what sorts of things are possible, what would happen **if** some of those possibilities were realised, and what the causal and mathematical constraints on possibilities are: this can lead to scientific and mathematical discoveries, story-creation, art, philosophy, ... **not all in the service of action.** (Sloman, 2006, 2010b)

Daniel Wolpert argues that 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.

What offline processing adds

(To be expanded ... clarified ...)

Extending Gibson's ideas about affordances:

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

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

Being able to do counter-factual reasoning

(e.g. about what could have happened and what the consequences would have been, about what you did not do and why not.)

This can lead to, or make use 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)

Online vs offline intelligence

This appears to be an example of 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?

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

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

Unfortunately the mathematical education many philosophers and scientists have had at school deprives them of the experience required to think about these issues.

For some simple examples of geometrical reasoning that seem to be beyond current AI, see
`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 compose stories, jokes, poems, tunes, etc. with eyes and mouth shut.

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.

16 How the nature of space/time controls and enables ever more complex biological information processing.

Problems relating to continuous spatial control

It is often thought that learning can be explained by imitation.

But that ignores some of the complex control problems in imitating continuous actions requiring precise control of change.

If a dancer, a violinist, a tennis player, or a parent tying shoe laces says to a learner: "this is what you have to do", then demonstrates the actions, it is not usually possible for the learner immediately to replicate what was demonstrated, even after several repetitions.

That's not because some subtle feature of the action was hidden (as with stage magicians).

It's partly because the learner needs to develop the appropriate ontology for fragments and features of continuous processes in order to be able to generate those processes.

Some researchers assume that's an ontology of patterns in sensory and motor signals (as in "multi-modal" learning), but I think that's totally wrong in most cases: the control requires an ontology of 3-D spatial relations and changes in relations of body parts of the performer.

E.g. a violinist learns to change a shoulder angle to make the bow rotate in a vertical plane so that it moves from one string to another and change the elbow angle to make the bow move on the current string to generate a tone, with more subtle variations of pressure and speed affecting the quality, volume, and discrete or continuous sound produced. (Changing nerve signals are means, not an end.)

A robot that learns to rotate a crank handle smoothly by generating a pattern of sensorymotor signals will fail if made to stand in a new position relative to the handle.

But if it has learnt to identify the vertical plane in which the crank rotates, and the changing tangent to the circle through which the end of the crank arm moves, then it can constantly alter the direction of the force on the hand to align with that tangent (perhaps using visual-servoing initially).

How could a 3-D ontology for motor control evolve? How is it represented in animals?

Example: Ways to put on a sweater

The example of putting on a sweater illustrates closely coupled internal and external processes with constantly changing relationships.

Think of different sequences of actions that can be used, such as

1. place sweater on a horizontal surface ensuring its front faces down, and the bottom edge is nearest you.
2. grasp bottom edge at two extremes and lift, leaning forward as you pull the bottom opening over your head, until your head goes through the sweater neck – with sweater compressed around your shoulders.
3. Use left or right hand to feel for the appropriate arm hole and push it through, then repeat with other hand.
4. pull waist down, and pull various portions of sweater in directions required to finalise the process.

Compare that with alternative sequences, e.g. first scrunching sweater so that your head can go directly through the neck hole, and then unravel as required to get hands and arms in through sleeves; or start by pushing one hand and arm through one sleeve, etc. (Has anyone studied the variety of strategies used by children and adults? What happens to a child who has never previously met a sweater?)

Different detailed motions are required with sweaters of different sizes, different thicknesses, different degrees of elasticity, different size neck holes, etc.

Even putting on the same sweater twice will not involve exactly the same physical processes, or exactly the same patterns of sensory and motor nerve signals, and muscle contractions.

Moreover you can put the same sweater on repeatedly while having different conversations at the time, or while standing, sitting, or with movement restricted by injury. (But some injuries make it impossible.)

The patterns of motion of portions of wool, skin, muscle, bone, if recorded in detailed measurements of location, orientation, speed of movement, speed and type of deformation, will vary enormously even if the same sweater is repeatedly put on in the same room – unlike the uniformity of patterns of action of a welding robot on a car production line. WHY? How much similar variation is already present in infant suckling?

What forms of representation and what sorts of information processing mechanism, in what kind of architecture, can support all that variability? (And how did they all evolve?)

Evolution of motivation (many transitions)

Another example of benefits of increasing gaps between sensing and acting is replacing reflex action triggers with goal triggers:

- 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.

For more on "Architecture-Based Motivation vs Reward-Based Motivation", see
`http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#907`

Digression? 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?)

The previous discussions and examples, should have made it clear that 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 using a compliant grasper to manipulate an object, some of the information in the shape of the object grasped is not transmitted but directly used as part of the sensing process: sensing by deformation of the grasper.

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 (or communication) do not fit some simple mathematical schema, like communicating by using strings of bits, or using sentences in a formalism defined by a grammar.
(this page needs revision)

Increasingly mediated control of physical processes

The earliest organisms had relatively “unmediated” interactions with their environment: they sensed, they reacted internally and they responded.

Increasingly, however, the gaps between sensing and selecting what to do, and the gaps between considering what to do and doing it became larger, and more diverse, e.g.:

- – temporal gaps
- – gaps in forms of representation
- – gaps in ontologies used.

Later I’ll present different architectures for dynamical systems, including some closely coupled with the environment, others with gaps/disconnections between information processing and the environment. (See the slides on dynamical systems, and on architectures.)

I suspect many important developments in evolution of information processing were concerned with increasing spatial, temporal and conceptual remoteness:

- – between acquisition of information and use of the information
- – between use of information and what the information referred to.

Including bigger and bigger gaps between

- – thought and perception
- – thought and action

Some gaps lead to development of mathematics (study of abstract entities), some to physics (study of very small, very large, near and far, past and future), some to history and geography, and some to philosophy (not an exhaustive list of types of gappy I-P).

As usual, more examples are needed. Some will be added to the M-M web pages.

Physical mechanisms for information processing

So far I have mostly discussed (in a very brief and sketchy fashion) varieties of functions, roles or purposes for information and information processing in organisms and varieties of types of information content – but we also need to take into account varieties of mechanisms for information-processing.

In the early 20^{th} century several types of abstract mechanism for I-P were investigated including propositional and predicate logic enhanced with axioms and rules of inference, the Lambda (λ) calculus (precursor of LISP), production systems (Precursor of rule-based programming languages), and Turing machines (with linear tape), and von Neumann machines (with randomly addressable memory).

Physical mechanisms were used in various kinds of control machinery, including sorting machines for punched cards, musical machines, programmable looms, calculators etc., then later the mechanisms were increasingly electrical, and transistor-based, with a sharp separation between processing units and storage units, though various kinds of analog, analog to digital and digital to analog devices continued being used.

Later there were various experiments using artificial neural nets, often ignoring most of the details of biological neural nets on the (rash) assumption that they are unimportant implementation details.

There have also been emulations of evolutionary processes and mechanisms, also ignoring most of the biological details, including the rich and varied interactions between individuals and complex environments.

It is at least possible, and now seems to me to be very likely, that ignoring some of the mechanisms made possible by the physical world in which evolution occurred may lead us to focus only on mechanisms of the sorts that emerged from research in mathematics, logic and computing –
inadequate for biological explanation, and possibly for advanced robots.

What sorts of underlying mechanisms are **required** to support, or **can** support, information-processing?

Chemical processes somehow grew more complex and varied, producing chemical forms of I-P and later other sorts, e.g. neural I-P, unintended use of stigmergy, intended social signalling, planning, theorising, philosophy, ...

- Organisms used chemical computation long before neural forms were available.
- Chemical mechanisms, unlike TMs, support both continuous changes in spatial and structural relations and also the ability to cross a phase boundary and snap into (or out of) discrete stable states that resist thermal buffeting and other potential disruptions.
- This stability relies on quantum mechanisms – but it is not clear whether other features of QM are so important (e.g. indeterminism, non-locality).
- In molecular structures, multiple constraints can be applied by complex wholes to parts, allowing some motions or rotations or chemical behaviours while restricting others.
- Some switches between discrete states, or between fixed and continuously variable states can be controlled at low cost in energy by fast-acting catalytic mechanisms.
- Even in organisms with brains, chemical information processing persists and plays a more fundamental role (e.g. building brains and supporting their functionality).
- As chemical structures expand the possible mechanisms and processes explode.

What's special about chemical computation/information-processing?

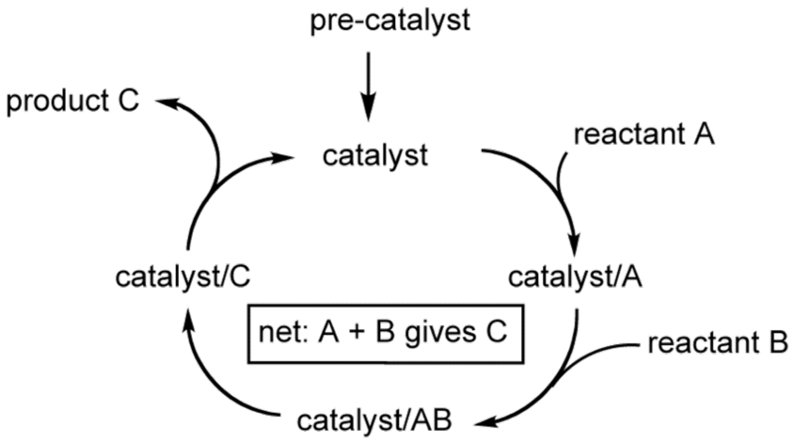
I have no definite answers now - perhaps others have? (Cf. Kauffman? (1995))

See Korthof's useful review of Kauffman: `http://home.planet.nl/~gkorthof/kortho32.htm`

Chemistry example: generation of structures

Catalytic mechanisms allow one chemical structure to assemble others.

The figure shows a catalyst repeatedly combining items of types A and B to form an item of type C.
After the product (e.g. C in the figure) is released, the catalyst is regenerated, and available for re-use.
From: http://en.wikipedia.org/wiki/Catalytic_cycle


pre-catalyst
catalyst
reactant A
catalyst/A
reactant B
catalyst/AB
catalyst/C
product C
net: A + B gives C

Catalytic cycle
for repeated conversion of A and B into C.
(From WikiMedia)

Example:

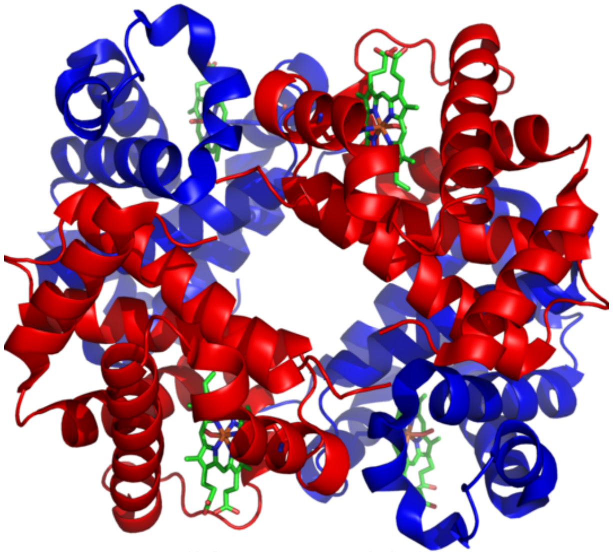
Haemoglobin (hemoglobin) molecules can absorb, transport and release oxygen, carbon dioxide and other substances, using catalytic reactions, as explained in this video:
https://www.khanacademy.org/science/biology/human-biology/v/hemoglobin

Collections of catalytic processes can form "autocatalytic networks" whose nodes continually produce catalysts or materials for other nodes.

Stuart Kauffman (and others) suggest that some of the earliest precursors of life might have used such networks, assembling and releasing molecules with similar capabilities.
(Reviewed here: http://home.wxs.nl/~gkorthof/kortho32.htm)

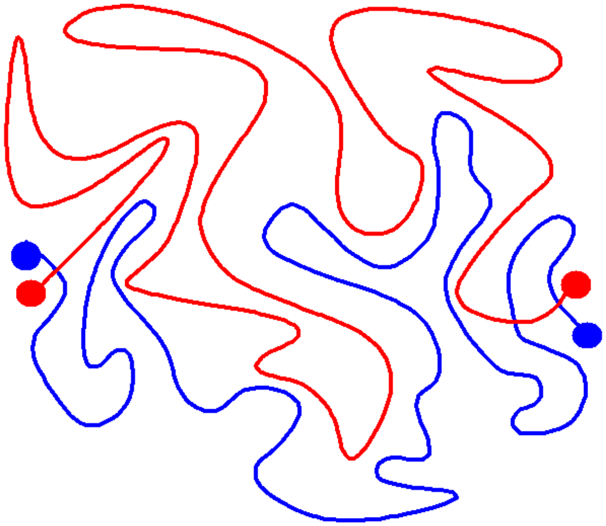
Perhaps billions of catalysts within an organism together function as a highly distributed highly parallel collection of (special purpose) processors: the original multi-core cpu??

Kauffman also suggests that the same principles of growth of complexity and control of processes by ever more complex interacting controllers apply to social and economic systems.


(Haemoglobin)
(Richard Wheeler, WikiMedia)

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.

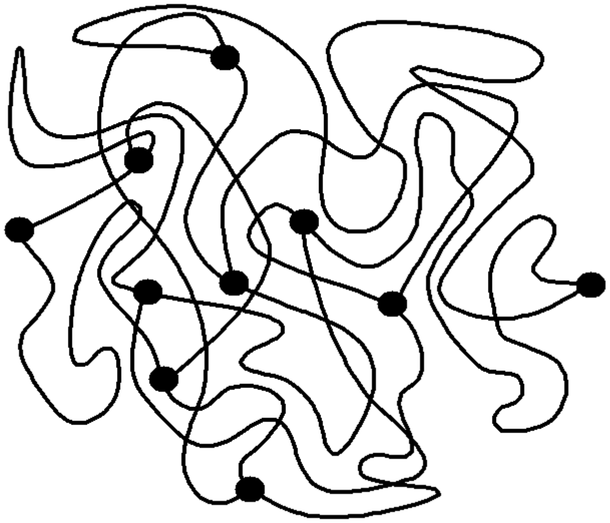
All of these mechanisms depend on propagation of causation through space over time, and the availability of mechanisms to detect, compare and use the resulting states.

(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 would be to measure all the lengths of the road segments linking towns then search for the shortest route, possibly 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.

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`

How is this relevant to AGI? (This was an AGI tutorial.)

It might seem that this concern with biological information processing systems is irrelevant to those working on Artificial General Intelligence.

There are several answers.

- Many AGI researchers seek inspiration from natural intelligence: so they will need to understand the varieties of natural intelligence and their supporting mechanisms.

  E.g. capabilities of biological systems are worth studying as part of the process of developing sets of requirements for new useful machines, since neither the requirements that drive evolution of animal intelligence nor the capabilities that result are obvious.

  E.g. Many AI vision researchers and language researchers have got both the functions and the mechanisms of vision and language badly wrong as argued in
  `http://tinyurl.com/CogTalks/#gibson`, and `http://tinyurl.com/CogTalks/#glang`.
  Also mechanisms of motivation and affect `http://tinyurl.com/CogTalks#cafe04`

- It is possible that aspects of the designs and mechanisms produced by biological evolution are pointers to unobvious solutions to the problems of designing artificial systems of similar (or greater) power and versatility.

  Are there biological precursors to current I-P mechanisms that are part of the indispensable infrastructure (or substructure), and a source of power in newer systems? Chemical I-P systems? Compare chemists studying very complex molecules while ignoring the simpler ones?

- Some of the earliest AI researchers (e.g. McCarthy, Simon, Newell, and before them Alan Turing) were interested in the science of natural intelligent systems, and wanted to use artificial systems (computer based) to help formulate and test theories about natural intelligence. (They used "artificial" mechanisms to study "natural" mechanisms.)

Summary:

Both AI as science, and engineering AGI, need a basis of good, deep, broad science.

Another justification for identifying transitions

An important reason for investigating transitions in biological information-processing, from microbes, and earlier, to humans (and beyond?) is that understanding such transitions is a requirement for bridging what Darwin's friend, T.H. Huxley, labelled "The explanatory gap".

- Bridging the gap requires explaining how minds and mental phenomena relate to physical (e.g. physiological) mechanisms, and how they could have evolved.
- The plethora of conflicting theories and problem statements regarding consciousness, emotions, etc., indicates deep muddle, requiring a major breakthrough. (Sloman, 2010c)
- We now have explanatory and modelling resources that Darwin could not have imagined, without which he could not possibly answer some of his fiercest critics.
  (However Ada Lovelace anticipated some of the answers. Did they ever meet?)
- Since Turing's time, we have designed and implemented increasingly complex virtual machinery, in which multiple concurrent non-physical mechanisms interact, performing functions that cannot be defined in the language of the physical sciences
  (e.g. correcting grammar and spelling, playing games, detecting threats and trying to avoid or minimise them, reasoning, forming motives, planning, selecting and executing plans, detecting unauthorised actions, learning grammars, formulating and testing theories, and many more). (Sloman, 2009b)
- Some of the important interactions include virtual machines that monitor and modify or modulate other virtual machines, or even themselves.
- Some virtual machines perform a single task that could be done by a physical machine (e.g. a virtual machine replacing a physical temperature controller), but others extend Turing's UTM idea by providing new platforms for supporting multiple additional machines (e.g. operating systems, and programming language interpreters).

Virtual machinery: Evolution got there before us?

Recently developed conceptual resources for specifying and reasoning about complex virtual machinery are not widely understood, even by people who have contributed to their development.

(Sloman, 2010a, 2010d, 2010e, 2010f)

See also `http://tinyurl.com/CogTalks/#talk95` and earlier talks.

- Trying to use introspection, evidence from psychological experiments, and research on brain mechanisms to infer the virtual machine functionality developed by evolution does not lead to good (deep, powerful, implementable) explanatory theories

  E.g. meeting the criteria for explaining possibilities formulated in chapter 2 of (Sloman, 1978).
- Research on evolutionary transitions where new kinds of mechanism were required, asking what benefits (in terms of new functionality) and what costs (e.g. in terms of physical resources, time for development, time for learning), can enhance our ideas about the space of possible designs (design space) and the space of possible sets of requirements (niche space).
- Because the evidence is so sparse, such research will have to include speculation inspired by observation and exploratory designs, justified if the speculations lead to design ideas that are testable both by empirical investigation and by working implementations.

  I.e. try to avoid theories expressed in vague verbiage and diagrams not usable by engineers.

  This may show that biological evolution encountered and solved hard design problems requiring use of self-monitoring, self-modifying virtual machinery that we have not yet identified, including machinery required to explain mathematical cognition. `http://tinyurl.com/CogMisc/triangle-theorem.html`

This should interest AGI researchers, though some focus on much narrower problems.

Types of transition in information-processing

Assembling a comprehensive and well organised collection of types of transition in biological I-P, and pre-biotic precursors, will require extending many of our ideas, a massive, long term, multidisciplinary task -

i.e. unlike typical funded 2-3 year projects with a few researchers and PhD students.

We should not presuppose a sharp, clear, boundary between non-biological and biological structures and processes:

seeking dichotomies and precise definitions often obstructs science.

The alternative to a dichotomy is not necessarily a continuum, with fuzzy transitions: it is also possible to have a very large collection of discontinuities, big and small, including branching and merging discontinuities.

This project overlaps with, but is different from attempts to understand transitions in (a) physical forms, (b) environments, (c) means of reproduction, (d) types of behaviour.

There is an incomplete, steadily growing, still disorganised, early draft, collection of different types of transition in I-P listed here

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

(with around 80 types distinguished at the time of writing this – and more to be added).

Comments, criticisms, suggestions, and substantive contributions to that list are welcome.

All suggestions incorporated will be acknowledged.

Eventually this could become an international, multidisciplinary collaborative project, comparable in some ways to the Human Genome Project, but more general.

AGI is only one of many research fields to which it is relevant.

Possible topics for more detailed discussion

(Most skipped in these slides, but discussed elsewhere on the Meta-morphogenesis web pages.)

Questions to be asked about biological information-processing:

- The main problem: What preconditions have to be satisfied for complex life forms and their mental processes to come from a cloud of dust? (not yet answerable!)

  All life involves **I-P**: many kinds of information, and many kinds of processing, on different spatial and temporal scales. What is "semantic" (non-Shannon) information? (Sloman, 2011a)
  Information-processing is (mostly) invisible, and (mostly) leaves no fossil records.
  Contrast varieties of: morphology, behaviours, environments (all usually easier to observe).
  Where/how does all the biological information originate? Discussed inconclusively in (Davies, 1999)
- What sorts of I & I-P can occur, in organisms and biological systems? Are there limits?
- What sorts of underlying mechanisms are required to support such I-P?
  - In what sorts of worlds could they occur? (Not all can occur in a 2-D world, like Conway's "Life".)
  - Which kinds of I-P can, and which cannot, be supported by Turing machinery?
    2-D and 3-D rotations can be described in TMs, but cannot have exact isomorphic models.
  - Which kinds of I-P can be supported by TMs, but have impossible space/time requirements?
  - Does chemical I-P have unrecognized powers? (E.g. efficiency gains for intelligent systems.)
- Is a systematic investigation of **transitions** in biological I-P feasible?
  - What **transitions in I-P** can occur, in evolution, development, learning, cultural change ?
  - We may need to understand biological precursors to current I-P mechanisms that remain part of the indispensable infrastructure (or substructure) for newer systems. (E.g. brains can't build brains?)
- Is this a worthwhile (massive) collaborative, multi-disciplinary project, or just foolish, extravagant, hand-waving, or...?
- **What are possible next steps?**

Most of the **requirements** are still to be identified

It is not easy to produce accurate descriptions of kinds of I-P that are

(a) products of biological processes, and

(b) still beyond what we can replicate in AI machines/robots

For example:

- Much of what is achieved by perception is not directly related to action, nor easily inferred from brain scans, e.g. perception of affordances, and their use in reasoning.

  Explained in more detail here: `http://tinyurl.com/CogTalks#gibson` (Sloman, 1996)

  What's vision for, and how does it work? From Marr (and earlier) to Gibson and Beyond

- Also geometric reasoning (e.g. about triangles, strings, part-built bird nests, worm & pinion gears).

  For example `http://tinyurl.com/CogMisc/triangle-theorem.html`
  Theorems About Triangles, and Implications for Biological Evolution and AI.
  How did biological cognition lead to Euclid's Elements? (Sloman, 2010b)

- Interplay of development and learning in altricial species.

  Why aren't all species "precocial" – with genetic information enabling independence from the start?
  Why are humans and hunting mammals born less competent than deer?
  (Sloman & Chappell, 2005; Chappell & Sloman, 2007) & also below.

- Human varieties of motivation and affect – and other types.
  See the Cognition and Affect Project `http://tinyurl.com/BhamCog#overview`

- and many, more!

- Capabilities that are specially hard to understand and replicate, include:

  The ability to enjoy music, or to find something funny, and use of self-monitoring self-modulating virtual machinery to detect existence of qualia (Sloman, 2010c). (The last is probably easiest.)

Biological transitions are discrete!

Reproductive changes involve new assemblages of molecules.

So the transitions must be discrete.

- It's not possible to add or remove a half, a quarter, a third.... etc. of an atom from a molecule.
- Between any two generations in a reproductive history there can only be a finite number of discrete changes.
  Continuous change would require infinitely many intermediate designs.
- Some transitions may make little difference to functionality at first, but later provide support for major changes, e.g. duplication of some information structure.
- Many of the discontinuities will be very minor and of little interest.
- Others are the main topic of this presentation: transitions in types of biological information-processing, in evolution, development, learning, and social processes.

  Conjecture:
  All major biological transitions build on collections of discrete transitions,
  some barely noticeable, others highly significant.

NB: Evidence for discontinuity is not evidence against evolution.

I don't deny that there are many unsolved problems, about what evolved, when it evolved, why it evolved, how it evolved, what the consequences were, what mechanisms were required, ... etc.

Some important transitions previously noted

John Maynard Smith and Eörs Szathmáry (1995) proposed that there are eight major transitions in evolution, summarised here:

`http://en.wikipedia.org/wiki/The_Major_Transitions_in_Evolution`

Transition from:	**Transition to:**
1 Replicating molecules	“Populations” of molecules in compartments
2 Independent replicators (probably RNA)	Chromosomes
3 RNA as both genes and enzymes	DNA as genes; proteins as enzymes
4 Prokaryotes (no nucleus)	Eukaryotes (nucleus)
5 Asexual clones	Sexual populations
6 Protists	Multicellular organisms - animals, plants, fungi
7 Solitary individuals	Colonies with non-reproductive castes
8 Primate societies	Human societies with language, enabling memes

They identified features common to these transitions:

1. Smaller entities come together to form larger entities.
2. Smaller entities become differentiated as part of a larger entity.
3. Smaller entities become unable to replicate in the absence of the larger entity.
4. Smaller entities become able to disrupt the development of a larger entity.
5. **New ways arise of transmitting information.**

   But that’s not the only kind of change in information-processing.

More transitions: we can generalise the last feature to include new information contents, new forms of representation and new ways of acquiring, processing, or using information.
These changes can have different sorts of trajectories in evolution and in development.
Many important information related changes have probably not yet been noticed.

The importance of transitions

Transitions in biological I-P involve different contexts, and different levels of abstraction, including changing relationships between genomes and I-P architectures and capabilities.

Besides the types of transition discussed by Maynard Smith and Szathmáry, there are others relevant to the evolution of information-processing capabilities:

- Transitions in the niches, or sets of requirements (including constraints and opportunities) that can play a role in natural selection (via changes in the environments in which organisms live, compete, collaborate, etc. or changes in their morphology).
- Transitions in the I-P competences and designs that are relevant to coping better or worse in meeting those requirements.
- Transitions in the implementations of designs: most abstract designs can be implemented in many different ways, and changes in implementation affect not only tradeoffs among the original requirements, but also possibilities for future changes.
- Developmental transitions in individuals – changes in requirements, designs, and implementations during development – only **facilitated** not **driven** by the genome.

  We need to take account of tradeoffs/interactions between
  – processes of evolution, and
  – processes of individual development.

Seeing how all these transitions interact in increasingly complex feedback loops may help to explain some of the drivers of evolution and perhaps thereby provide a better understanding of the products of evolution.

Analysing such processes may require new kinds of mathematics.

The view of (Farnsworth, Nelson, & Gershenson, 2012)

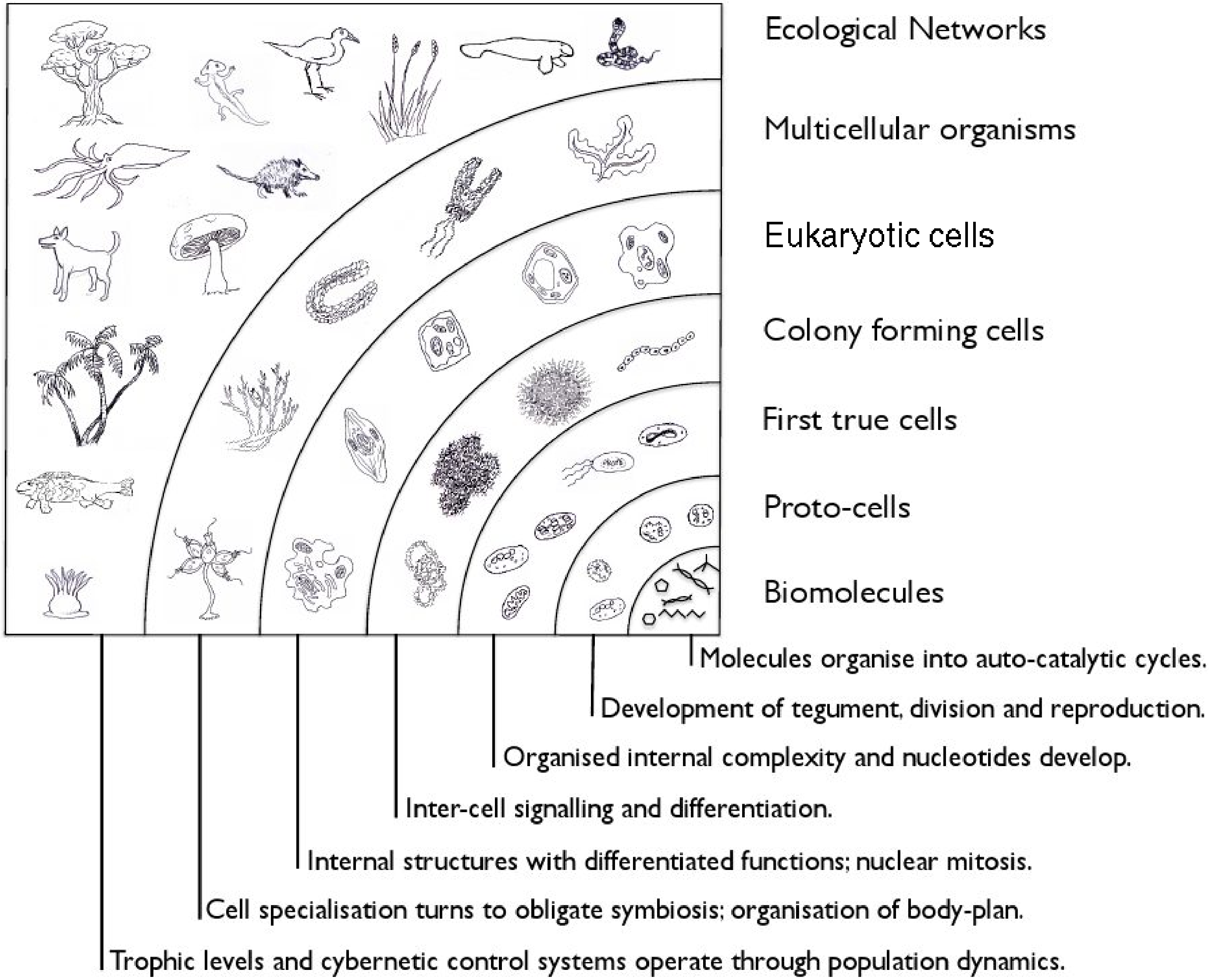
Their paper: **“Living is information processing; from molecules to global systems”**

Layered
products
of evolution:

Architecture of
biosphere?

What’s missing?

How many
intermediate
stages with
differences in
information
processing
are there?



The authors state:“...all inner layers exist concurrently at each level”

(Just one of many research groups now stressing biological information processing.)

Interactions at different organisational levels

(Farnsworth, Lyashevska, & Fung, 2012)

Organization Level	Interactions
life as a whole	global bio-geochemical networks
ecological communities	interspecific material and energy flows
populations - species	gene-flow, dispersal, evolution
multi-cellular organisms	organism societies + interspecific, e.g. parasitism
tissues, organs and organ systems	cellular communication and organ function
cells	specialisation and ontogeny: e.g. immune system
sub-cellular structures	catabolic autopoietic processes
molecular networks	metabolic and information processing
DNA sequences: codons to genes	coding and expression control
molecular surfaces	lock and key - enzymes

Table 1 (From the Farnsworth et al. paper)

A ten-level hierarchy of biocomplexity. Left column names the level of organization and right column gives examples of the complex interactions and processes that take place at that level, contributing to biocomplexity. Complexity is also added by interactions among levels, both upwards and downwards, producing feedback circuits. Interactions at every level and among levels constitute information processing.

(Adapted from Farnsworth, K., Lyashevska, O., and Fung, T. (2012). Functional complexity: The source of value in biodiversity. *Ecol Complex*, 11:46-52.)

Within each level we can look for many more subdivisions concerned with varieties of information-processing.

Other slides present examples.

Biology and AGI

Researchers who ignore the very many “design decisions” leading to new mechanisms, providing new functionality in biological information processing, risk continually (re-)inventing vastly inferior systems.

Worse: they risk failing to understand what the problems are, and failing to detect the inadequacies of their own designs.

Examples include AI/AGI vision researchers, and over-optimistic brain emulation researchers:

> Who fail to notice the main functions of vision, and the need for concurrent perceptual processing at different levels of abstraction. not merely concerned with spatial scale, i.e. “multi-window perception” vs “peephole perception”, illustrated later and in:
> `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#gibson` – also (Sloman, 2000, 2011b)

Other examples:

- Failing to understand requirements for offline (e.g. mathematical) intelligence.
- Failing to understand requirements for inner languages before communicative languages. (Sloman, 2008b)
- Failing to understand the diversity of types of motivation and affect required in intelligent systems (operating in different architectural layers) (White, 1959; Simon, 1967a; Sloman, 2009a; Wright, Sloman, & Beaudoin, 1996)
- Failing to understand the architectural requirements for meta-semantic competences needed for self-understanding and interaction with others (including offspring, prey, predators).
- Failing to understand the roles of meta-semantic competences in perception and action.
- Failing to understand the requirements for virtual machinery with self-monitoring and self-modulating capabilities. (Sloman, 2010a)
- Failing to understand the genetic prerequisites for matching or exceeding human rates, and human variety, of learning and development. (Chappell & Sloman, 2007; Karmiloff-Smith, 1992)

(Many AI/AGI developers seem to think of learning as data-mining.)

Biological processing of information = use of information

A common error is to think of computation merely as the manipulation of information structures, ignoring the fact that information is used in multiple, very different, ways.

A famous example of this error is John Searle's claim that computers merely perform syntactic operations and are incapable of understanding symbols they manipulate.

He distinguished the Strong AI thesis that such operations can replicate understanding and other thought processes from the Weak AI Thesis that they can at best simulate or model those processes, as a running computer program models, but does not replicate a tornado.
(Searle, 1980, 1984). Contrast (Sloman, 1986a) (Sloman, 1985)

(Since then the labels "Weak AI" and "Strong AI" have been hijacked for use with shallower distinctions.)

This narrow (erroneous!) view of computation is in part based on the theoretical models of computation arising out of the work of Frege, Russell, Church, Turing, Post and others: all of whom specify in great detail various combinations of abstract, formal (symbolic) operations that enable a machine to have the power to generate processes that can be interpreted as operations on numbers, equations, steps in a mathematical proof, etc.

Those authors (unlike Craik) did not consider how information-processing could actually be used in animals perceiving and acting on their environment, and used in future robots.

As a result there are still many thinkers (including philosophers who should know better) who write as if computation were merely a matter of performing mathematical/logical and similar operations on symbols (including bit-patterns) inside a machine.
I'll refer to "information-processing", instead of "computation" to help counter this neglect.

What is information – for organisms?

Recap: A key feature of (pre-Shannon) information is that it can be **used**.

Shannon stretched the word "information" in a new sense only loosely related to its (much older) main sense in ordinary language, and unfortunately confused many people who were not as clever as he was. (Shannon, 1948)

E.g. many now think information is something to be measured, which must be expressed in bit-patterns, and is intrinsically connected with computing machinery.
However, information content is often richly structured, used by humans and other animals, and usually needs to be described rather than measured: Information content is more like **shape** than like **length** or **volume**
Information content is more like a **machine** than like a **piece of string**.

What will counting bits tell you about this: "Never send to know for whom the bell tolls; it tolls for thee"?

The older, familiar concept of information **content**, allows information items to stand in various information relations:

If **I1** and **I2** are information items, possible relations include the following (just a sample):
– **I1** subsumes, entails, contradicts, is consistent with, is sub-contrary to, is evidence for, or generalizes **I2**.
– **I1** controls the derivation, interpretation, recording, invocation, or use of **I2**.
– **Above all information content can be used, and is used, for many purposes, including control.**

A measure of **amount** of information is much less useful – except for designers of transmission equipment, storage mechanisms, etc., who don't need to know what the information is about, what it is to be used for, why it will be used, etc. – they only need requirements for speed, storage capacity, reliability, etc.

You don't need to read a news report to find out how much (Shannon) information it contains, but you do need to read it to find out what the information is.

A theory of biological information and biological information processing needs to take account of what information contents are used and what they are used for.

Examples of the pre-Shannon notion of information

For organisms, information content is often relevant to selection and control of actions of various kinds.

Having a measure of the information available about something does not tell you whether you need to do anything, or what you can do: for that you need the information content.

Example:

Consider environmental information about a tree trunk or a boulder, available to an animal or robot that needs to climb over it, demolish it, move it out of the way, move it into a position where it can serve a particular purpose etc.

Different information subsets (about parts of the obstacle, relationships, constraints, possible bits to push, pull, rotate, or twist, and possible movements to make before or after manipulating it) are used at different stages of perceiving, reasoning, motive formation, planning, acting, and reflecting on what happened.

Those are information contents, where the structured information content is biologically important, rather than a numerical measure of amount of information.

Unfortunately most work on machine perception and robotics ignores most of the biologically important uses of information, especially use of information to understand what's possible. Not: what's probable. See also (Sloman, 1996, 2011b).

A bit is all you need (in a suitable context) to express whether a path is or is not blocked.

An agent who needs to unblock a path, or select an alternative route needs something much richer than a collection of bits, and also something more directly useful than an internal **model** (proposed in (Craik, 1943)).

A more detailed answer to "What is information?" is in this book chapter:

What's information, for an organism or intelligent machine? How can a machine or organism mean?
`http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#905`

Subtle harm sometimes done to science by technology

Many machine vision and robotics researchers assume that visual information enters an organism or robot in the form of a sequence of 2-D arrays of bit-patterns registering information about wavelengths (or frequency) and intensity.

This seriously distorts thinking about what vision is and how it works – and may contribute to inadequacies of current machine vision systems.

E.g. usually a camera is calibrated using some test grid when it is switched on and that calibration is used thereafter – but perhaps constant task-relevant re-calibration of orderings, not absolute values, is more appropriate to mobile animals?

> Questions about representations of straightness, curvature, angles between lines, etc. in images are trivial for current digitised images but they are far from trivial for brains – as Kenneth Craik noticed in his (1943), written before modern digital cameras were available to mislead researchers.
>
> It may turn out that not making use of the metrical structure inherent in current digitised images can lead to far more powerful and general visual systems for dealing with most environments, including environments without straight lines or plane surfaces.
>
> But different primitive optical sensors may be required, e.g. concentric rings of varied resolution sensors, and comparators for length, curvature, angle, rather than measures, etc.

We may learn more about what's required for matching the power and generality of visual functioning in humans, other primates, squirrels, birds, octopuses, etc., if we try to understand the transitions that occurred in visual information-processing in biological evolution, especially what differences they made to the organisms involved.
See also Gibson's ideas summarised in (Sloman, 2011b).

Compare spewing out or reading in bit-streams with moving/rotating a 3-D manipulator.
(Possibly controlled by bit-streams!)

Dynamical system architectures for embodied agents.

Focusing on information processing for online intelligence often leads to a focus on dynamical systems closely coupled with the environment.
(Gelder, 1998; Beer, 2000)

But that does not include requirements for offline intelligence.

The next few slides indicate some different possible ways of viewing an agent's (an animal's or a machine's) interaction with its environment.

Details may be far more complex than indicated here.

The simplest case is a system very closely coupled with its environment.

More complex designs can involve (among other changes):

- more sensors and effectors constantly interacting with the environment, using "online" intelligence.
- more advanced sensors and effectors interacting with the environment e.g. eyes with controllable vergence, higher resolution, etc., skin with more sensors, grippers with more degrees of freedom or more mobility relative to the torso, or greater strength,
- more diverse, or more sophisticated uses made of sensors and effectors (Sloman, 1993)
- more subsystems within the information-processing architecture de-coupled from the environment and performing tasks relevant to various kinds of offline intelligence.

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://videolectures.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.

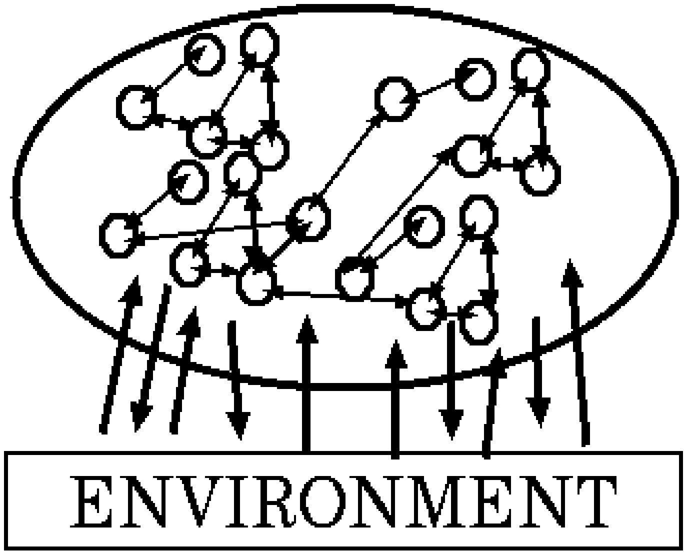
Processes involving use of information, not only about what does or doesn’t exist, but also information about things to achieve, and information about how to maintain, achieve or prevent things, have increasingly dominated changes on this planet.

(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.

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,
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.

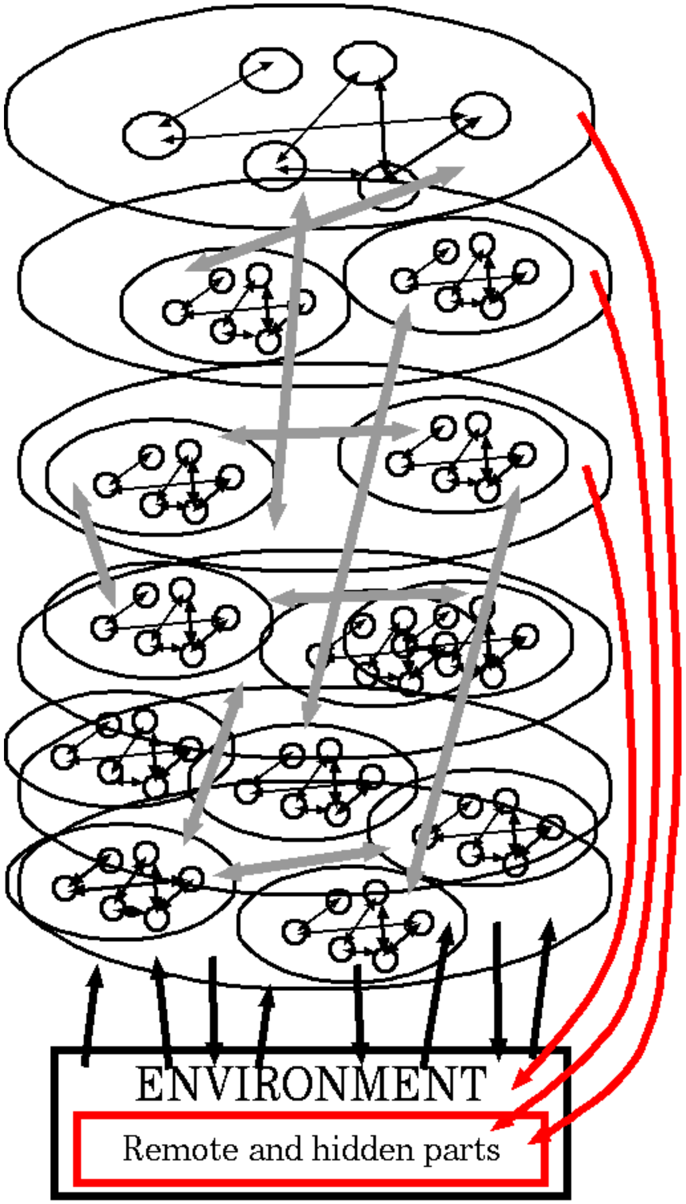
Dynamical systems II (including offline intelligence)

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?



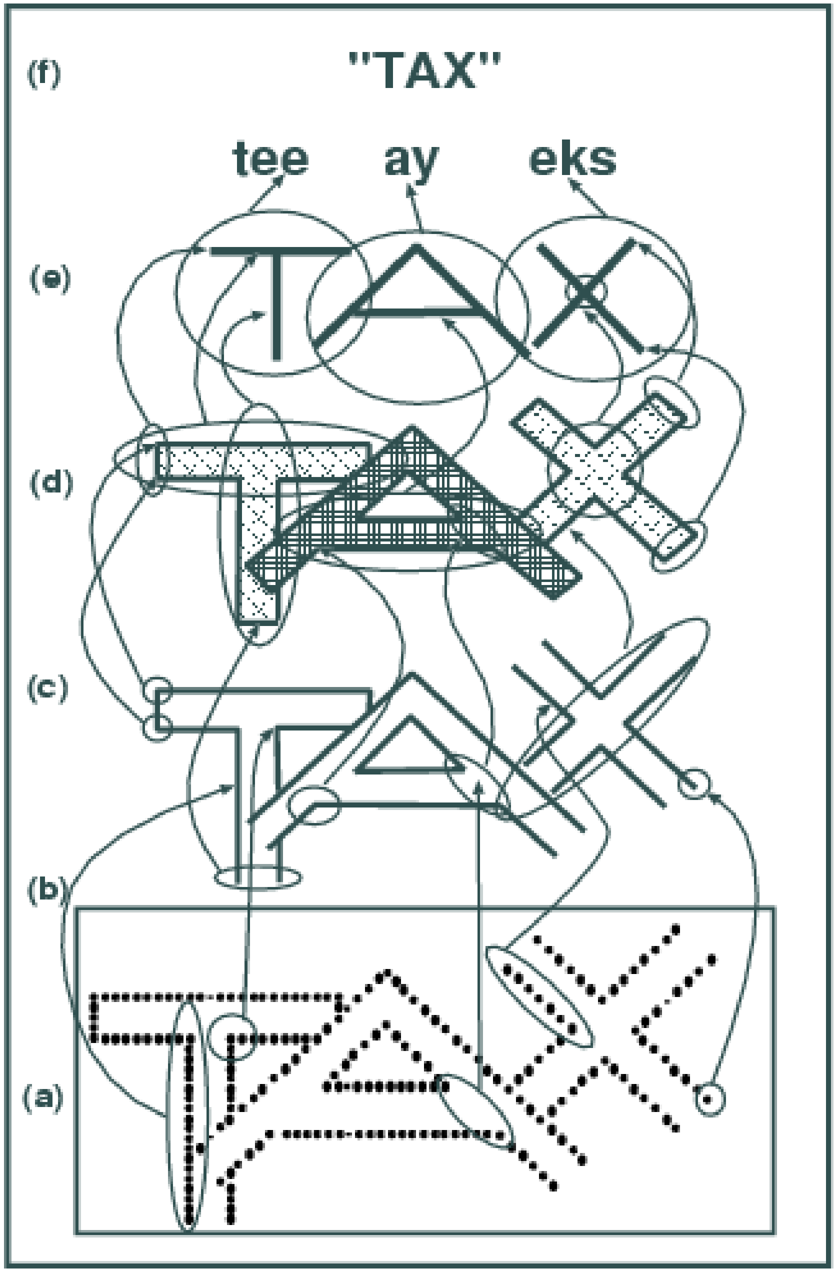
Do not believe symbol-grounding theory: use **theory-tethering** instead. (Discussed later.)
`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models`
(The “tethering” label was proposed, in discussion, by Jackie Chappell).

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 dynamical 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 and causal interactions.



Dynamical systems III

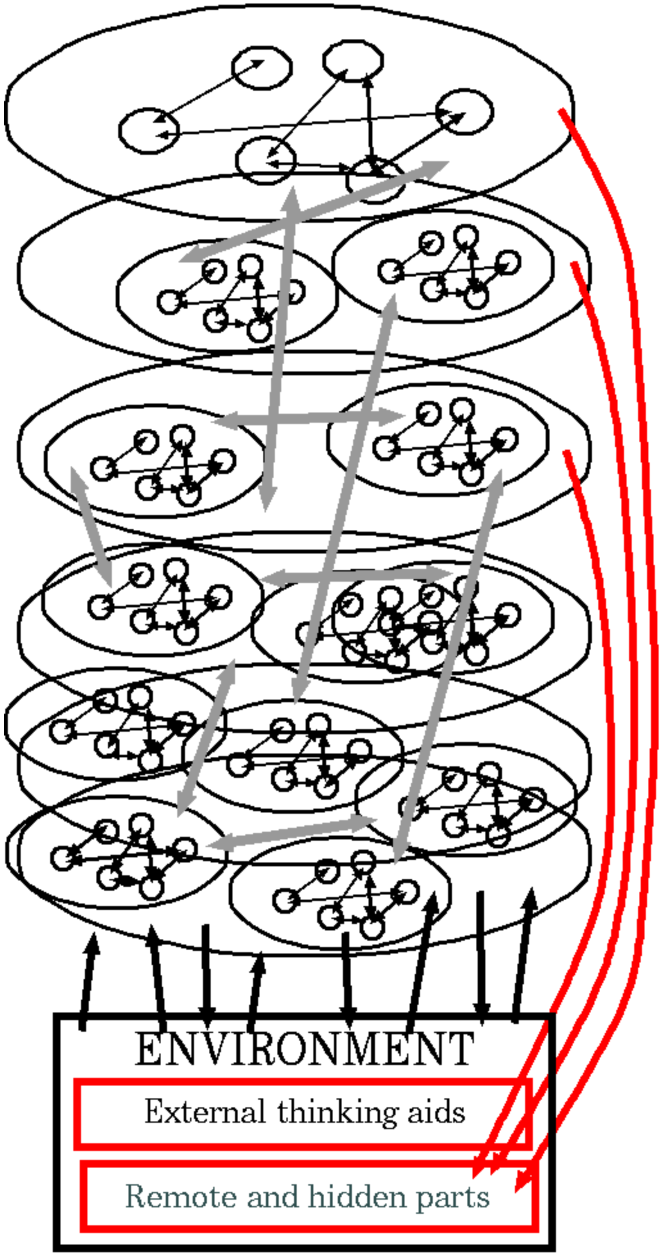
A point that is obvious to mathematicians, but usually ignored by roboticists:

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

Concepts are the building blocks of sentence-meaning, which can be used in constructing propositions (true, false, incoherent), questions, motives, plans, theories, arguments, etc.

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.)

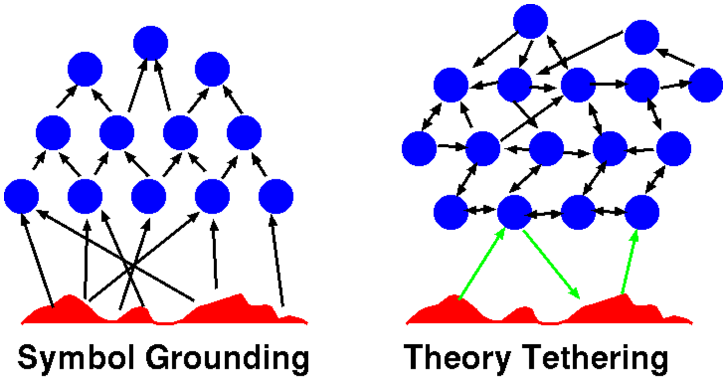
A set of available concepts, plus modes of composition, determine an ontology.

"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, 2007).



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, 2007). 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,

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, 2008c) 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`)

Alternative ways of viewing information-processing architectures are presented later: every complex system has multiple views.

Information for control vs communication

Control can involve communication, e.g. when a controller sends instructions to a agent, who obeys the instructions.

But control information may be acquired and used without communication.

Many key ideas about control were developed by Norbert Wiener (1961)

A link between communication and control exists in a computer with a component receiving a stream of instructions, including simple instructions to be obeyed, e.g. halt, or get next instruction, and complex instructions with items of data and an operator to be applied to the data, e.g. an operator specifying that a new list be constructed containing all items, or that the data be treated as numbers and added up.

That sort of functionality can be mapped onto operations in a Turing machine, or a computer.

However there were types of information-based control long before there were computers

E.g. music boxes, mouse-traps. thermostats controlling heating systems, watt governors controlling steam engine speeds, and windmills controlling the direction of their vanes, so as to get most energy from available wind,

And, long before those, there were far more sophisticated control mechanisms in living organisms, including homeostatic control of temperature, controlled transfer of waste products and other materials, release of chemical energy, control of infections, and, of course, controlled reproduction.

Communication requires a sender S, a receiver R, and a message M (or, more generally, something sent that can be interpreted as expressing the content M).

But an agent can simply acquire and use information to control its actions, without any sender, e.g. acquiring information about your path being blocked by a chair.

Analogical, pictorial doesn’t imply isomorphic

(This is a place-holder)

People discussing the contrast between

– Fregean forms of representation (using the distinction between functions and arguments) and

– Analogical forms of representation where not symbols but properties and relations denote properties and relations

often jump to the mistaken conclusion that the second sort must be isomorphic with what they represent.

This ignores the counter-arguments and counter-examples presented in the paper introducing the Fregean/Analogical distinction
(Sloman, 1971),(Sloman, 1978, Ch 7).

E.g. part of the usefulness of maps comes from their **not** being isomorphic with what they represent, and in many cases not even isomorphic with 2-D projections of what they represent, e.g. London Tube maps.

A 2-D photograph or painting is mathematically incapable of being isomorphic with the 3-D scene it depicts.
(A fact cubist painters noticed and attempted to partially compensate for.)

Ontological short-sightedness in researchers

Many researchers who rightly assume that organisms and robots need to be understood as information processing systems make wrong assumptions about the types of information processed.

In engineering terms: they get the design requirements wrong, and so produce theories and designs that meet the wrong set of requirements: a well known form of IT disaster, but not so well known as a form of scientific disaster.

Examples of erroneous, but not uncommon, assumptions:

- All intelligence is concerned with online interaction with physical environments.
- All knowledge of the environment is based on discovering structures and relationships between bit patterns in streams of sensory input and motor outputs.
- Agent ontologies need to be "grounded" in sensory-motor signal patterns.
- There is one best/right architecture for human-like intelligent systems.
- The function of perceptual systems (e.g. intelligent vision) is pattern recognition based on training on sets of examples.
- Languages are primarily required for communication – the use of languages for thinking, perceiving, planning, learning is derivative on communicative language becoming internalised. This ignores representational requirements for many animals.
  See discussion of primacy of internal languages. (Sloman, 1979, 2008b)
- All learning is data-compression using uniform algorithms.
- All motivation is based on anticipated reward.
  See discussion of "architecture-based" motivation. (Sloman, 2009a).

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?

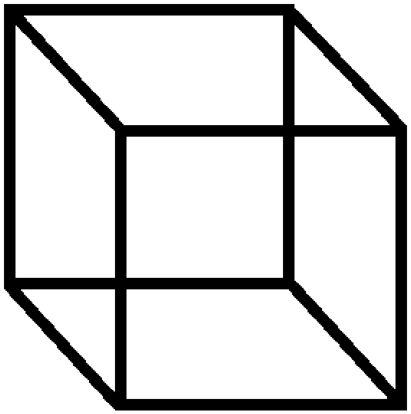
What causal relationships are perceived?

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

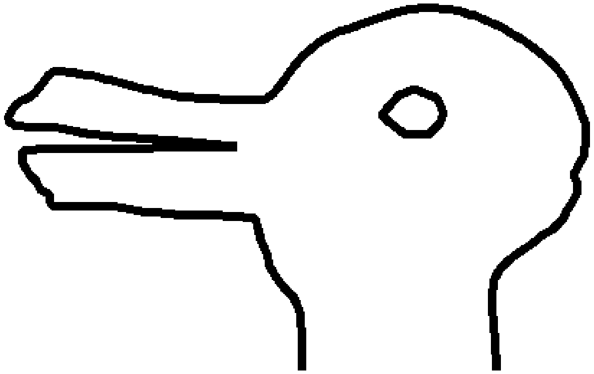
Ontologies for vision II

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



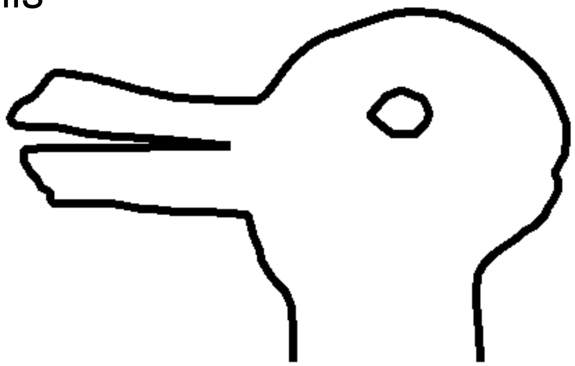
Duck-rabbit

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.

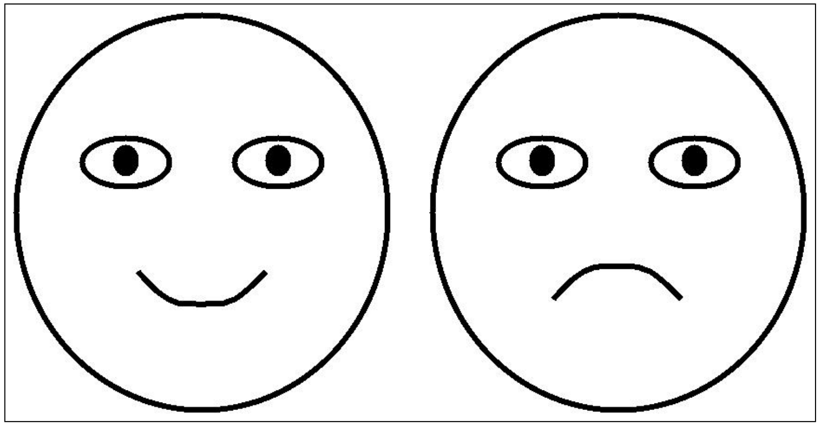
How would such possibilities, and constraints on possibilities be represented in the machine (or in an animal’s mind).

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?

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 different?

(They do for most people I've asked!)

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 above, 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. (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, ...

The next few slides aim to undermine the belief that seeing 3-d structures involves building 3-d models.

Visual information processing – for planning actions

Human vision includes many capabilities that are proving very difficult to replicate in current machines.

This may be because of our limited understanding of what the problems are.

(a) (b) (c)

You can look at the scene in (a) and think about how your fingers might have to move to produce the arrangement (b) (using your right hand, or your left hand, or both).

You can also think about a sequence of finger movements that would be able to perform the reverse rearrangement, from (b) to (a).

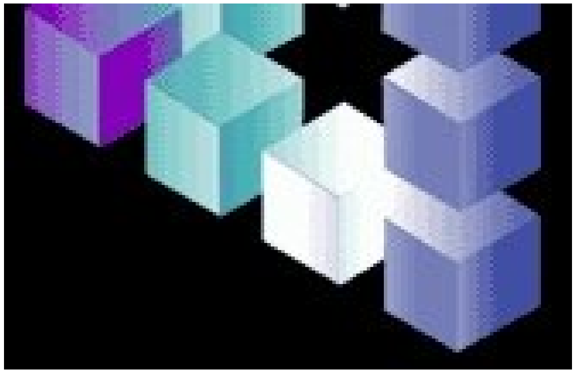
Current robots cannot do that unless highly trained on particular problem classes.
What is missing in their design?

It's not too hard to get a machine to derive image (c) from image (a): could that be a first step???

Notice that examples like this show that seeing is FAR more than attaching labels to portions of images – which is all that some current "vision" systems can do.

Seeing should not be confused with recognising: you can see things you don't recognise.

3-D structures – and possible actions



Here (on the right) is part of a picture by Swedish artist, Oscar Reutersvärd (1934) which you probably see as a configuration of coloured cubes.

As with the Necker cube you have experiences of both 2-D lines, regions, colours, relationships and also 3-D surfaces, edges, corners, and spatial relationships.

You probably also experience various affordances: places you could touch the surfaces, ways you could grasp and move the various cubes (perhaps some are held floating in place by magnetic fields).

E.g. you can probably imagine swapping two of them, thinking about how you would have to grasp them in the process – e.g. swapping the white one with the cube to the left of it, or the cube on the opposite side.

Although this is just a picture, you could see, and may actually have seen a collection of blocks in a similar relationship, except that the blocks on the right would normally be resting on blocks below, not floating above them.

However there could be blocks held in the depicted relationships supported from behind.

What happens when you see such a scene?

A common view (one interpretation of David Marr's theories in (Marr, 1982)) is that seeing such a scene involves building a model of it, or at least a description of the 3-D structures in some formalism that allows pictures of the scene to be generated from different viewpoints. (A common test for 3-D stereo vision systems.)

But that cannot be what your visual system does.... as we'll see shortly.

Experienced Possibilities, and representations

All of this suggests that much of what is experienced visually is collections of possibilities for change and constraints on change –

where these possibilities are attached to various fragments of the scene (via fragments of the image).

I.e. the possibilities and the constraints (e.g. obstacles) are located in 3-D space, in the environment, just as the objects and object fragments are.

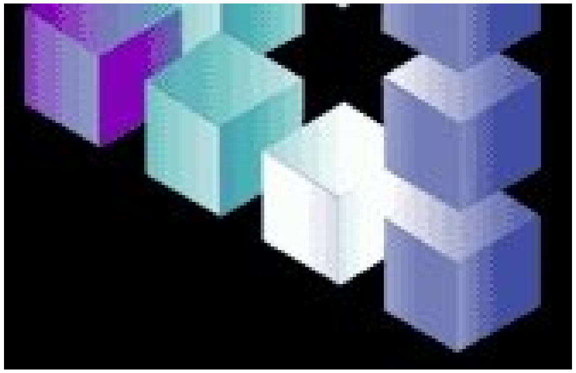
Question: how are all these information fragments represented?

- Anyone who thinks it is possible for organisms or robots to do without representations must think it is possible for them to do without information.
- For information must somehow be encoded and whatever encodes the information is a representation – which could include a physical or virtual structure and much of its context.
- The representations used may be physical (neural, chemical, electronic) structures or structures in virtual machines – like the information a chess program has about the current state of the board, and the information about possible moves and their consequences.

However, the picture on the previous slide can be extended to show more fragments of the same scene.

The original Reutersvaard picture, like the Penrose triangle produced later, is a picture of an **impossible** scene composed of cubes in a triangular configuration.

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

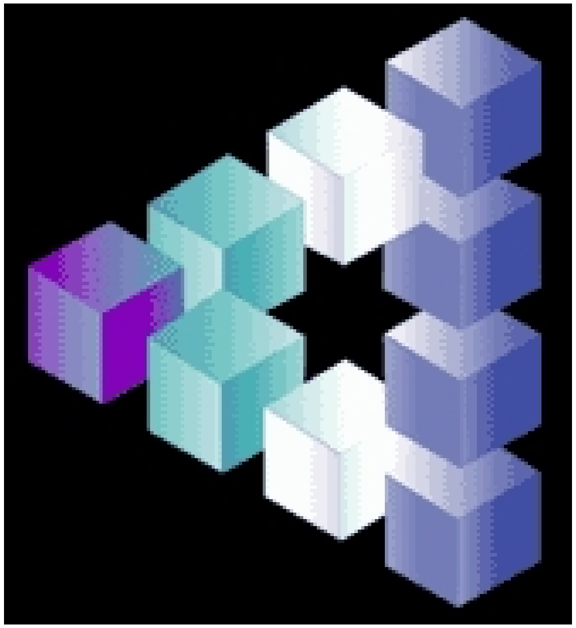


As before, here is part of a picture by Swedish artist, Oscar Reutersvärd (1934) which you probably see as a configuration of coloured cubes.

As with the Necker cube you have experiences of both 2-D lines, regions, colours, relationships and also 3-D surfaces, edges, corners, and spatial relationships.

You probably also experience various affordances: places you could touch the surfaces, ways you could grasp and move the various cubes (perhaps some are held floating in place by magnetic fields).

E.g. you can probably imagine swapping two of them, thinking about how you would have to grasp them in the process – e.g. swapping the white one with the cube to the left of it, or the cube on the opposite side.



The second picture on the right (from which the first one was extracted) has a richer set of 2-D and 3-D contents.

Again there is a collection of 2-D contents (e.g. a star in the middle), plus experience of 3-D structures, relationships and affordances: with new possibilities for touching surfaces, grasping cubes, moving cubes.

The picture is outside you, as would the cubes be if it were not a picture. But the contents of your experience are in you: a multi-layered set of qualia: 2-D, 3-D and process possibilities.

But the scene depicted in the lower picture is geometrically impossible, even though the 2-D configuration is possible and exists, on the screen or on paper, if printed. **The cubes, however, could not exist like that.**

So your qualia can be inconsistent!

That's impossible according to some theories of consciousness.
(Qualia can also go unnoticed: See `http://tinyurl.com/uncsee`)

Representational implications

When we see the whole impossible scene (which young children may not notice is impossible) it’s clear that despite impressions given by introspection, what we see cannot be a 3-D model isomorphic with the perceived structure.

That’s because the perceived structure cannot exist, so a model isomorphic with it cannot exist.

This implies that the internal representation of the information gained from Reutersvaard’s triangle and from Escher’s pictures, is more abstract than introspection reveals.

Logical sentences can easily describe something impossible:

That cat is black all over and the cat is not black.

I suspect that vision (and more generally animal representations of spatial structure) uses not sentences but a hybrid form that preserves some of the structure of the projection of the scene, in combination with something like symbolic, logical descriptions,

because that’s what’s useful for recording affordances that can later be used in selecting or carrying out actions.

Precise details of that proposal are still not clear: it could involve a data-structure with parts partly in registration with the optic array whose contents are something like condition-action rules – if this object moves in direction X then...

(Compare aspect graphs.)

Thanks to Jeremy Wyatt for useful, still incomplete discussions of these ideas.

There are many sorts of things humans can see besides geometrical properties:

- that one object is supported by another,
- that one object constrains motion of another (e.g. a window catch),
- that something is flexible or fragile,
- which parts of an organism are ears, eyes, mouth, bill, etc.,
- which way something is facing,
- what action some person or animal is about to perform (throw, jump, run, etc.),
- whether an action is dangerous,
- whether someone is happy, sad, angry, etc.,
- whether a painting is in the style of Picasso...
- what information is available about X
- which changes in the scene or changes of viewpoint will alter available information about X.

and other epistemic affordances.

Reading music or text fluently are important special cases, illustrating how learning/training can extend a visual architecture. (Sloman, 1978, Ch.9)

Open questions:

- What architectures, mechanisms, forms of representation, make these possible?
- How do the above competences develop in individual humans?
- Which of these capabilities (if any) are present in other species?
- How can these capabilities be given to machines?

Attaching affordance information to scene fragments

I have argued that a lot of the content of visual perception (at least in adult humans) takes the form of information about what is and is not possible (possibilities and constraints) in various locations in the perceived scene – along with realised possibilities when processes are perceived

(e.g. objects or surfaces moving, rotating, deforming, interacting).

- The Reutersvaard image shows how quite rich collections of possibilities can be “attached” to various scene fragments,
- Where some of the information is about possibilities and conditionals:
  - X is possible, Y is possible, Z is possible
  - if X occurs then, ...
  - if Y occurs then, ...
  - if Z occurs then, ...
- including “epistemic affordances”

  (if motion M occurs the information I will become available/or become inaccessible).
- This can be seen as (yet another) generalisation of aspect graphs which encode only information about how changes of viewpoint (or rotations of a perceived object) will change what is and is not visible.

  `http://people.csail.mit.edu/bmcutler/6.838/project/aspect_graph.html#1`
- An important special case of this idea concerns the effects of different kinds of material on what would happen if various things were to happen:

  e.g. material being: rigid, plastic, elastic, liquid, viscous, sticky, etc.
  See `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#babystuff`

The idea of generalising the concept of “aspect graph” as suggested here arose in a discussion with Jeremy Wyatt several years ago.

The dynamics of co-evolution

As the next slide suggests, in biological evolution it's not only designs and implementations that change, but also requirements, or niches, sometimes as a result of purely physical changes in the environment, e.g. global freezing, volcanic eruption, or an asteroid impact, and sometimes as a result of biological changes, e.g. changes in predators, prey, plants used for grazing, etc.

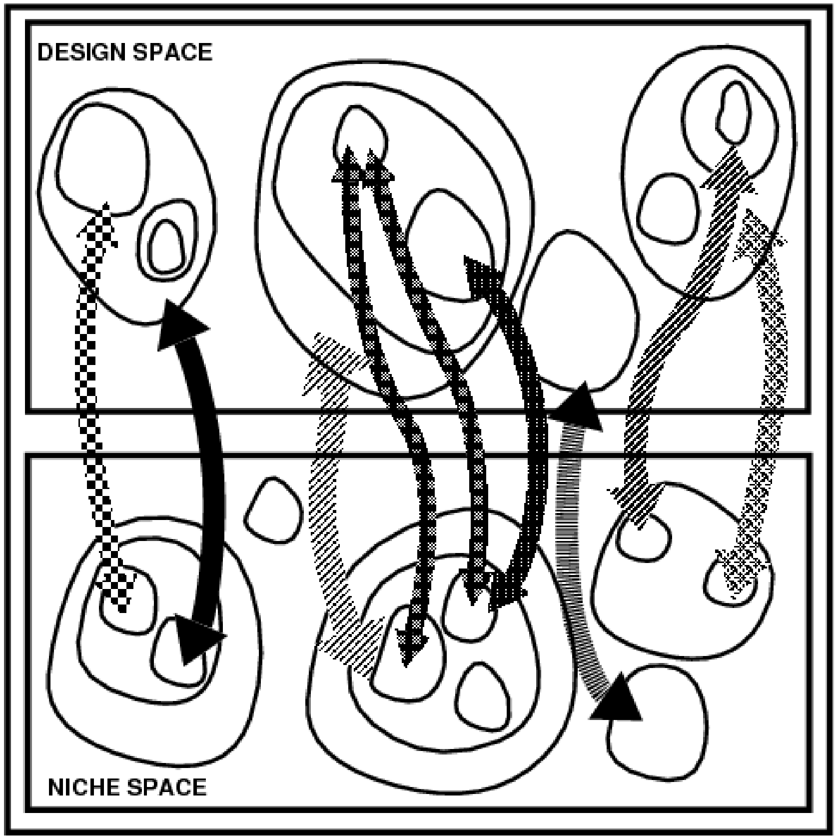
So within the biosphere, there are many species changing in parallel and many niches changing in parallel and many causal interactions (including feedback loops) linking those changes.

I suspect we do not have appropriate mathematical or computational tools for modelling or reasoning about such processes in any detail, though we can reason about them informally.

Some aspects of the complex network of relationships that change during evolution, are depicted crudely in the next slide.

[References needed]

Varieties of designs and varieties of requirements



Relationships between designs for information-processing systems and niches (sets of requirements) are not scalar values, such as fitness measures.

Instead, structured descriptions of matches and mismatches are needed.
(Compare consumer product reports and this analysis of “better”: (Sloman, 1969).)

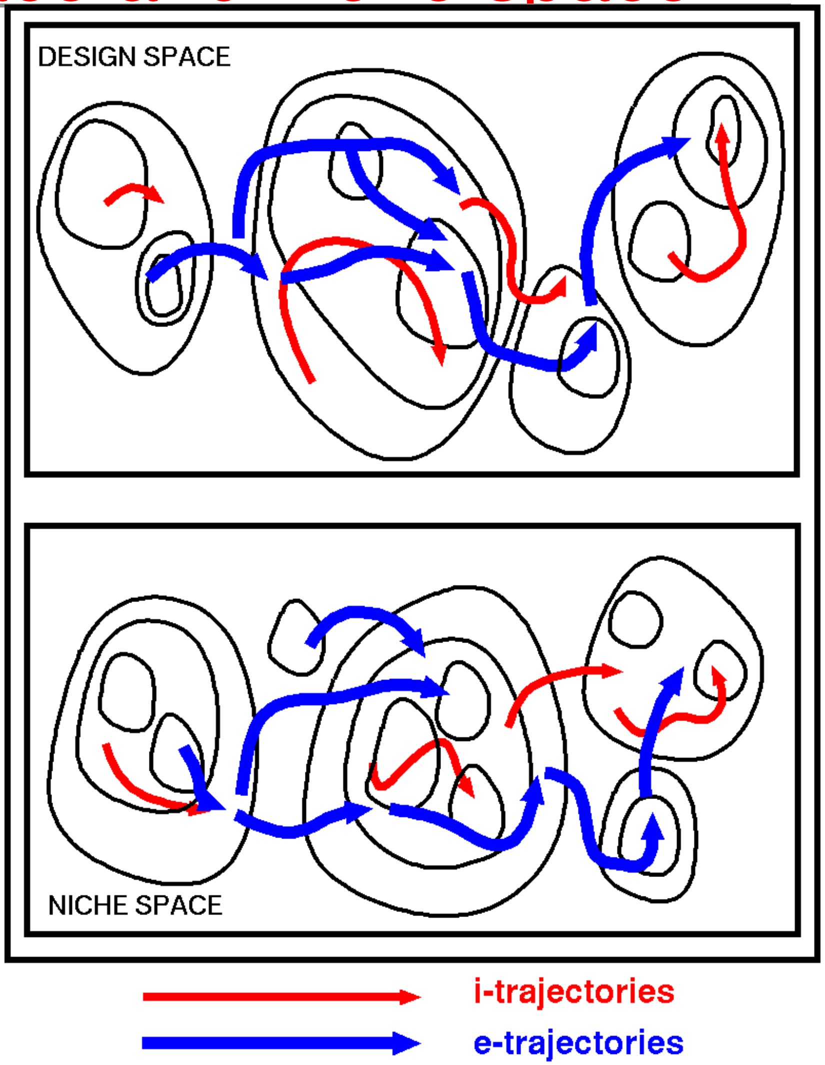
The diagram indicates crudely that the set of possible designs has various regions with commonalities across discontinuities, and likewise the set of possible niches (sets of requirements).

There are not one-one but many-many relationships between designs and niches.
(The next slide depicts trajectories.)

During evolution, individual development, individual learning, and social/cultural development and learning there are trajectories of change (including discontinuous change) both for the designs and for the niches.

There are also complex causal influences between designs and niches with feedback loops that modify the feedback mechanisms, unlike most human-designed feedback control systems. (Compare (Kauffman, 1995))

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.)

Varieties of designs meeting various sets of requirements

Increasingly sophisticated information-processing architectures result from accretion of mechanisms and resulting competences – driven by changes in the environment, including changes in other organisms, and also previously evolved structures, mechanisms, and competences.

- Oldest – purely reactive systems: microbes to plants and insects and beyond?

  Getting more complex: different sensors, effectors and reactions; some reactions become conditional, some become patterns or schemas that are instantiated according to context, and some become temporally extended (e.g. feeding behaviours, mating rituals). Some become learnt not innate.

- Behaviour selection expands to use more than current sensory inputs

  Using: previously acquired information (particular facts and generalisations), inferred and predicted information, preferences, values, goals,... all of which impact on information-processing requirements.

- Increasing sophistication in deliberative competences

  Goals generated further into the future; planning more and more steps ahead (using increased temporary memory), exploring randomly selected alternatives, then systematically generated alternatives, then optimised planning strategies, ...

- Addition of meta-semantic and metacognitive competences.

  Using information about self- and other- related information contents.

Trajectories in design space, and in niche space can interact, producing very complex forms of feedback in ecosystems containing multiple evolving types of organism.

I think that's another way of making the same sort of point as Kauffman makes in the later chapters of this book: (Kauffman, 1995)

However, it should be possible to be more specific about evolutionary changes in information processing functions, mechanisms and architectures.

The structure of design space for architectures

The next few slides present the CogAff schema, which provides a simplified way of thinking about different classes of information processing architecture in terms of the sorts of competences they include.

A large variety of different I-P architectures can be thought of as all providing different ways of instantiating the same abstract schema for architectures, the CogAff architecture (so-named because it allows for Affect in addition to core forms of cognition).

By presenting more recently evolved architectures as derived from older versions with more restricted resources we invite consideration of how those older systems might have been successful.

That can open up the question whether more recently evolved architectures still contain some of the earlier mechanisms that made the older simpler architectures self-sufficient.

If some of the older mechanisms are still required for the more complex specifications to be met, that should influence proposed designs for architectures.

We start with a simplified model of what an ancient species might have done, and then add more sophisticated components.

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/`

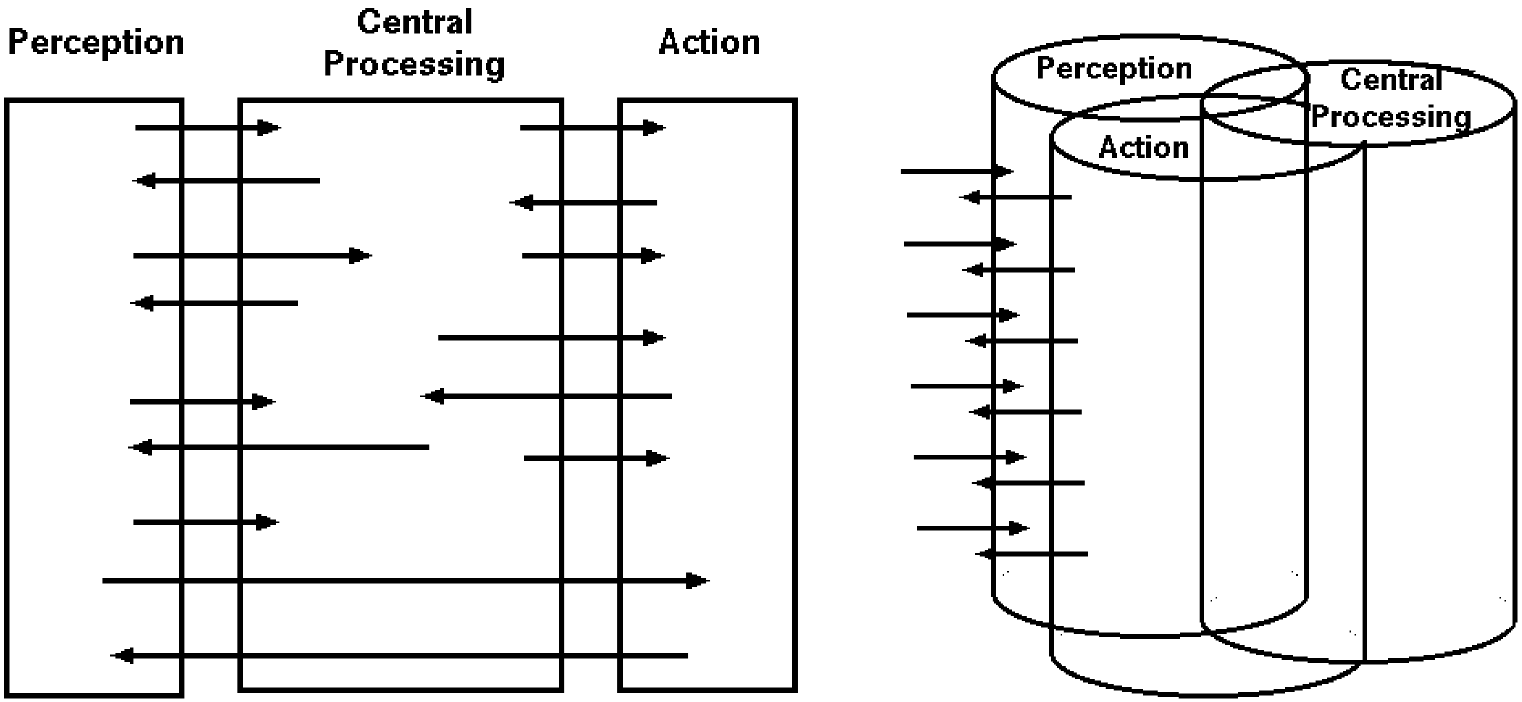
NB: this is not an architecture. I call it the “CogAff Architecture **Schema**” because it can be instantiated in many different ways, producing different I-P architectures meeting different sets of requirements.

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):


(a) Separate Towers

(b) Overlapping Towers

Patterns of information flow may be different in different cases.

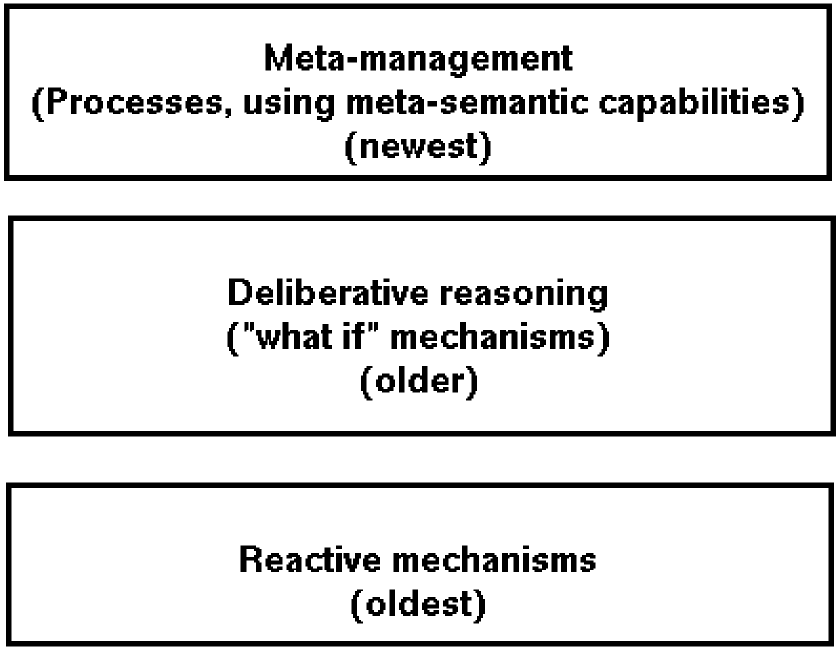
J.J. Gibson and others have emphasised the need for the right hand view of perception: perception is generally active, not passive reception.

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):



Meta-management
(Processes, using meta-semantic capabilities)
(newest)

Deliberative reasoning
("what if" mechanisms)
(older)

Reactive mechanisms
(oldest)

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.

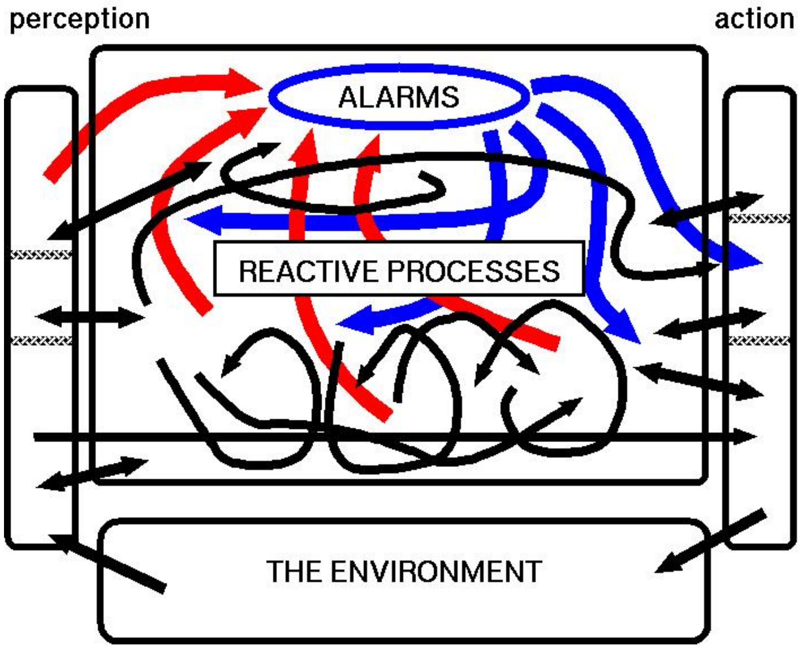
There are usually trade-offs: no single best design for any environment.

Special case of the CogAff Schema: Purely reactive

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?

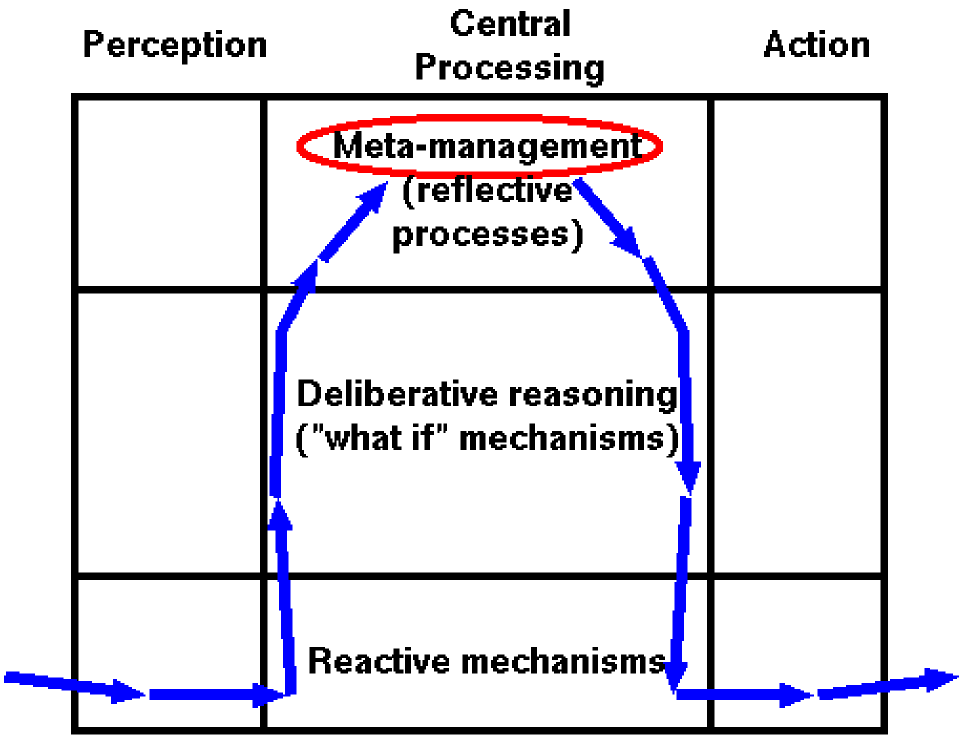
A fly that detects the approaching swatter will immediately abandon finishing its lunch and buzz off.

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)

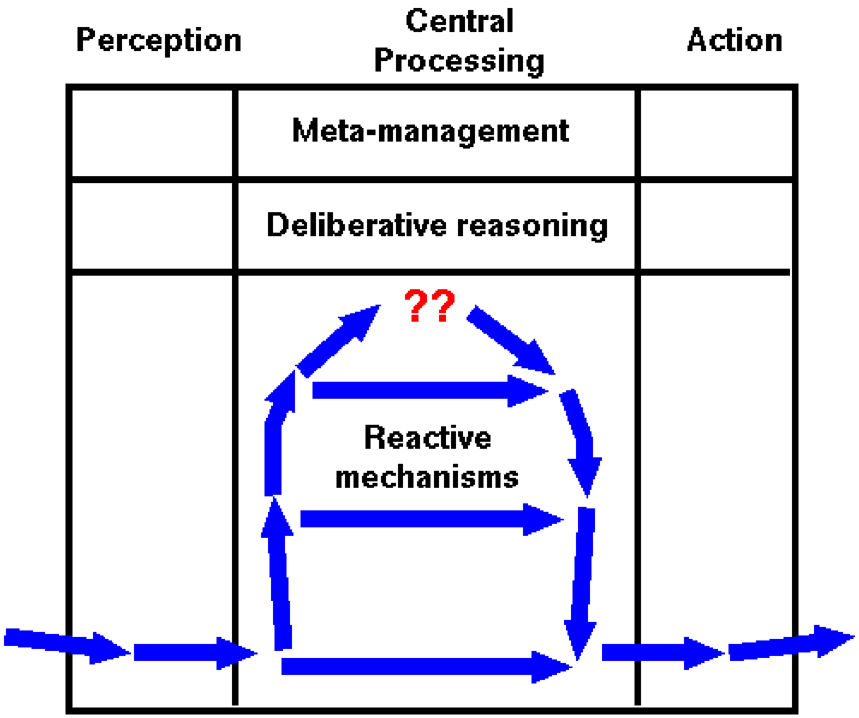


“Omega”-type architectures, like many others, use “peephole” perception and action, restricting input/output processes to very low level sensing and acting signals.

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)”

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?
(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`

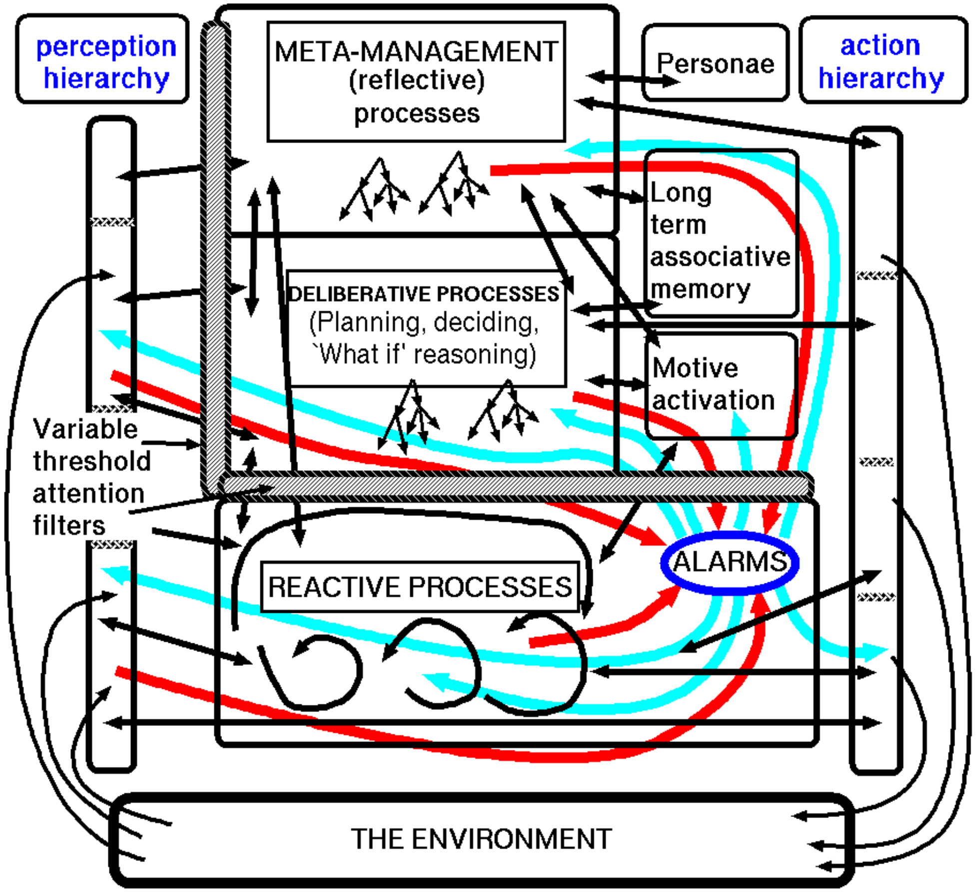
Compare I-P requirements for: larva → pupa → butterfly stages of development.

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.

Recapitulation

The Universe contains
structures and processes, involving

— Matter

— Energy

— Information

in various combinations.

And various kinds of virtual machinery....

“Information” not in Shannon’s sense, but something like the ordinary sense (information is about something, actual or possible. (Sloman, 2011a)

All forms of life involve
Matter, energy, and Information.
Forms of Information-Processing evolve –
producing new forms and new mechanisms of change

If we can identify many of the unobvious transitions from earliest to most recent forms of information processing that may help us understand better what sorts of competences and what sorts of information-processing mechanisms are involved in the newer, more complex organisms, for instance humans.

The study of such transitions seems still to be in its infancy: that’s not what researchers have been looking for.

High level overview

Transitions can occur in parts of organisms, in whole organisms, within a species, in interacting groups of species, in societies, and in environments (though organisms are part of the environment for conspecifics and for others).

Sample Types of transition:

1. Change of physical shape (in individual, in species)
2. Change in physical behaviour (in individual, in species)
3. Change in information processing (in individual, in species)
   (including control of growth, metabolism, immune system, processing of perception, motive formation, motive selection, action selection, action control, learning, reasoning, ...)
4. Change in developmental trajectory (physical, non-physical)
5. Change in what can be learnt (in individual, in species)
6. Change in type of interaction between individuals
   (in same species, across species, within 'family unit', prey, predators, others...)
7. Change in type of social organisation
   (including forms of collaboration, forms of nurturing, forms of education, forms of competition)
8. Changes in mechanisms of evolution (evolution of evolvability (Dawkins, 1988))
9. Changes in mechanisms of development
10. Changes in mechanisms of learning
11. Changes in mechanisms of interaction
12. Changes in mechanisms of self-monitoring, self-control
13. Introduction of new virtual machines, new forms of representation, new ontologies, new architectures
14. And others ...

These changes can interact and influence one another...

Information-processing can be invisible

In the past, the objects of scientific study could be perceived, handled, measured using physical devices.

But the complexity and rapidity of state changes required for processing of information rules out the use of physical machinery that cannot reorganise itself fast enough.

Turn your head at right angles very quickly to look in a new direction.
How long does it take for new percepts to form?

In computers, the need for very rapid reorganisation led to the construction of running virtual machines (RVMs) whose nature and function is very different from the nature and function of the underlying physical machinery – and can be used for very different virtual machinery at different times.

Conjecture:

Some organisms had the same problem (need for rapid structural reorganisation) and evolution produced the same general sort of solution – but

- the problems are more complex and diverse than those solved by human engineers
- the solutions are more complex and diverse and still barely understood
- the processes by which the biological designs were produced and tested are themselves among the most complex, and least understood.
- the underlying mechanisms used are very different from computers – and some of the differences may play a crucial role in biological intelligence.

Virtual machinery is invisible, but very powerful.

(Sloman, 2009c) (Sloman, 2010a) (Sloman, 2013 In Pressa) (Sloman, 2013 In Pressb)

Information and control 1

The most important feature of information is its potential for control.

Biological computation is primarily concerned with control, though one of evolution's side effects is that sometimes information is processed without actually being used for control.

The use of information for control can take several different forms, as discussed previously.

- Ballistic control: information is processed and immediately used in some action such as jumping, or throwing, that initiates a process where no control, or in some cases very little control, is possible after initiation.

  Examples: jumping over a gap, jumping up to reach something, throwing something at a target, throwing something to get rid of it.

  Indirect ballistic control occurs when a command is given to some other agent who then obeys it without the commander having any further influence over the process.

- Online control: information is continually taken in and processed during performance of some extended action or sequence of actions.

  Examples: many varieties of homeostasis, including temperature control, speed control, flow control, direction control, control of approach to something to be grasped (where several changing relationships may have to be monitored in parallel)

- Offline control: information about some object or process or state of affairs is acquired and used in order to gain understanding of possibilities for action by the processor, or possibilities for various kinds of externally initiated process.
  This can include deriving constraints on possibilities, or consequences of realising some possibilities.

  Examples: planning, explaining, predicting, designing, wondering about, reminiscing.

Many varieties of offline intelligence are important for humans, and still missing in AI.

Information and control 2

The most important feature of information is its potential for control.

There are many different ways in which information can be used to control something – though only a small subset is typically taught in CS degrees.

This very old concept of (semantic) information is not defined in terms of bit patterns, discreteness, or computers (recent man-made information processing machinery), although those can also provide examples of information processing.

Semantic content is not a scalar measurable – not a quantity.

Contrast Shannon's idiosyncratic use of "information".

Ways in which information contents can be related to other contents

- I1 includes I2
- I1 overlaps with I2
- I1 is inconsistent with I2
- I1 can be derived from I2 with additional assumptions
- I1 makes I2 possibly true
- I1 prevents I2 becoming true
- I1, I2, I3, ... can be abstracted to form generic I(x)
- I1 refers to I2 (as in meta-cognition, self-awareness, etc.).

.... and many more

What's (Semantic) Information?

See a preliminary answer in a chapter in the book *Information and Computation* (also online). `http://tinyurl.com/BhamCog/09.html#905` (Sloman, 2011a)

This is very different from, but loosely related to Shannon's notion of "information" (a scalar)

Detached/abstract computation

The various types of control-related or potentially control related can be contrasted with completely abstract, “detached”, computations,

e.g. based on a specification of

- a n-ary function F,
- a set of inputs X1, X2, ... Xn, and
- an instruction to compute F(X1, X2, ...Xn)
- which initiates a process of working out the value of F(X), treated as a purely mathematical (though not necessarily numerical) problem.

For example, X1 could be a set of rules for a grammar, X2 a lexicon, and X3 a list of symbols from the lexicon. F could name a function that takes a grammar, a lexicon and a sentence, and returns either a parse for the sentence, or the symbol “failed”.

The instruction F(X1, X2, X3) could specify an exhaustive search for a way of parsing X3 in accordance with X1 and X2, which returns a parse tree if one is found, otherwise “failed”.

A more sophisticated version could allow for multiple parses and return a possibly empty list of parse trees.

Alternatively X1, X2, ... Xn-1 could specify syntactic rules, and a set of axioms and inference rules while F(X1, X2, ...Xn) is an instruction to search for a derivation of Xn, starting from axioms and rules X1,...Xn-1.

Such “detached” computations may relate to important mathematical problems whose solutions can potentially play a role in engineering design, or planning an action, or analysing sensory inputs, but those applications need not be taken into account when the computation is performed, or checked.

Many people, including John Searle (1980), regard all computations as essentially detached, with the machine concerned taking no account of any meanings or truth conditions of the axioms, proof steps or result Xn, and not using any derivation found for some other purpose.

Online vs offline intelligence

Often control requires working things out in advance

E.g.

- Working out which materials are needed to build something before fetching the materials
- Anticipating something that needs to be prevented and working out how to build something to prevent it happening.

An important type of product of biological evolution:

increasing the ratio of non-interactive to interactive cognition.

The non-interactive computations are closely related to generic aspects of the environment,

e.g. trying to explain what happened, or did not happen, or trying predict what will happen, or to generalise from previously observed cases.

Those are more concerned with possible structures and processes, past, future, in remote places: not necessarily problems here and now.

Enthusiasts for embodied cognition, and enactivism tend to ignore or forget offline intelligence.

(I'll return to this later if there's time: if not, look at these triangle theorems: `http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html`)

Examples 1

If the spiral rotates so that the hole on the left moves up, what will happen to the cog wheel?

Assume both are on axles

free to rotate

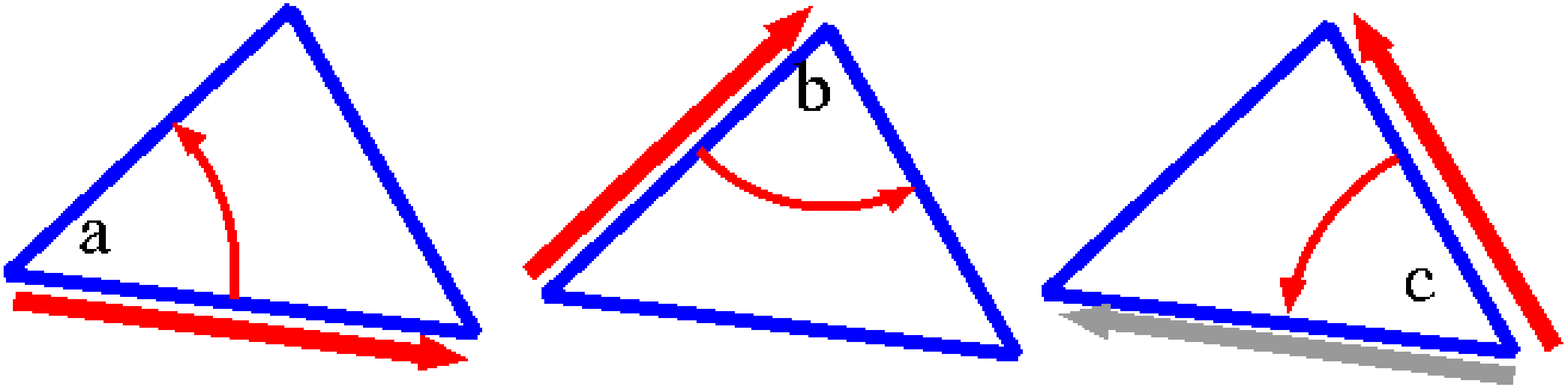
in fixed locations.

Question:

What sort of (offline) reasoning mechanism can work out how the cog will move as it is rotated, or how the behaviour and usefulness of some other mechanism will be affected by how it is assembled?

Picture from Wikimedia. [REF]

Examples 2



What does this tell you about the sum of the three angles, a, b, and c?

(Mary Pardoe's idea.)

Some examples of products of the process that we still cannot replicate in AI machines/robots

How could a machine be made to understand what's going on and use its understanding to design a new machine or re-deploy this one?

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

Mathematical competences and biological evolution

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

A QUESTION:

Who taught the first mathematicians?

ANSWER

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

That's the process educators are contributing to.

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

But the opportunities are not being taken.

From reaction to motivation

We need a better understanding of the diversity of types of motivation and affect required in intelligent systems (operating in different architectural layers) (White, 1959; Simon, 1967a; Sloman, 2009a; Wright et al., 1996)

TO BE ADDED

More on varieties of motivational mechanisms and meta-mechanisms.

Including

- motive generators
- motive comparators
- motive generator generators
- motive comparator generators
- motive generator comparators
- motive comparator comparators
- ..... and so on recursively.

We can’t do it all at once

I can’t yet go all the way back to a cloud of dust:

But I think there’s something important to be learnt by studying chemical information processing mechanisms – required before brains existed and still used, for many biological control mechanisms, including mechanisms for building brains.

Important features of chemical mechanisms

- Structures capable of moving around and rearranging themselves continuously are also capable of “snapping” into fixed relationships that are resistant to perturbation.
- Mechanisms for information processing that can “interface” directly with some aspects of the environment, instead of requiring special purposes interfaces like AtoD and DtoA converters required to enable turing machines to interact.

Products of evolution include many mechanisms with those properties that are able to build additional mechanisms with those properties.

Meta-morphogenesis:

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

- E changes B
- B changes E
- Bi changes Bj
- Combinations of Ei, Bi, Bj, ... cause changes in B .... 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:

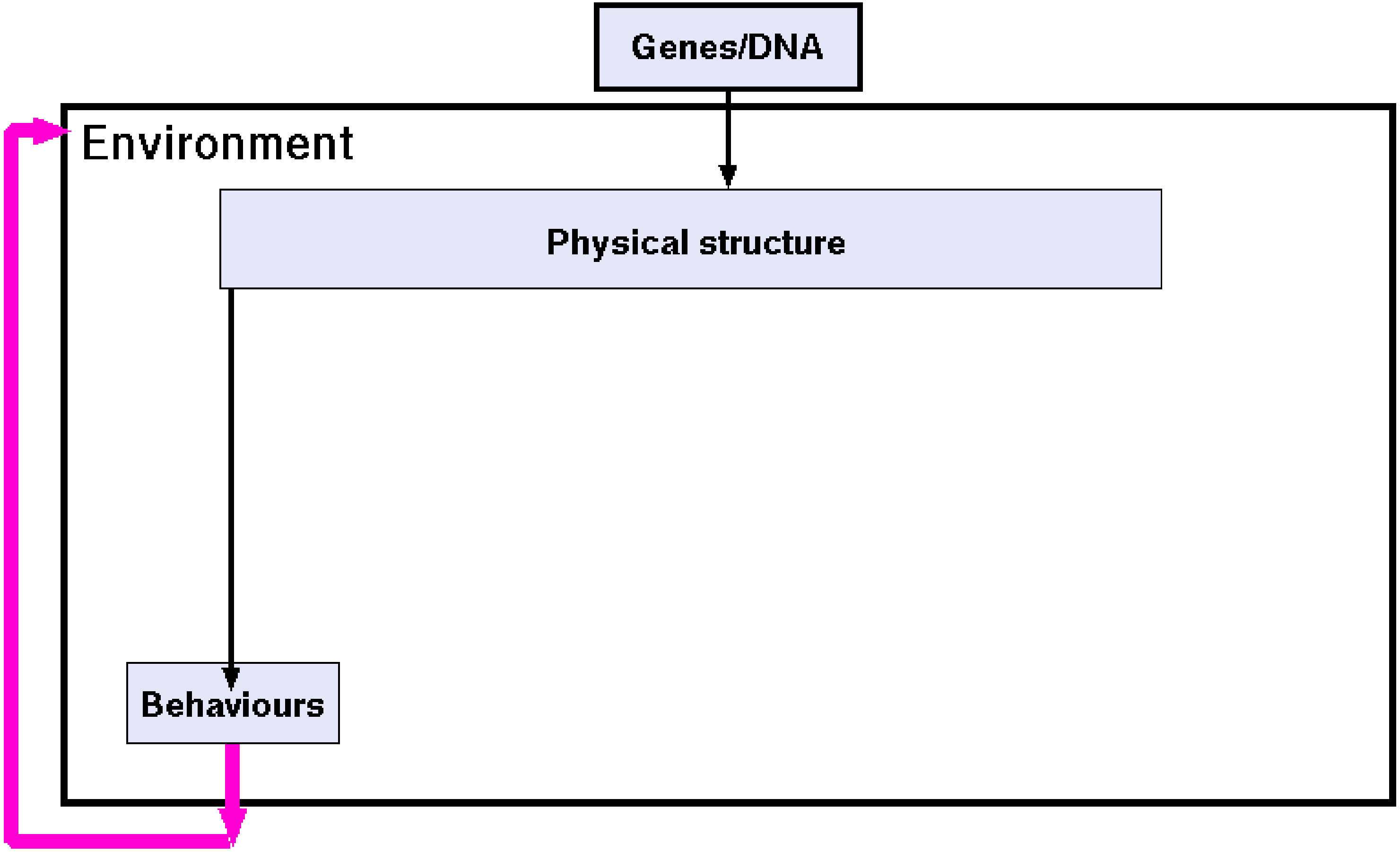
Evaluation in natural evolution keeps changing, as environments change.

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.

(It’s a satisficer, not an optimizer – H.A. Simon)

Individual developmental trajectories I

Routes from genome to behaviour : the direct model.

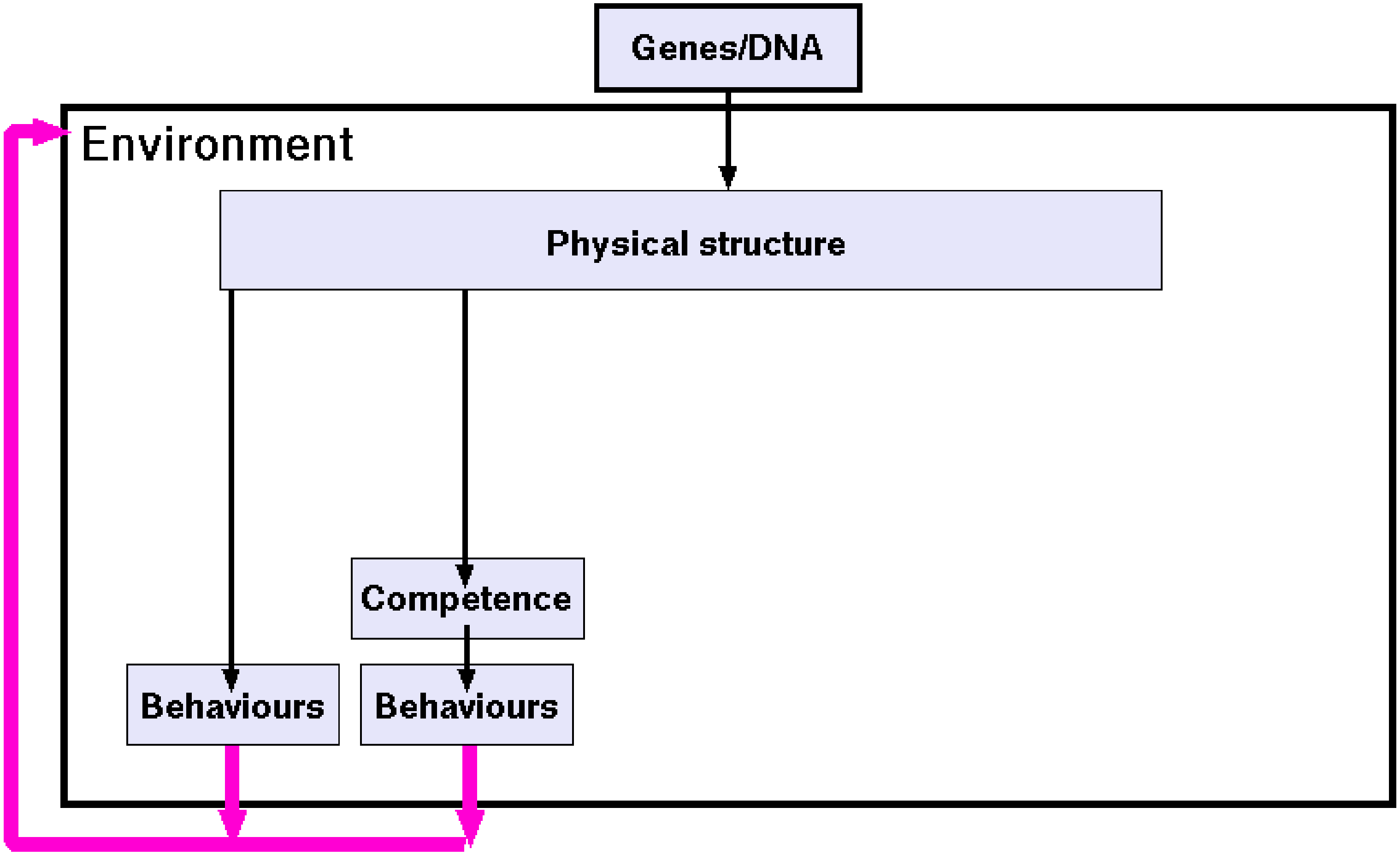


The vast majority of organisms (including micro-organisms) are like this.

Many don’t live long enough to learn much – they have to make do with innate reflexes. Other organisms have more “inside the box”.

Individual developmental trajectories II

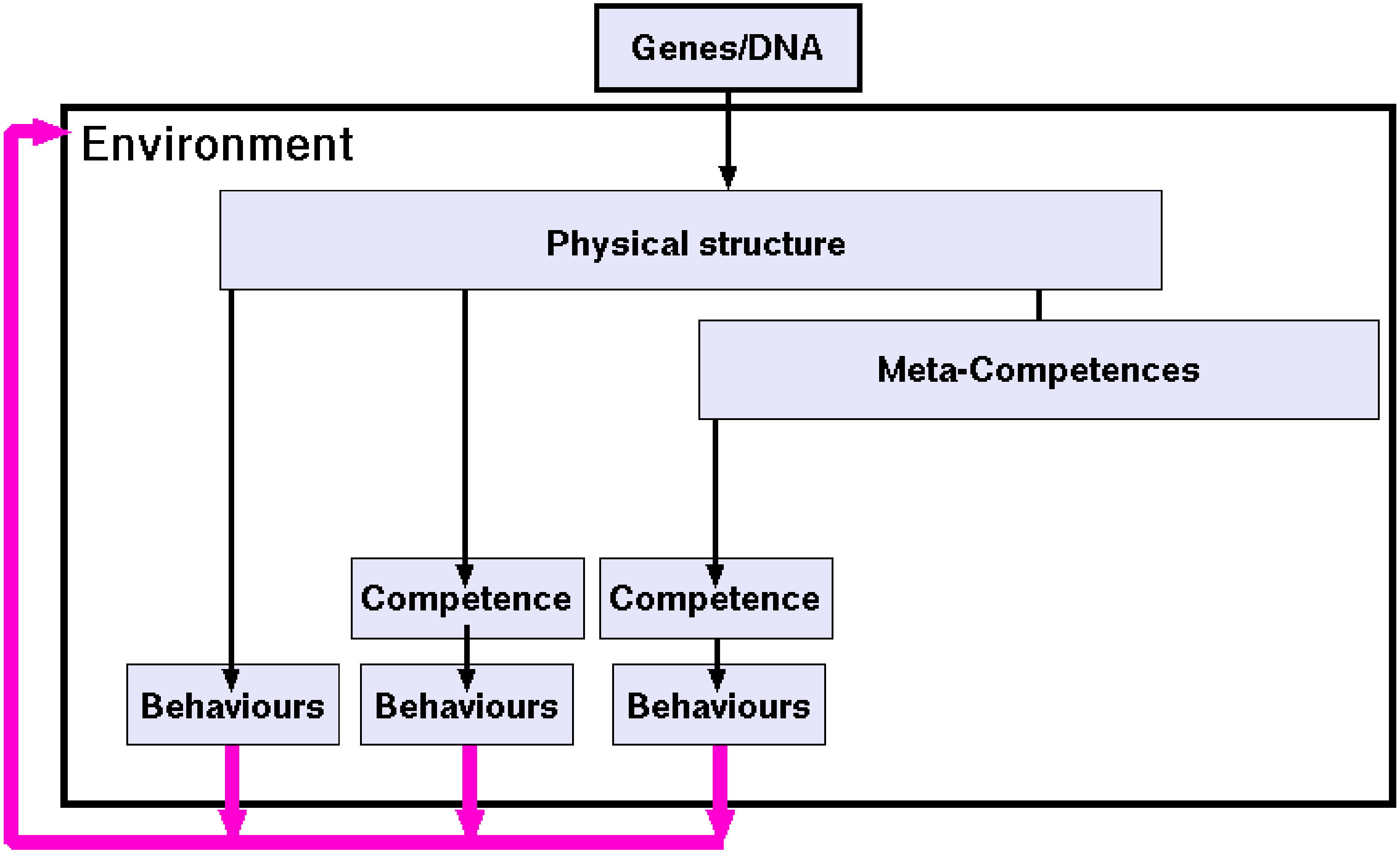
Routes from genome to behaviour : the two-stage model.



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.

Individual developmental trajectories III

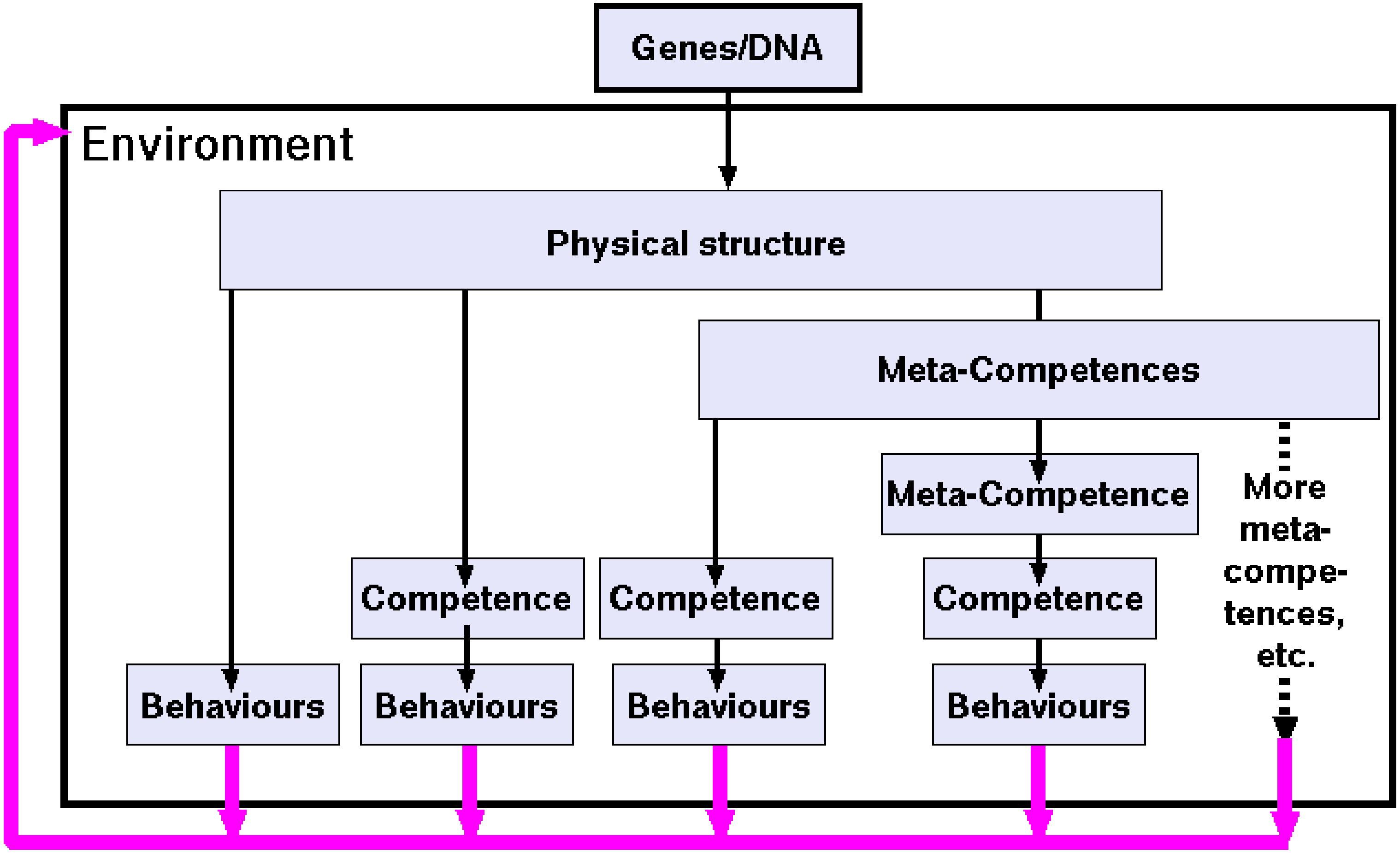
Routes from genome to behaviour : stages added by learning.



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

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.

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.

Are there I-P mechanisms that “naturally” support this process?

(A sort of “reverse Baldwin” effect.)

This is also a recurring feature of mathematical discovery.

Compare Annette Karmiloff-Smith on “Representational redescription”.

Warning: diagrams don’t constitute a theory

Those diagrams are not to be taken as expressing 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) during development and learning.

It looks as if 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.

Non-trivial, non-definitional ontological extension

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

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

“Non-trivial” ontological extensions introduce terms that are 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.

Concepts used in a robot doing SLAM.

Simultaneous Localisation and Mapping

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

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

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

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

Ontologies, evolved, or just learnt?

The first exosomatic ontologies were probably produced by evolution,

E.g. organisms that construct primitive maps recording spatial occupancy, 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.

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

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

See also:
The Computer Revolution in Philosophy, 1978. Now online, revised, especially chapter 2.
`http://www.cs.bham.ac.uk/research/projects/cogaff/crp/`

Ontologies required by living things keep growing

Sometimes evolution

(aided by individual development, learning, culture, ...)

produces new things whose description requires non-trivial ontological extension.

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

concepts of kinds of structural relationship

- topological (topological temporal relations, as well as spatial)
- metrical (Metrics can be used for topological temporal relations, as well as spatial)
- causal
- functional

Concepts of different kinds of agency

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

Compare developments in computer systems engineering.

Placeholder: Mechanisms, ontologies & competences using time

A very rich collection of competences relating to time develops in humans.

We need to understand how far those competences, and related competences, extend to other species.

This will include forms of representation, collections of mechanisms, algorithms required, as well as tools referring in more and more indirect and diverse ways.

See also
`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/evolution-info-transitions.html#time`

Why is evolutionary history relevant?

Some biologists adopting an evolutionary approach focus on questions of the form:

> How does the collection of behaviours (or behaviour generating competences) in members of a species maximize the genetic fitness of the species.
>
> How does an individual's behaviour maximize that individual's genetic fitness?

This project is not concerned with questions of that sort!

Instead I am concerned with questions of the form:

- What exactly are organisms doing when they do X?
- What are the problems that have to be overcome in doing X?
- What information-processing capabilities and mechanisms are used?
- How did those capabilities and mechanisms come to exist,
  - in the species
  - in current individuals
- Can we learn more about how they work by understanding older prior mechanisms on which they were built, or how older mechanism had to be modified to enable the current functions to be performed?

NOTE:

I see no reason to believe that evolution optimizes or maximizes anything. (Simon, 1967b, 1969)

Instead, the enormous diversity of products of evolution indicates that evolution explores possibilities.

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`

Additional materials in the Appendix

An appendix to these slides will contain additional related materials, also accessible at

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk107`

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.

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`

Gerd 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://home.planet.nl/~gkorthof/`

Top level: ”Towards The Third Evolutionary Synthesis”

`http://home.planet.nl/~gkorthof/korthof.htm`

Introduction to the Evolution literature

`http://home.planet.nl/~gkorthof/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

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

Beer, R. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, *4*(3), 91–99. ((http://vorlon.case.edu/˜beer/Papers/TICS.pdf))

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

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.

Davies, P. (1999). *The Fifth Miracle: The Search for the Origin and Meaning of Life*. Simon & Schuster.

Dawkins, R. (1076). *The Selfish Gene*. OUP.

Dawkins, R. (1982). *The extended phenotype: The long reach of the gene*. Oxford, New York: Oxford University Press.

Dawkins, R. (1988). The Evolution of Evolvability. In C. Langton (Ed.), *Artificial Life: Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems* (pp. 201–220). Reading, MA: Addison-Wesley.

Farnsworth, K., Lyashevska, O., & Fung, T. (2012). Functional Complexity: The source of value in biodiversity. *Ecological Complexity*, *11*, 46–52. Available from http://dx.doi.org/10.1016/j.ecocom.2012.02.001

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.

Gelder, T. van. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*(5), 615–628.

Gibson, J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

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.

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`

Marr, D. (1982). *Vision*. San Francisco: W.H.Freeman.

Maturana, H. R., & Varela, F. J. (1992). *The Tree of Knowledge: the Biological Roots of Human Understanding (Rev. ed.)*. Boston, MA: Shambala Publications, Inc.

Maynard Smith, J., & Szathmáry, E. (1995). *The Major Transitions in Evolution*. Oxford, England:: Oxford University Press.

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.

Nahle, N. (2004, November). *Synopsis: Compilation of New Data from Physics and Cosmology and its Application for Supporting the Theory of Abiogenesis.* Monterrey, N. L., Mexico;. Available from `http://www.biocab.org/Abiogenesis_Synopsis.html` (Conference on February 27, 2004)

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.

Numerico, T. (2010). The New Machine: from Logic to Organization. *The Rutherford Journal*, *3*. Available from `http://www.rutherfordjournal.org/article030102.html`

Searle, J. (1980). Minds brains and programs. *The Behavioral and Brain Sciences*, *3*(3). ((With commentaries and reply by Searle))

Searle, J. (1984). *Minds brains and science*. BBC Publications. ((The Reith lectures))

Shannon, C. (1948, July and October). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423 and 623–656. Available from `http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html`

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. (1969, Jan). How to derive “better” from “is”. *American Phil. Quarterly*, *6*, 43–52. Available from `http://www.cs.bham.ac.uk/research/cogaff/papers.html#1969-02`

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. (1986a). Did searle attack strong strong or weak strong AI. In A. Cohn & J. Thomas (Eds.), *Artificial intelligence and its applications.* John Wiley and Sons. Available from http://www.cs.bham.ac.uk/research/projects/cogaff/00-02.html#70

Sloman, A. (1986b). *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. (1993). The mind as a control system. In C. Hookway & D. Peterson (Eds.), *Philosophy and the cognitive sciences* (pp. 69–110). Cambridge, UK: Cambridge University Press. Available from http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#18

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

Sloman, A. (2008a). Architectural and representational requirements for seeing processes, proto-affordances and affordances. In A. G. Cohn, D. C. Hogg, R. Möller, & B. Neumann (Eds.), *Logic and probability for scene interpretation.* Dagstuhl, Germany: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany. Available from http://drops.dagstuhl.de/opus/volltexte/2008/1656

Sloman, A. (2008b). *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. (2008c). *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, Sep). *Virtual Machines and the Metaphysics of Science.* Nottingham. (http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mos09, PDF presentation for AHRC Metaphysics of Science Conference)

Sloman, A. (2009c). What Cognitive Scientists Need to Know about Virtual Machines. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1210–1215). Austin, TX: Cognitive Science Society. Available from http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#901

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. (2010c). Phenomenal and Access Consciousness and the "Hard" Problem: A View from the Designer Stance. *Int. J. Of Machine Consciousness*, *2*(1), 117–169. Available from `http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#906`

Sloman, A. (2010d). *Supervenience and Causation in Virtual Machinery.* University of Birmingham, School of Computer Science. Available from `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk86` (Online tutorial presentation)

Sloman, A. (2010e). *Talk 85: Daniel Dennett on Virtual Machines.* University of Birmingham, School of Computer Science. Available from `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk85` (Online tutorial presentation)

Sloman, A. (2010f). *Using virtual machinery to bridge the "explanatory gap"; Or: Helping Darwin: How to Think About Evolution of Consciousness; Or: How could evolution (or anything else) get ghosts into machines?* University of Birmingham, School of Computer Science. Available from `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk84` (Invited keynote talk at SAB2010)

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. (2013 In Pressa). Paper 1: Virtual machinery and evolution of mind (part 1). In S. B. Cooper & J. van Leeuwen (Eds.), *Alan Turing - His Work and Impact* (pp. ???–???). Amsterdam: Elsevier. Available from `http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1106a`

Sloman, A. (2013 In Pressb). Paper 2: Virtual machinery and evolution of mind (part 2). In S. B. Cooper & J. van Leeuwen (Eds.), *Alan Turing - His Work and Impact* (pp. ???–???). Amsterdam: Elsevier. Available from `http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1106b`

Sloman, A., & Chappell, J. (2005). The Altricial-Precocial Spectrum for Robots. In *Proceedings IJCAI'05* (pp. 1187–1192). Edinburgh: IJCAI. (http://www.cs.bham.ac.uk/research/cogaff/05.html#200502)

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., Chappell, J., & CoSy PlayMate team the. (2006, January). *Orthogonal Recombinable Competences Acquired by Altricial Species (Blankets, string, and plywood)* (Research Note No. COSY-DP-0601). Birmingham, UK: School of Computer Science, University of Birmingham. Available from `http://www.cs.bham.ac.uk/research/projects/cosy/papers#dp0601`

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.

Sussman, G. (1975). *A computational model of skill acquisition*. San Francisco, CA: American Elsevier. Available from `http://dspace.mit.edu/handle/1721.1/6894`

Swinton, J. (2004). Watching The Daisies Grow: Turing and Fibonacci Phyllotaxis. In C. Teuscher (Ed.), *Alan Turing: Life and Legacy of a Great Thinker, 2nd Ed* (pp. 477–498). Berlin: Springer.

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proc. London Math. Soc.*, *42*(2), 230–265. Available from `http://www.thocp.net/biographies/papers/turing_oncomputablenumbers_1936.pdf`

Turing, A. M. (1952). The Chemical Basis Of Morphogenesis. *Phil. Trans. R. Soc. London B 237*, *237*, 37–72.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*, 297–333. (5)

Wiener, N. (1961). *Cybernetics: or control and communication in the animal and the machine*. Cambridge, MA: The MIT Press. (2nd edition)

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`