

Orthogonal Recombinable Competences Acquired by Altricial Species (Blankets, string, and plywood)

Aaron Sloman

Originally begun in the context of the CoSy project
with help from Jackie Chappell and many others.

CONJECTURE:

Alongside the innate physical sucking reflex for obtaining milk to be digested, decomposed and used all over the body for growth, repair, and energy, there is a genetically determined information-sucking reflex, which seeks out, sucks in, and decomposes information, which is later recombined in many ways, growing the information-processing architecture and many diverse recombinaible competences.

JUMP TO CONTENTS LIST

NEWS/NOTES: 20 May 2014; Updated 18 Mar 2015, 24 Jun 2015

- More history at the end of this document.
- The word "orthogonal" here is used fairly loosely and may be a little misleading in suggesting complete independence between different competences that can be creatively combined to solve problems or make discoveries about how things work.
- The first draft of this paper was written in 2006 as part of the EU-Funded COSY project: www.cognitivesystems.org. Since then there have been several new papers and presentations related to the discussion here. References to some of them are inserted below where relevant, but the cross referencing is far from complete.
- The Turing-inspired Meta-Morphogenesis project started in late 2011, includes many examples related to this topic. An overview is here:
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html>
- Relevant materials include discussions of reasoning in Euclidean geometry, and unsolved problems in machine vision (including uses of visual reasoning in mathematics):
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html>
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-sum.html>
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/trisect.html>
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/shirt.html>
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/knots/>
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vision>
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/chewing-test.html>
The Chewing test for intelligence.
- Much of the discussion here is closely related to this KR-1996 Conference Paper:
<http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15> on "Actual Possibilities".

NEWS: 26 Jul 2012

Links to some related work, done since this web page was created in 2006.

- The polyflap domain provides a space in which to develop and explore a collection of related competences.
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/polyflaps>

It is an artificial domain designed to give human adults and future robots some of the experiences of a toddler learning about a new domain, including discovering "[toddler-theorems](#)".
- The idea of a "toddler theorem" arose out of consideration of requirements for a robot that can learn about its environment empirically, then begin to reorganise what it has learnt into a generative system in which theorems can be derived. For more on toddler theorems see:
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html>
- This work is now part of the Meta-Morphogenesis project:
<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html>
- Some of the ideas, about how knowledge acquired empirically can be re-organised into a non-empirical generative/deductive form are related to the ideas about representational redescription in Annette Karmiloff-Smith's 1992 book [Beyond Modularity](#) (reviewed [here](#)).

NEWS: 5 May 2006

In response to a question from Alain Berthoz, suggested some experiments to probe the ability to handle different kinds of visual representations in parallel [here](#).

NEWS: 2 May 2006

Added a separate discussion note on the limitations of learning about [sensorimotor contingencies](#) (contrasted with learning about 'objective' or 'external' condition/consequence contingencies).

NEWS: 19 Apr 2006

[Dean Petters](#) drew my attention to this book

Philippe Rochat
The Infant's World
Harvard University Press
Cambridge, MA 2001

It is extremely well written, and provides much evidence related to the topics discussed on this web site, including evidence that even a two month old child has the ability to make use of the equivalence of grasping with a hand and grasping with the mouth. (page 53). Compare [the discussion of grasping, below](#).

Here are two extracts from the book added here on 27 Apr 2006

What we perceive visually as mature individuals is not light stimulation per se. In addition to being sensitive to light we are objective perceivers. We perceive layouts: surfaces made of objects and things cluttering the environment. These things are sometimes static, sometimes dynamic. (Page 85)

(I suspect 'surfaces made of objects' could have been 'objects made of surfaces' though of course objects are more than surfaces -- they have insides too!)

and

... despite their poor visual acuity, contrast sensitivity, and chromatic discrimination abilities, infants are born perceiving and discriminating an "objectified" world: a world made of rich layouts and surfaces, furnished with physical objects and events, that can be differentiated and specified. from birth infants perceive --- they do not merely sense and react. (Page 88)

People who find that view surprising should reflect on the fact that newborn chicks can follow their mothers and peck for food, and deer can run with the herd soon after birth. (For more on that see [this paper on the altricial-precocial spectrum](#), and the papers co-authored with Jackie Chappell mentioned [below](#).)

Compare also the quotes from McCarthy below.

CONTENTS

- Influences
 - [John McCarthy on The Well-Designed Child](#)
 - [The Naive Physics Project](#)
 - [The work of Jean Piaget](#)
- Introduction: Competences of a five-year old
- Examples of orthogonal competences concerned with physical situations
- Example: properties of blankets, plywood, and string
 - [Maria Petrou's Ironing Robot](#)
- Example: different levels of abstraction - grasping
 - NOTES
 - [Mirror neurons or abstraction neurons?](#)
- Example: Research on lemurs
- More examples: Lids, parallelograms, apes, rubber-bands
- Not all dimensions are equal
- Questions
 - [Q.1. What kinds of learning mechanisms can extract all the different specific kinds of knowledge \(the orthogonal competences, or basis functions\) from the totality of what is perceived over a period of time?](#)
 - [Q.2. How is the information represented and what mechanisms allow the different kinds of information to be combined in novel combinations?](#)
 - [Q.3. Are there any known neural mechanisms and learning processes that can achieve the kind of decomposition into orthogonal competences...](#)
 - [Q.4. Are there any known neural mechanisms that can achieve recombination](#)
 - [Q.5. What exactly *are* the constraints on 'precocial' species ...](#)
 - [Q.6. How are the consequences of novel combinations derived? How is the causal knowledge represented and used?](#)
 - [Q.7. Can we use an analysis of logical/structural dependencies between competences to determine possible developmental pathways?](#)
 - [Q.8. What are the clinical implications of these ideas?](#)
 - [Q.9. What are the implications for ethology?](#)
 - [Q.10. Could the ability to acquire orthogonal competences and recombine them have been a precursor to the evolution of human language?](#)
 - [Q.11. Is all this talk about orthogonal recombinable competences just a trivial re-formulation of what AI theorists have been doing for decades?](#)
 - [Q.12. Can these ideas be extended to include observation and interaction with other agents with percepts, beliefs, preferences, desires, hopes, ...](#)
 - [Q.13. How does all this relate to the theory that perception provides information about affordances?](#)
- Architectures
- NOTE: How all this started
- Origins of this document

Originally CoSy Discussion paper: COSY-DP-0601

- [Thanks and links](#)
-

Influences

John McCarthy

After writing most of what is below I suddenly remembered reading an important unpublished paper by John McCarthy.

John McCarthy, 'The Well Designed Child' (1999)

<http://www-formal.stanford.edu/jmc/child1.html>

I'll quote a few portions the paper that I think are specially relevant:

Evolution solved a different problem than that of starting a baby with no a priori assumptions.

.....

Instead of building babies as Cartesian philosophers taking nothing but their sensations for granted, evolution produced babies with innate prejudices that correspond to facts about the world and babies' positions in it. Learning starts from these prejudices. What is the world like, and what are these instinctive prejudices?

.....

We ask what the world is like at the level at which people and robots interact with it. Particularly important is what we call the common sense informatic situation. It relates the partial information about the world that can be obtained and the kinds of results that can be achieved in the world with these actions.

.....

Footnote 4:

We emphasize the effect of the actions on the world and not the new sensations that result from the action.

.....

Section 7 ...

Ever since the 1950s, people have suggested that the easy way to achieve artificial intelligence is to build an artificial baby and have it learn from experience. Actual attempts to do this have always failed, and I think this is because they were based on the Cartesian baby model.

.....

It is the essence of our approach that appearance and reality are quite distinct and the child is designed to discover reality via appearance.

.....

Suppose something appears and disappears. There are two kinds of mental models a person or robot can have of the phenomenon--flow models and conserved quantity models. Flow models are more generally applicable and apparently are psychologically more primitive. When there is no way of quantifying the substance, as with the water flowing from the tap, the notion of conservation of water is of no help in understanding the phenomenon.

Naive Physics Project

Another precursor to the present paper is the 'Naive Physics' Project of Pat Hayes.

Hayes, PJ 1985. The Second Naive Physics Manifesto. In *Formal Theories of the Commonsense World*, 1-36, eds. JR Hobbs & RC Moore.

The work of Jean Piaget

Piaget had a deep influence on my thinking many years ago, even though I did not believe most of his specific theories (he lacked the conceptual tools required for the task -- as did everyone else at the time). But he looked in many of the right places for problems to be explained. For a taster, see [this online excerpt \(the last chapter\)](#) from *The Construction of Reality in the Child* translated by Margaret Cook, 1955, Routledge and Kegan Paul.

There have been many other influences, some of them listed in [the presentation on vision as process simulation](#).

Introduction: competences of a five-year old

By the time a typical child is about five years old, there is much detailed knowledge and skill concerned with many aspects of the environment (some explicit, some implicit, and some a mixture). This includes competences of several distinct kinds, which are orthogonal in the sense that

- those competences can be used in various combinations in perceiving, thinking about and acting on the world,

(e.g. watching someone folding a sheet can use perceptual sub-processes concerned simultaneously with seeing different kinds of physical materials with different properties, different kinds of surfaces and surface structures, different kinds of objects, different kinds of processes and sub-processes, different sorts of causal relations between events and processes and the role of another intelligent agent in the environment)

- the competences required in those combinations can vary (more or less) independently,
- those competences can be combined in novel ways in creatively perceiving, understanding, and acting in the environment, i.e. these competences are used both in perceiving and in producing novelty.

Examples are presented later.

NOTES:

The competences should be described as 'nearly orthogonal' as there are some dependencies that limit variability, which I shall not discuss here.

The phrase "perceiving, understanding, and acting" is used here to gesture towards a variety of types of process that can occur in a child, robot or animal, in which the different competences can be combined.

Perceiving

A type of physical motion is perceived if sensors affected by the motion cause internal structures to be created that represent various aspects of the motion so as to enable information gained to be used for a variety of purposes, including producing and controlling actions. (Not just recognition, labelling or classification, as is assumed in much research on perception).

Understanding

Understanding has many forms and in the simplest cases may just involve perceiving in a coherent way. In more sophisticated cases understanding may include knowing what the causal relationships are between things perceived, being able to predict what will and will not happen in the near or remote future, being able to explain why something does or does occur, and being to reason hypothetically about what would have happened if something had been different. Special cases of this can be labelled perception of positive and negative affordances, as discussed below. Primitive affordance perception involves learning and using only sensory-motor contingencies, where as more sophisticated cases, including perception of 'vicarious affordances' (affordances for others) may require more abstract information to be represented, as discussed below in connection with grasping.

Acting

Acting can include many kinds of physical action, including reaching, touching, grasping, pushing, pulling, throwing, catching, twisting, inserting, extracting, assembling, disassembling, rearranging, etc. In addition there are communicative acts including pointing, talking about, asking about, hearing about,

Learning

All of the above competences may be products of processes of learning and development, or could be innately instilled in a robot or animal as a result of evolution or design by an engineer. Any non-trivial learning will require sophisticated innate meta-level competences, as discussed in the papers listed below.

Everything said here about examples should be regarded as indicating just the tiny tip of a huge, still mostly invisible, iceberg whose study requires enormously patient dissection of varied collections of examples.

Humans are not the only animals that can learn such things, though they excel in the variety and richness of the competences they acquire. This contrasts with animals (the majority of species, I suspect) that can only learn associations involving total sensory arrays, or relatively global sub-patterns within such arrays (e.g. large blob getting bigger fast).

Why sensory-motor contingencies do not provide an adequate representation of knowledge about the environment in a child or domestic robot is discussed in connection with different levels of grasping competence (and the requirement for perceiving 'vicarious affordances'), below.

Development of re-usable orthogonal competences through interaction with the environment requires special innate mechanisms. They are innate initially, i.e. genetically determined, but in humans seem to be extendable, possibly by applying the mechanisms to themselves. (This required evolution of an architectural extension beyond what most animals have which we have labelled 'meta-management' elsewhere.) Some speculations about requirements for these innate mechanisms in altricial species can be found in these three papers co-authored with Jackie Chappell.

<http://www.cs.bham.ac.uk/research/cogaff/05.html#200503>

COSY-TR-0501: Altricial self-organising information-processing systems (AISBQ 2005)

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

COSY-TR-0502: The Altricial-Precocial Spectrum for Robots (IJCAI 2005)

<http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#717>

Natural and artificial meta-configured altricial information-processing systems (IJUC 2007)
(And other things we have co-authored since then.)

The conjecture is that the learning requires use of creative play and exploration to find out how to decompose complex effects, and to find out how previously learnt competences can and cannot be combined in novel ways. It is well known that learning from passive experience does not achieve the same results as active learning. In our framework a partial explanation for this is that active learning includes forming and testing hypotheses. Hypotheses generate 'test-actions' that are capable of revealing mistakes. Active learning enables mistakes to be discovered quickly leading to new or revised hypotheses. Passive learning requires waiting for something to turn up that reveals mistakes: and that may take a very long time. (The implications of this will be obvious to students of Karl Popper.)

This process makes use of 'Architecture-based motivation' that is distinct from reward-based motivation, as explained here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/architecture-based-motivation.html>

in APA Newsletter on Philosophy and Computers, 2009.

What follows is a first (abbreviated) attempt to summarise some of the distinct kinds of competence involved in dealing with the physical environment, some of which are illustrated in more detail in the full slide presentation (and further elaborated in papers related to the Meta-Morphogenesis project.)

Nothing is said here about adding competence at dealing with *animate* entities in the environment. That, and discussion of self-understanding, would require considerable extensions to the points made here.

A feature of the competences listed is that they can be combined in various ways that are creative insofar as they are novel to the individual. Ela Claridge has suggested that a good name for them might be 'basis functions', in the sense explained [in this short tutorial](#), though that tutorial restricts itself to basis functions expressible in certain sorts of mathematical notations, which may not be general enough to model animal spatial reasoning competences.

Creativity and fluency

Although 'creativity' can be used as a label for the important ability to combine competences in new ways in order to perform new actions or perceive new complex structures and processes, or cope with new situations, it should also be remembered that creatively developed competences can, through much practice, be 'compiled' into various sorts of habitual, or routine, specialised skills, perhaps represented in the same ways as the genetically determined, relatively inflexible, specialised skills of precocial species depending on associations between global patterns. Such skills allow fast fluent performance, but with reduced flexibility.

Self-awareness

I should also emphasise that I am *not* claiming that if a child has a competence or goes through a process of acquiring and using new competences, then it is necessarily *aware*

- that it has the competences that it has
- that it is acquiring collections of competences,
- that the competences are (nearly) orthogonal and recombinable,
- that it is using a particular competence when it uses it.

Such self-awareness may develop gradually, and requires an information processing architecture that does not seem to be available at birth in humans, but develops piecemeal during the first decade of life (or longer).

Some kinds of self-awareness require that the child can notice, and make inferences from, aspects of its own behaviour observed during performance (e.g. noticing that a small change can have a big effect when pushing a block near the edge of a surface). Other kinds of self-awareness require that its information-processing architecture has a component that can *observe* what is going on in various parts of the system while those things are going on (what I have elsewhere referred to as 'meta-management', following Luc Beaudoin -- also one of the kinds of 'reflection' discussed by Minsky in *The Emotion Machine*).

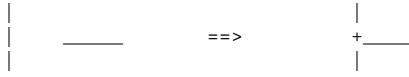
Having various orthogonal competences affects how the child experiences the world (e.g. what it can notice), but that does not mean that the child experiences the underlying causes of how it experiences the world.

Causation

Human children, like many other animals have some innate predispositions to act as if they know what causes what and also the ability to learn things about what causes what. But that assumes that 'causes' is a well defined concept, whereas there is a long and deep history of discussion and controversy about what it is and how it works. To a first approximation there are two views of causation

- **Humean:** 'A caused B' means that A and B are instances of classes for which there are reliable (but probabilistic) correlations, allowing occurrences of the type of A to be used to predict occurrences of the type of B (subject to certain conditions holding). The most sophisticated versions of this deal in networks of conditional probabilities in 'Bayesian nets', as expounded, for instance, in the work of [Judea Pearl](#).
- **Kantian:** 'A caused B' means that there is something in the structure of the bit of the world involving A and B which meant that the occurrence of A *necessitated* the occurrence of B in something like the same way as adding exactly three new objects to a collection of exactly five objects produces a collection of exactly eight objects (if the objects are more like stones than like drops of water), or moving a vertical line towards a horizontal line until the gap

between them is eliminated necessitates production of an intersection as illustrated in this diagram, where '==>' represents a temporal transition



All animals capable of any sort of learning or adaptation that involves generalisation from examples are capable of acquiring information of the first kind (Humean causation). Members of a relatively small subset of species, including humans, squirrels, elephants, some nest-building birds, cetaceans, and octopuses, seem to be capable of grasping and making use of Kantian causation.

I've tried to elaborate on this in [Two views of child as scientist: Humean and Kantian \(PDF, COSY-PR-0506\)](#), and later presentations with Jackie Chappell on [Kantian and Humean causal competences in animals and machines](#).

Some of the examples below are Humean, some Kantian. Distinguishing them is left as an exercise for the reader.

Examples of orthogonal competences concerned with physical situations

A typical child about five years old has much detailed knowledge of several distinct kinds, which can be combined in different ways in perceiving, understanding and planning actions in the environment:

- Many kinds of physical stuff with different physical properties (e.g. water, ice, snow, sand, mud, wood, string, rope, paper, metal, stone, plastic, human skin, cotton wool, hair, butter, treacle, plastic film, aluminium foil, various kinds of food, wind, breath, fire and many more -- should we include shadows here?)

The kind of stuff of which something is made will affect a cluster of properties of various kinds including whether the object is rigid or not, elastic, compressable, easily broken or not, foldable, chewable, scrunchable, etc., and which actions produce reversible results.

NB

Thus learning about kinds of stuff involves both learning about dimensions in which kinds of stuff can differ (hardness, rigidity, flexibility, density, texture, elasticity, etc) and also about particular kinds of stuff which have specific combinations of features in these different dimensions. (The ontology of kinds of stuff, is more complex than this implies: e.g. there are different kinds of flexibility, illustrated by the differences between cloth and paper.)

The contents of this and the other lists will depend heavily on the environment in which the child develops. Some of the mechanisms for acquiring this knowledge are innate, and perhaps some general features of the structure of the knowledge. But the variety of kinds of stuff encountered will be heavily dependent on epoch, geographical location, and culture. I did not encounter snow till I was about 21 (though I had learnt a lot about it by report before then). Moreover, the learning mechanisms themselves may develop and be accommodated to the requirements of the child's life, which change with age, and vary from one culture to another. Explaining how learning mechanisms develop is part of the task of a theory of learning.

- Different kinds of local surface features - flat curved, smooth, rough, sticky, slimy, wet, sharp, textured in different ways, with ridges, furrows, dents, colour, etc. etc.

This decomposes further into yet more (partly) orthogonal sub-spaces. Note that there are dual aspects to all these features: how they can be recognised, and how they affect actions. In addition local features have the ability to be combined in different ways to produce whole objects: and perceived objects can be seen as 'decomposable' into these local features. For some discussion of how this relates to unsolved problems in vision see this [challenge to vision researchers \(PDF\)](#).

Some surface features will be dependent on kind of stuff.

- Different shapes of whole objects, varying in topological and metrical aspects, with both continuous and discrete sub-spaces, at different levels of abstraction,

E.g. there are discrete differences between numbers of holes, between being symmetric or not, being formed from one, two, three, four etc. rectangular blocks stuck together, having a long axis or not, etc. as well as a huge variety of types of continuous variation, e.g. size, angles between components, curvature, weight, degrees and kinds of flexibility, friction at joints....

- Different ways in which new, possibly more complex, wholes can be formed by combining or modifying things (in ways that depend on their shape, material, etc.)

We could include 'negative' combinations, e.g. gouging out, carving, punching a hole, to make a new shape as in sculpture.

Other shape-making transformations include bending, twisting, etc.

- Different sorts of spatial relations between different objects of similar or different material, or different parts of the same object (e.g. containing, touching, being glued to, being hooked round, being a certain distance apart, resting on, being mixed, attracting, repelling, etc. etc.)

There's a particularly important difference between 'rigid' containment (e.g. the streak of metal in a rock, the screw in a plank) and 'fluid' containment, e.g. water, sand or a small ball in a mug, a river flowing in its bed.

Different kinds of spatial regions in which things can be located and processes occur, 2-D, 3-D, large, small, thin, wide, convex, regular, irregular, etc. The scale and shape of an empty space makes a large difference to what can happen within it, including what actions can occur in it, and what their consequences will be. (E.g. how is it that a child can push its head between railings, then be unable to pull it out.)

- Different kinds of force that can be applied to things, e.g. prodding, poking, stroking, squeezing, twisting, pulling, pushing, screwing, patting, moulding....
- Different sorts of process that can occur, either passively observed, or produced through the observer's actions, including moving, rotating, changing shape, entering, coming out of, passing between, pushing, pulling, stretching, swaying, wave-motion, covering, uncovering, putting on (clothing), flocking, swarming, as well as applying forces, and changing the application of forces

Different kinds of causation may be involved in these processes, and different causal connections between parts of a process, e.g. initiating, maintaining, propagating, preventing, stopping, speeding up, slowing down, redirecting, and modulating in other ways.

(E.g. compare the difference between moulding a static lump of potter's clay and a spinning lump.)

Some of the processes may result from the individual's actions, some merely observed.

N.B. More complex things can be observed by an individual than produced by that individual, e.g. a busy street scene, a waterfall, a football match.

There may also be behaviours an animal (e.g. an insect) can produce that it cannot perceive because its perceptual mechanisms lack the required sophistication. I suspect that that is true of more human actions than we realise, especially in infancy. In particular, what we notice and label using verbs (e.g. 'grasping'), is often just a high level abstraction from a complex more detailed reality most of which is not consciously noticed even though it is seen and used in fine-grained control.

These lists are illustrative, not definitive or exhaustive, and do not include social abilities or the ability to represent things that represent. There are many further subdivisions, not specified here.

An example of a mechanism for recombining both competences relating to the physical environment and also competences relating to mental and social entities, allowing for kinds of uncertainty that can be qualitatively expressed, is [John Barnden's ATT-Meta system](#) for reasoning about propositional ATTitudes and METaphors in a unified framework. Much of our use of metaphor can be seen as a special case of creative recombination of competences. (ATT-Meta deals only with phenomena that can be represented in discrete propositional structures.)

Some questions about these competences will be listed below. Major questions include how the information is obtained by the child (assuming that most of it is not innate), how the information is represented in the child, how different items information can be combined in different ways in different contexts, in seeing, planning, reasoning and acting.

Although competence of a child of five years is mentioned above, large subsets of this collection of competences begin to be visible much earlier, e.g. when a child can pick things up and play with them, though there can also be curious gaps, as indicated in this video of a 19 month old child defeated by the problem of using hook and ring to join up a toy train:

[4.2Mbytes mpeg](#)

[11Mbytes mpeg](#)





Even at 11 months a child ostensibly feeding a physical need, seems also to be playing with and exploring a variety of kinds of stuff, kinds of objects, and kinds of actions in the environment, in this video:

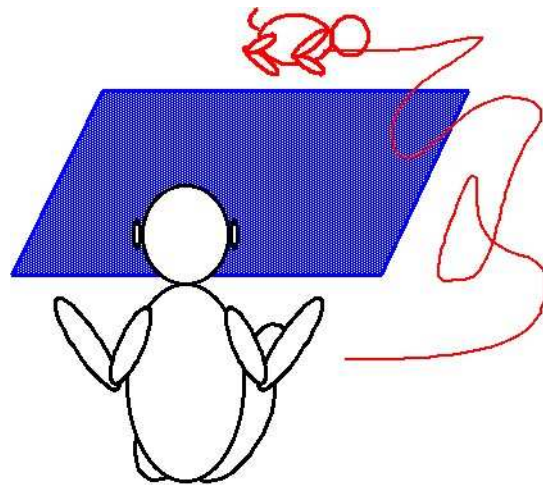
[3.6Mbytes mpeg](#)

[19Mbytes mpeg](#)

What that video illustrates is that the process of learning these orthogonal competences may be driven as much by 'intellectual motivations' (curiosity, a preference for 'interesting' actions and percepts) as by more obviously biological motivations (need for food, drink, warmth, avoidance of pain, physical comfort, mating, etc.). Perhaps the intellectual motivations are more important since, in humans, the biological drives are relevant only a small subset of the time.

Example: properties of blankets, plywood, and string

A child can learn that if a toy is on a blanket the blanket can be pulled to get hold of the toy. But different cases have to be learnt about.



Pulling a blanket:

In the simplest situation, simply stretching forward and grasping the blanket once and pulling the grasped point towards the body will bring the toy within reach. In more complex cases a sequence of grasp and pull motions is needed. (The child may also learn that an efficient way to do this involves using left and right hands alternately, so that each new pull can begin as soon as the previous one ends.)

This can work only because the blanket is made of a kind of stuff that folds so that even after the edge of the blanket has been brought up to the body another portion of the blanket can be grasped and pulled. I.e. the blanket is highly flexible. (String, discussed below, has a similar property, though the grasping actions required are different.)

Another important feature of the blanket's flexibility is that even a portion that is flat and has no visible graspable protrusion can be given a graspable protrusion by 'scrunching'. This depends not only on flexibility but on the textured surface providing friction, so that if a flat hand is pressed on the surface and fingers moved towards the palm, a portion of the blanket will be made to curve to form a graspable protrusion.

Plywood (or stiff cardboard, etc.):

Some objects that are not flexible but provide sufficient friction can also be brought closer. E.g. if the toy is on a piece of plywood it can be brought closer either by grasping the edge of the plywood and pulling towards the body, or by pressing the hand on a part of the flat surface and moving it to the body, using friction to make the plywood move. However, because the plywood is rigid, its length can impose a limit to the movement that the length of the blanket or a piece of string does not. Once the edge of the plywood is brought up to the body, if the toy is still not in reach it cannot be brought closer by bringing the plywood closer. In that case, the only option may be to crawl over the plywood to the toy, or to move round it. A child whose competence is still under development may get itself into an impossible position by bringing the edge of the plywood over its legs towards its belly, making it impossible then to crawl forward or move round the plywood: an ideal opportunity to learn new kinds of actions.

Orthogonality and creativity are demonstrated when a child who has learnt to get hold of a remote toy by using a blanket encounters a similar situation involving a sheet, or towel, of a different shape, size, colour and texture, with a toy resting on it near the far edge. If the appropriate features have been extracted from previous experiments with the blanket the previously acquired competence can immediately be deployed to fetch the toy. Likewise if the toy is on a plastic sheet that is transparent, flexible, and provides enough friction to allow scrunching. Anyone reading the above paragraph on plywood who has never seen the sort of process described there, but who understands the point being made is also creatively recombining competences.

String:

If the toy is beyond a blanket, but a string attached to the toy is close at hand, a very young child whose understanding of causation involving blanket-pulling is still Humean, may try pulling the blanket to get the toy.

At a later stage the child may either still be Humean about causation, and have extended the ontology used in its conditional probabilities, or may have progressed to being Kantian and be able to reason about physical processes, e.g. by simulating the process of moving X when X supports Y. In either case the child does not try pulling the blanket to get the toy lying just beyond it, but uses the string instead.

However the ontology of strings is a bag of worms, even before knots turn up.
(For more on knots see the discussions, photographs, and problems presented in <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/knots/>)

There are many little details to be learnt about the effects of the thickness, stiffness, and spatial arrangement of the string and its relations to other things.

Pulling the end of a string connected to the toy towards you will not move the toy if the string is too long: it will merely straighten part of the string.

The child needs to learn the requirement to produce a straight portion of string between the toy and the place where the string is grasped, so that the fact that string is inextensible can be used to move its far end by moving its near end (by pulling, though not by pushing).

In some cases the child can make the toy move (e.g. further away or up and down) by flapping the string to make a wave. What circumstances permit that?

Try analysing the different strategies that the child may learn in order to cope with a long string, and the perceptual, ontological and representational requirements for learning them. What happens if the string goes round a table leg on the far side of the toy? How does the child's brain (or yours) represent the process perceived when a long string coiled on the floor is gradually straightened by pulling one end?

Maria Petrou's Ironing Robot

Many of the issues addressed here were raised in Maria Petrou's humbling web site for roboticists:

[Specification for an 'ironing robot'](#)

By [Maria Petrou \(Imperial College\)](#)

[The background story](#) (Editorial of the IAPR Newsletter, Volume 19, Number 4, 1997, with cartoons).

Thanks to [Dean Petters](#) for drawing my attention to experiments with infants involving these issues.

Example: Different levels of abstraction - grasping

I've claimed (in [a discussion of vision](#)) that an important step in the evolution of new cognitive competences occurred as a result of the development of hands that can move independently of eyes. This provided the opportunity for animals to manipulate objects without manipulation causing global changes in visual appearance. Global visual changes necessarily accompany manipulation when the only manipulator available is the mouth or beak, which cannot be moved without moving the eyes and thereby generating global changes to the visual field. (Ignoring [mobile lips and tongue](#) for now.)

The evolution of use of hands for manipulation could have led in two directions

1. Development of totally separate competences involving manipulation using mouth or beak and manipulation using hands
2. Development of integrated understanding of the commonalities between the two cases, requiring a form of representation that is more 'objective', i.e. less based on correlations between sensory and motor signals.

The second option allows for more economical representations, sharing information between the two cases, e.g. information about what is common between grasping things with a mouth and grasping them with a hand.

To discover what is common an animal has to find the right level of abstraction. If the actions and their consequences are described in terms of motor signals and sensory input, or in terms of very detailed physical processes, there will not be much in common between the two sorts of grasping, so the representations will have to be done separately and there is no gain in economy. (Compare [the quotations from McCarthy](#), above.)

However, at a higher level of abstraction there are things in common like this (simplifying somewhat):

1. The grasping entity (hand, beak, mouth) has two parts that can move together or apart (e.g. finger and thumb, or upper and lower jaw/beak).
2. During grasping the 'inner' surfaces of the two parts will come into contact with the grasped object.
3. There's (usually) a maximum gap between the inner surfaces. (Maximum grasping gap: MGG)
4. The grasped entity must either be smaller in diameter than the MGG, or must have a protruding part that is narrower than the MGG.
5. If the maximum diameter of the grasped entity is smaller than the MGG the grasping entity can approach the grasped entity from any direction in which there are no obstructions, otherwise the approach must be in a direction that allows a suitable protrusion to come into the gap.

(That's easy for us (adult researchers) to say. How does a toddler or a monkey represent such a generalisation? Can they?)

6. (Cutting a long story short): There's lots more about what can happen after the grasp is achieved (gap closed so that inner finger surfaces are in contact with the grasped object), e.g.
 - effects of moving the grasping object on the motion of the grasped object (e.g. moving grasper from side to side, back and forth, up and down, in a rotary motion, etc.)
 - facts about how the grasp has to be adjusted to cope with weight, size and material of grasped object (e.g. different pressure-to-weight requirements for picking up a rigid block of polystyrene and picking up a block of something much more compressable -- e.g. cotton wool?)
 - what things cannot be grasped even if they are of the right size, e.g. a puff of smoke, a shadow, a column of water coming out of a tap (faucet).

An animal that cannot represent those generalisations because it can represent only sensorimotor contingencies will have to store vastly more information and will not be able to generalise to new cases, e.g. the case where grasping is done using both arms and two palms, or grasping done by others.

Compare the difference between being able to represent translations and rotations of a moving rigid 3-D wire-frame cube, and representing the patterns of motion of all the 12 apparently independently moving shadows produced when the rotating cube is projected onto a 2-D screen.

These ideas are elaborated in two files elaborating on the contrast between sensorimotor contingencies and 'objective' or 'environmental' condition-consequence contingencies, [here](#) and [here \(more elaborate\)](#).

Implications for clinical developmental psychology

Certain kinds of genetic brain malfunction might interfere with the process of learning what is common between the different cases. E.g. if the abstractions listed here cannot be achieved, then the learning about different cases (grasping with mouth and with hand) will have to be duplicated, and inferences made in one case will not be transferred to the other. (One aspect of Autism?)

This paper

Jody Culham, '[Human Brain Imaging Reveals a Parietal Area Specialized for Grasping](#)' Book Chapter. in press, to appear in: *Attention and Performance XX: Functional Brain Imaging of Visual Cognition* (2004), N. Kanwisher & J. Duncan (Eds.) Oxford: Oxford University Press

reports finding a brain area specialised for grasping, but as far as I can tell they have not investigated whether there is anything specialised for grasping whether done with mouth, or left hand or right hand, or both hands, etc. The theory developed here implies that there is (at least in normal humans above a certain age), but it need not be detectable using current brain imaging techniques.

I think a child learns very many such things in the first few years, and a domestic robot will probably also have to. Of course the child cannot *state* any of this at first, so that raises questions about how the information is stored, accessed used, extended, etc., and also how the competence is used later on when the child learns to talk about such things.

NOTES:

1. I am not claiming that all species that use both manipulators on limbs and have a mouth that can grasp things will necessarily develop this abstract understanding of grasping. E.g. there are crabs that use claws to bring food to their mouth, and I have no reason to believe that they understand or make any use of what is common between grasping with claws and biting. Their competence may be restricted to the use of specific sensory-motor contingencies.

2. These ideas about representing what is common to different competences by using a more abstract representation undermine any simple idea that all knowledge of the environment is encoded in terms of sensory-motor contingencies, for relations between sensor and motor signals would be incapable of expressing the powerful, widely re-usable kinds of knowledge we are discussing. For that we need action-consequence contingencies where the actions and consequences are described in a manner that refers to the 3-D spatial and structural relations (and changes therein) that are independent of detailed control and sensing mechanisms. Likewise it needs notions of force and pressure rather than notions of signals to muscles, or muscle tension, etc.

Of course, sensory-motor contingencies remain relevant both in testing hypotheses and in controlling specific actions. But knowledge about what the contingencies should be can be derived from more fundamental and 'objective' representations of what is in the environment. (This is related to what I've called learning about Kantian causation as opposed to Humean causation [here](#) (COSY-PR-0506 - PDF).)

3. The point about discovering and using abstractions, as one of the wide variety of orthogonal competences is relevant to debates about the role of embodiment in cognition. For animals that are incapable of discovering, representing and using the abstractions that allow the same core competence to be applicable in physically varied contexts, including actions done using different body parts, or actions done by different individuals, brain mechanisms will be very closely tied to body structure and specific sensors and effectors. In that case there is deep synergy between the brain mechanisms and the physical structures that sense and act on the world and there is plenty of scope for shifting representational functions between them. (E.g. if extension of a muscle adequately encodes information about the weight of an object there may be no need for that information to be represented centrally.) This is probably true of all microbes and insects (with a few highly specific exceptions such as the representation and transmission of route information in bees). It is probably true of all or most non-vertebrates and may even be true of most vertebrates. In that sense their 'cognition' is highly embodied.

In contrast, for organisms that can store, manipulate and use information that abstracts from the detailed states and processes of particular body parts, even to the extent of being able to think about past, remote or future events that are not currently being produced or observed, the connection between cognition and the body is much less direct. So in that sense humans are much less embodied than insects.

4. Amongst people who join in the (silly) debate about whether to use discrete symbols or other forms of representation it is common to claim that discrete symbols are too 'brittle', that the knowledge they represent cannot cope with the complexities of the real world. What many such people fail to realise is that there is no one right form of representation, and different sorts have different benefits (as I tried to [show in 1971](#); compare Minsky's 'causal diversity' matrix in [here](#)).

In particular, continuously changing forms of representation may be required for fine-grained control of physical processes, especially multi-strand processes where several different metrical and topological and other spatial relations are changing concurrently, e.g. in grasping an object, pulling string so that bits of it straighten out, scrunching a portion of blanket --- whereas discrete, more logical, formalisms (which I called 'Fregean' in 1971) are appropriate for representing competences at higher levels of abstraction.

(It's a pity that so many intelligent people waste effort on factional battles about what is the *right* way to do something instead of expanding knowledge by analysing trade-offs.)

5. I am grateful to Henrik Christensen for drawing my attention to the work of Alain Berthoz [The Brain's sense of movement](#), which provides much relevant information about perception, action and learning to do them, though it does not discuss thinking about what would happen if, or explaining how something might have happened, a competence that seems to grow in young children as they learn about new domains.

6. There is also relevant material in Philippe Rochat's [book, mentioned above](#).

Mirror neurons or abstraction neurons?

The mechanisms involved in abstracting from details of how something is done or even who is doing it must be implemented in brains, and so it is not surprising that there are some neurons implicated. Moreover it is important that there be overlap in the semantics of perception/recognition and the semantics of intention, so it is to be expected that the implementation of that overlap will overlap with the implementation of the overlap due to abstraction.

Conclusion: the so-called discovery of mirror neurons was probably really a discovery of 'abstraction neurons' and had they been properly labelled some spurious research and speculation might have been avoided. It could have led to deeper studies of imitation. See also [Beyond Sensorimotor Contingencies](#).

Example: Research on lemurs

NOTE:

Jackie Chappell, with whom I have been discussing many of these issues, drew my attention to some research at Yale university involving Lemurs

Laurie R. Santos, Neha Mahajan, and Jennifer L. Barnes

How Prosimian Primates Represent Tools: Experiments With Two Lemur Species (*Eulemur fulvus* and *Lemur catta*)
Journal of Comparative Psychology 2005, Vol. 119, No. 4, 394 - 403

The concluding section states:

Our overall pattern of results, which suggests that lemurs solve the cane-pulling task like other tool-using primates, poses a puzzle for the view that differences in primates' natural tool use reflect differences at the level of conceptual ability. Our results, instead, provide support to an alternative account -- namely, that many primates share an ability to reason about the functional properties of different objects, irrespective of whether they exhibit tool use in their natural behavioral repertoire (see also Hauser, 1997; Hauser, Pearson, & Seelig, 2002; Spaulding and Hauser, 2005). These data fit with a growing view that both tool-using and non-tool-using primates share a suite of domain-specific -- possibly innate (e.g., Hauser, Pearson, & Seelig, 2002) -- mechanisms for reasoning about physical objects and that performance in simple pulling tasks may tap into these abilities.

The ideas about orthogonal competences presented here provide a framework for decomposing collections of competences that allows far more fine-grained questions to be posed about the abilities of children at different stages of development, or animals of different species. From this viewpoint notions like 'tool use', 'understanding functional properties' 'simple pulling tasks' turn out to be very coarse-grained, since each can be further subdivided to identify different collections of competences that can exist.

A loose analogy:

Compare the role of the periodic table of the elements and its relation to a theory of the architecture of matter?

More examples: Lids, parallelograms, apes, rubber-bands

Ideally many more examples should be provided.

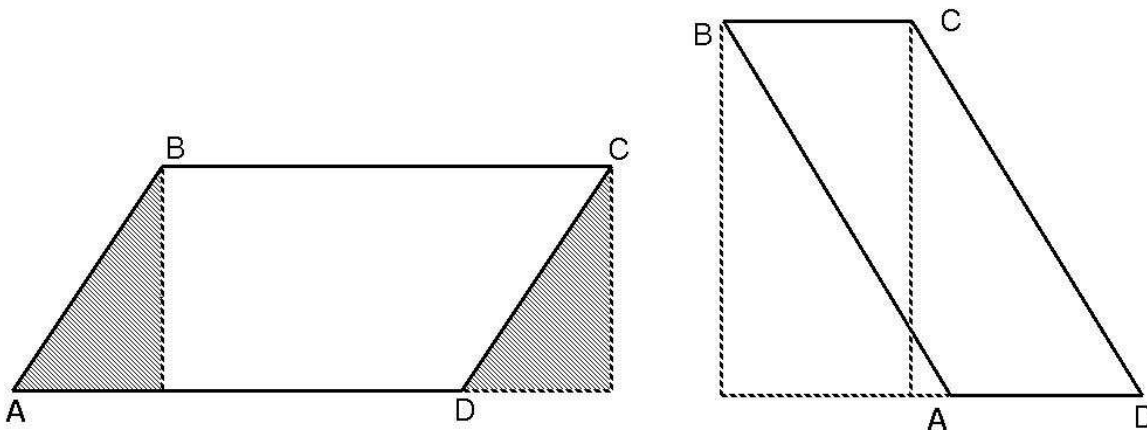
1. A surrogate screwdriver

A small child tries to remove the lid of a large 'restaurant size' coffee tin which contains toys. Normally he uses a screwdriver, by inserting one end under the flange and levering the flange up until the lid is loose. But there is no screwdriver in sight. Suddenly he picks up the circular lid of a smaller coffee tin, inserts an edge under the flange and uses the smaller lid to lever up the edge of the bigger lid.

He has discovered another orthogonality: the functional properties of a screwdriver that suit it to the levering role can be separated from the specific shape of the screwdriver. Exactly how he represented what was common is not clear to me. (This happened many years ago.)

2. Max Wertheimer's tilted parallelogram

Max Wertheimer tells the story in his book *Productive Thinking* of a mathematics teacher showing students how to compute the area of a parallelogram by dropping two perpendiculars and seeing that the two triangles that result are congruent, as in the parallelogram on the left.



After the teacher and the children claimed that the teaching was successful, Wertheimer drew a parallelogram where the shorter side was horizontal (as on the right). Most of the students claimed not to have learnt that. A few realised that the figure could be rotated to produce the same sort of figure as they had previously mastered. Some also noticed that dropping perpendiculars still worked, provided that negative areas were allowed for.

This shows how the very same learning context can produce learning at different levels of abstraction in different learners. Sometimes subtle probes are required to reveal the differences.

3. Kohler's apes

W. Kohler's observations, in *The mentality of apes* seemed to show considerable variation in the ability of chimpanzees to factor out and recombine competences.

4. Rubber bands [added 3 Feb 2006]

This example is inspired by an example in the little book by Jean Sauvy and Simonne Sauvy *The Child's Discovery of Space: From Hopscotch to Mazes an Introduction to Intuitive Topology* (1955?). They use the example of making an "H" with a rubber band in connection with teaching children about the topological differences between closed curves and open curves, and other topological differences.

Imagine that you have a rubber band lying flat on a horizontal board with some pins nearby and you wish to use the pins to hold the band in the form of an 'outline' capital "H".

1. How many pins will you need to stick into the board to do this?
2. How many of the pins will be inside the closed curve formed by the rubber band and how many will be outside?
3. When you have the H constructed what are the different shapes that can be produced by removing exactly one pin? Exactly two pins?
4. If you have only one hand available, what sequence of actions could you use to get the rubber band from the initial state of lying loose with a circular shape to forming the H?

I suspect few, if any, children could answer these questions by the age of 5. I suspect many, though perhaps not all, could answer them by the age of 10. What would have to change in between?

Not all dimensions are equal

(PLACEHOLDER for a longer comment.)

In many spaces composed of orthogonal dimensions the dimensions are all similar in kind, e.g. the space of locations represented by three coordinates.

However the roles and structures of different dimensions of reality discovered by a child vary, and that is a major feature of the ontology the child has to develop.

E.g. a simple rotation can switch a horizontal dimension of length to a vertical dimension of length of a 3-D object. But material stuff and shape are not interchangeable, and neither are rigidity and size, or elasticity and weight, etc.

This both complicates and simplifies what has to be learnt. The complication is that different things have to be learnt about different dimensions.

The simplification is that the constraints on different dimensions can help control searches for interpretations of what is seen and searches for actions or plans to achieve a goal.

(Or something like that: that's why specialised expertise gives people power that the generalists lack.)

Questions

Q.1. What kinds of learning mechanisms can extract all the different specific kinds of knowledge (the orthogonal competences, or basis functions) from the totality of what is perceived over a period of time?

(How important is it for the learning process to be driven by active exploration rather than passive observation, and what is it about the mechanisms that make this important?)

Q.2. How is the information represented and what mechanisms allow the different kinds of information to be combined in novel combinations?

E.g. you can reason in some detail about the effects of making a hat out of lead, and putting it on your head, even if you have never encountered such a thing? You can imagine roughly where it will hurt, and what effects it will have on your head movements. Likewise consider making a table out of butter, then doing various things with it (at normal room temperature), including leaning on it, putting a hot iron on it, trying to screw a lamp-holder onto it.

Constraints on appropriate forms of representation for the information discussed here will come from various requirements, e.g.

- the need to acquire the information by observing and experimenting on the environment,
- the structure of the information-space (e.g. is it continuous, discrete, and in how many different ways/dimensions can variation occur? how many different levels of abstraction are useful?),
- the requirement to retrieve previously stored information on the basis of (a) perceived situation, (b) goals to be achieved, (c) partially solved problems,
- the requirement to be able to notice and repair bugs in previously stored information,
- the requirement to use the information in planning, selecting and controlling actions.
(*Searching* for a sequence or other combination of actions may require parallel representations of alternative possibilities, along with backtracking capabilities, etc. Representing two or more alternatives simultaneously in order to be able to compare and evaluate them adds extra constraints on forms of representation and representational mechanisms.)

The nature of actual sensor and motor devices used will also be relevant to how the information should be stored, but not as relevant as the 'embodiment' bandwagon suggests.

NB:

The ability to factor out components of complex situations and to recombine them in different ways for different purposes requires a form of representation with a *syntax*, in the generalised sense that allows information structures to be combined in larger structures in such a way that the information encoded in the larger structure is a function of both the components and the form of composition (and in some cases also the wider context -- which allows ambiguity to be a source of economy). A special case of this is Tarskian compositional semantics, which is relevant to Fregean formalisms (where all complexity comes from function/argument composition). But there are probably many other forms of syntax with their own modes of composition, including what I've called 'analogical' representations and also neural mechanisms, dynamical systems of various sorts, chemical composition (which may turn out to be a special case of one of the others), genome composition and perhaps things that still remain to be discovered.

It is possible that evolution discovered forms of representation and mechanisms for combining representations that have been in use for millions of years for internal processing in many animals, and which have not yet been discovered by neuroscientists, computer scientists, psychologists, or mathematicians.

Q.3. Are there any known neural mechanisms and learning processes that can achieve the kind of decomposition into orthogonal competences (or basis functions) described here?

Q.4. Are there any known neural mechanisms that can achieve recombination to solve new problems or interpret/understand new perceptual phenomena?

Q.5. What exactly *are* the constraints on 'precocial' species whose competence is almost all genetically determined?

E.g. how much ability to distinguish different kinds of stuff (sand, mud, grass, rock, leaves, snow?) is present in chickens, cows, nest-building insects? Can the set of sub-categories be extended during the lifetime of such individuals? How is their ability to distinguish different kinds of stuff related to perceiving different affordances and performing different actions?

For some insects the performance of mating seems to require an elaborate understanding of a small subset of structures and movements. How is that understanding encoded, and how much variation can it cope with?

Presumably there has to be substantial reconstruction of the 'brain' between larval state and adult state, given the very different body shape, varieties of movements possible, kinds of entities interacted with, etc.

Q.6. How are the consequences of novel combinations derived? How is the causal knowledge represented and used? (See some suggestions.)

For cases where the information combined can be expressed in a logical form, e.g. as a conjunction of conditions

X is round and hard and small

standard mechanisms using rulesets or bayesian nets may be used.

However in situations where what is important includes a collection of structural features which change during the performance of an action it may be that a different form of representation better suited to representing 'multi-strand' processes is needed, E.g. if a long piece of string is lying on the floor, as in the picture, the curves and coils will form a unique configuration that will gradually change as the loose end is pulled. During the process there will be some continuous changes, e.g. a loop getting smaller, the string getting straighter, and some discrete changes, e.g. a loop disappearing, or a limit reached because the string goes round a table-leg, or a portion becomes straight, or the string is fully stretched and cannot get straighter without moving the table.

How many animals can understand such changes? At what stage can a child understand them?

For more on this see [This Paper](#)

COSY-PR-0506: Two views of child as scientist: Humean and Kantian (PDF)

And a much older (1971) [paper about reasoning using analogical representations.](#)

Q.7. Can we use an analysis of logical/structural dependencies between competences to determine possible developmental pathways?

There will obviously not be a single sequence of stages of learning, but there may be constraints that determine a partially ordered network of learning/developmental trajectories. Compare Waddington's notion of an 'Epigenetic landscape'. He presented this as a continuum in his diagrams, but from our viewpoint the child's development includes the acquisition of many *discrete* and *re-usable* information-based competences.

So instead of a developmental 'landscape' we need to think in terms of a collection of partially ordered networks (of trajectories) at different levels of abstraction, including many discontinuities.

Q.8. What are the clinical implications of these ideas?

If all the sorts of competences alluded to here really are orthogonal, then in principle it seems likely that sufficiently local brain damage or malfunction might be able to disable a very specific competence leaving others intact. This idea could lead to empirical investigation of new kinds of 'double dissociations' (cases where competence X is damaged but not Y and cases where competence Y is damaged but not X).

More interestingly we may be able to find cases where *specific* competences are disabled by brain trauma and cases where *generic* competences are disabled, e.g. the ability to learn new competences, or the ability to combine competences to deal with novelty. If such cases were found empirically they would surely include some surprises that would help to refine the theory.

Q.9. What are the implications for ethology?

I suspect these ideas could stimulate a host of empirical investigations into the kinds of competences and the kinds of orthogonality that evolved in different species, including the trade-offs between 'preconfigured' collections of competence determined mostly genetically, and 'meta-configured' competences developed by interacting with the environment under the control of a powerful meta-level learning competence (or collection of competences), as discussed in these papers attempting to extend the biologists distinction between precocial and altricial species:

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/index.php#tr0501>

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/index.php#tr0502>

Q.10. Could the ability to acquire orthogonal competences and recombine them have been a precursor to the evolution of human language?

I've argued in [The primacy of non-communicative language \(1979\)](#) that before external communicative language could develop there had to be internal forms of representation with syntax and semantics. I think the phenomena described here, insofar as they are found in non-human animals and prelinguistic children, strengthen that argument.

Q.11. Is all this talk about orthogonal recombinable competences just a trivial re-formulation of what AI theorists have been doing for decades?

Anyone who has claimed that all the information an intelligent robot (or child?) needs to use can be expressed and manipulated in a logical form will regard it as trivially obvious that

- Logical assertions (particular and general) can to a large extent (i.e. subject to consistency requirements that can be difficult to mechanise) be independently added and removed from a database -- which is a way of implementing orthogonality.
- Given any rich collection of logical assertions, including both generalisations as well as particular propositions, there are in most cases indefinitely many logical consequences that can be derived from them, so that a theorem prover can be seen as a mechanism for implementing novel, creative, combinations of competences.

While correct as far as it goes, this logicist point (as famously presented in papers by John McCarthy since [1959](#) including the important [1969 paper by McCarthy and Hayes](#), does not address the question of where all the information (propositions) in the database come from, and in particular how it can be acquired using physical sensors and effectors such as eyes, ears, touch-sensors, along with fingers, arms, legs, muscles, and mechanisms for control their movements.

It is also open to the challenge that other forms of representation, including analogical representations, neural mechanisms and perhaps forms of representation used in mechanisms not yet thought of by scientists and engineers may be epistemologically and heuristically more powerful than logical representations for an important subset of tasks concerned with the physical environment, as I and many others have argued (e.g. in [the 1971 paper](#) and [elsewhere](#)).

One of the requirements for human learning is having a form of representation of the environment that allows different ontologies to be developed on top of it, including different ways of carving up continuous spaces to meet different requirements, and also different ways of adding abstractions. Such extensions allow percepts, thoughts, hypotheses, etc. to be expressed that were not expressible in the earlier form of representation. This is not the same as adding a new symbol to abbreviate what was previously expressible. A survey of kinds of ontological extensions and their representational requirements is beyond the scope of this paper, but two examples are

- carving up a continuum in a new way
- adding a new collection of primitive symbols not definable in terms of previous ones, e.g. in developing a new theory to explain old phenomena.

I suspect human development includes both sorts of processes. Logic is obviously usable for the second process, but may not always be suitable for the first.

Q.12. Can these ideas be extended to include observation and interaction with other agents with percepts, beliefs, preferences, desires, hopes, intentions, fears, etc?

The answer is not obvious, but I think it is affirmative, though the representational apparatus acquired by a child or animal to cope with the physical environment needs modification to cope with entities that themselves contain representational apparatus (a requirement for the mental states and processes mentioned in the question). Similar modifications are required for an individual to be able to represent its own mental states and processes. Such 'meta-semantic' competence makes possible many kinds of individual and social processes not mentioned here.

This is a topic for discussion on another occasion, though a useful recent paper that is highly relevant can be found here (PDF): [Sharon Wood, 'Representation and purposeful autonomous agents' *Robotics and Autonomous Systems* 51 \(2005\) 217-228](#)

Q.13. How does all this relate to the theory that perception provides information about affordances?

J.J. Gibson introduced the important idea that for an organism the important information about the environment is not what the objective physical properties (shape, size, material, distance, speed of motion, etc.) of objects are but rather what actions the organism can and cannot perform and what the consequences of performing or not performing them are, in respect of the organisms goals or needs. (This includes positive and negative affordances.)

This does not make affordances subjective: rather, affordances are subtle *relations* between portions of the environment and organisms. This is a deep and important point and leads to a much richer view of the functions of vision (and perception in general) than the sort of view put forward by David Marr in his 1982 book *Vision*, and assumed by many machine vision researchers who assume that vision can be done independently of considerations of who or what the vision is for.

One of the corollaries is that insofar as an organism or robot has a complex architecture performing multiple functions in parallel, many of them using visual input (e.g. posture control, route planning, enjoying the view, reading a map etc.) there may be quite different sorts of affordances being processed at the same within different parts of the architecture derived from the same retinal input -- with most of the processing done unconsciously (e.g. posture control).

Affordances and proto-affordances

Detailed analysis of processes that occur when a child or robot manipulates 3-D objects reveals that as part of the process of perceiving affordances the viewer needs to be able in some cases to perceive changes of physical structure or relationships, at different levels of abstraction -- e.g. seeing two surfaces moving together or apart, seeing a long thin object going through a narrow gap, etc. Some organisms (e.g. an insect) may be capable only of percepts that are 'holistic', because they are not able to use orthogonal competences to separate what happens in the environment from how their goals or actions are enabled or hindered. In contrast, as explained above in connection with grasping, some organisms and robots can separate out (a) processes that are occurring in the environment and (b) the relevance of those processes to their own concerns and actions. We could use the term 'proto affordances' for these perceivable fragments of affordances.

One consequence of perceiving such proto affordances independently of how they relate to one's own goals and actions is that it may be possible to see how they relate to goals and actions of another. I call this seeing 'vicarious affordances'. Being able to see vicarious affordances is a requirement for an adult caring for a child and also for predators trying to see how best to catch its prey, or for a predator to see how best to avoid a predator. It can also be used in learning by watching others do things. This is discussed further in the [CoSy report on requirements for representation](#).

For more on the need to generalise Gibson's ideas, and the evolutionary implications of some requirements for human visual capabilities see this online tutorial presentation:

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

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

This research note (work with Jackie Chappell) is also relevant:

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

Talk 52: Evolution of minds and languages.

What evolved first and develops first in children:

Languages for communicating, or languages for thinking (Generalised Languages: GLs)?

Architectures

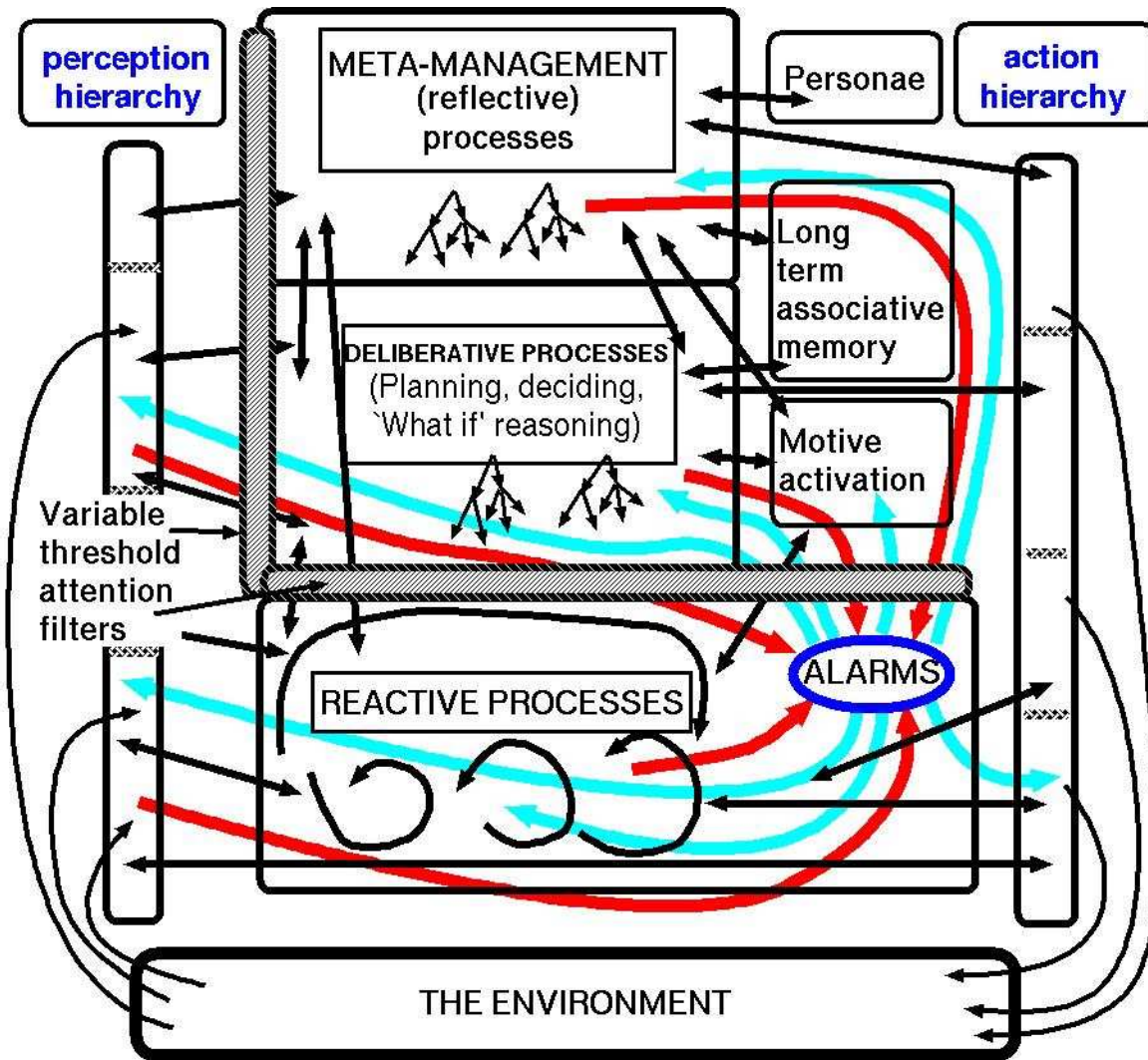
It is a fundamental assumption of our work that a human (or animal, or robot) mind should not be thought of as a unitary system that can be described as, for instance, having one visual process. Rather there is a complex architecture in which different kinds of tasks are being done in parallel with various subsystems *sharing* the same sensors and effectors, sometimes in parallel and sometimes sequentially.

A simplified framework for talking about these different types of components and their functions is the CogAff schema, depicted crudely here, indicating different types of functionality that may or may not be present in different parts of the perceptual, central and action sub-systems (which are not necessarily as distinct as the diagram suggests). Different examples of the schema will have different

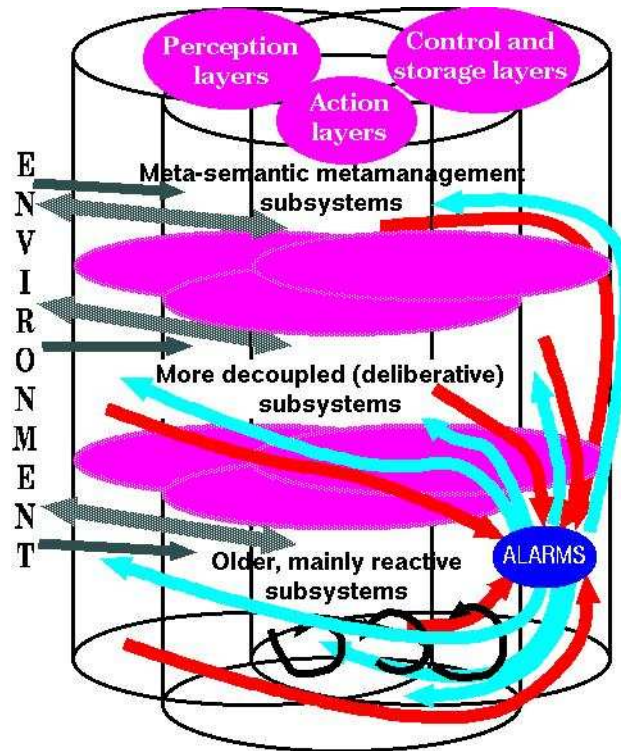
concurrently active sub-mechanisms, with different communication links between the subsystems. The bottom layers are the oldest in evolutionary terms, and insects, for example would not have anything but the bottom layer.

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

A simplified depiction of a theory of the human architecture, which is an instance of the CogAff schema is here



Below is another view of the CogAff schema, allowing for overlaps between sensor, effector and central subsystems and with "alarms" added (Thanks to Dean Petters for this version):



What all this means is partly explained in the presentations on architectures here (e.g. Talk 31) <http://www.cs.bham.ac.uk/research/cogaff/talks/>

A corollary of all this is that there is not a unique answer to the question 'What does so and so see?' since different parts of the architecture (e.g. evolutionarily old and new parts) may see different things at the same time.

NOTE: How all this started

The work presented here grew, in part, out of work in Birmingham on analysis of requirements for representation in the PlayMate scenario in the CoSy robot project, as reported in COSY-TR-0507: DR.2.1 Requirements study for representations, especially the requirements for something like a domestic robot capable of perceiving, understanding and manipulating 3-D objects found around a typical home.

The ideas about 'multi-strand relations' and 'multi-strand processes' discussed there proved particularly fertile and took off rapidly in new directions in October 2005, as reported here and in a presentation on a (possibly) new theory of vision (PDF) viewing vision as being essentially concerned with process-perception, and making use of multi-level process simulations. About six seminars on this topic were presented in different places in the following weeks and the comments and criticisms received provoked further development of the ideas sketched below. Working with colleagues at Birmingham in the CoSy project on the implications of all this for the 'PlayMate' robot scenario led to the idea of the robot requiring a collection of *(nearly) orthogonal re-combinable competences* of different sorts, for which we started to develop a methodology and a web-based tool, as explained here.

This file arose from attempts to answer objections and questions brought up in discussions, and then continued to grow independently. The version in the presentation and this version will continue to change independently.

Origins of this document

Originally CoSy Discussion paper: COSY-DP-0601

(Based on work in progress in the CoSy project in 2006)

Available at <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/orthogonal-competences.html>

Now part of the [Meta-Morphogenesis project](#).

[Other reports, papers and presentations related to the CoSy project](#)

This is also part of the Birmingham CogAff project, summarised here:

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

With a messy, steadily growing, collection of discussion notes here:

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

and presentations (mostly PDF and/or slideshare) here:

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

THANKS AND LINKS

Thanks to Frederic Glanois for help in clarifying some of these ideas.

Some relevant work led by Helene Intraub at the University of Delaware is [here](#).

To be extended

Installed: early 2006

Last updated: 2 May 2006; 26 Jul 2012; 24 Jun 2015(format)