Notes from CogX kick-off

Assembling bits of stuff and bits of process, in a baby robot's world

A Kantian approach to robotics and developmental psychology

Aaron Sloman

School of Computer Science, University of Birmingham

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

These PDF slides will be in my 'talks' directory:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks#babystuff

Assumptions of this presentation

These slides are based on the observation that current machine perceptual abilities and machine manipulative abilities are extremely limited compared with what humans and many other animals can do.

There are mobile robots that are impressive as engineering products,

e.g. BigDog – the Boston dynamics robot

`http://www.bostondynamics.com/content/sec.php?section=BigDog`

and some other mobile robots that are able to keep moving in fairly rough terrain, including in some cases moving up stairs or over very irregular obstacles.

However, they all seem to lack any understanding of what they are doing, or the ability to achieve a specific goal despite changing obstacles, and then adopt another goal.

For more detailed examples of missing capabilities see these web sites

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

`http://www.cs.bham.ac.uk/research/projects/cosy/photos/crane`

As far as I know, none of the existing robots that manipulate objects can perceive what is possible in a situation when it is not happening, and reason about what the result would be if something were to happen.

Neither can they reason about why something is not possible.

I.e. they lack the abilities underlying the perception of positive and negative affordances.

They cannot wonder why an action failed, or what would have happened if..., or notice that their action might have failed if so and so had occurred part way through, etc., or realise that some information was available that they did not notice when they could have used it.

The Problem

Why are the achievements of both AI robotics and AI vision so disappointing as contributions to the goal of explaining and modelling the competences of humans and other animals (e.g. nest-building birds, primates, squirrels, hunting mammals, ...)?

SUGGESTION: a major reason is that instead of

- first trying to identify the problems that need to be solved,
- then analysing the problems, and working back to: techniques, forms of representation, mechanisms, architectures and physical designs that could play a role in working solutions to the collection of problems.

researchers normally

- start from already understood mathematical and computational techniques and forms of representation, and try to find promising ways to use them
- often focusing on a specific problem whose solution is wrongly assumed to be a necessary part of a solution to the general problem (e.g. doing well on some benchmark test!).

This mode of research seems to be a consequence of inadequate understanding of the possible primitive (but extendable and recombinable) competences that might be produced by millions of years of biological evolution (or by biologically inspired robot designers), as part of the initial repertoire of competences of a new-born animal or robot.

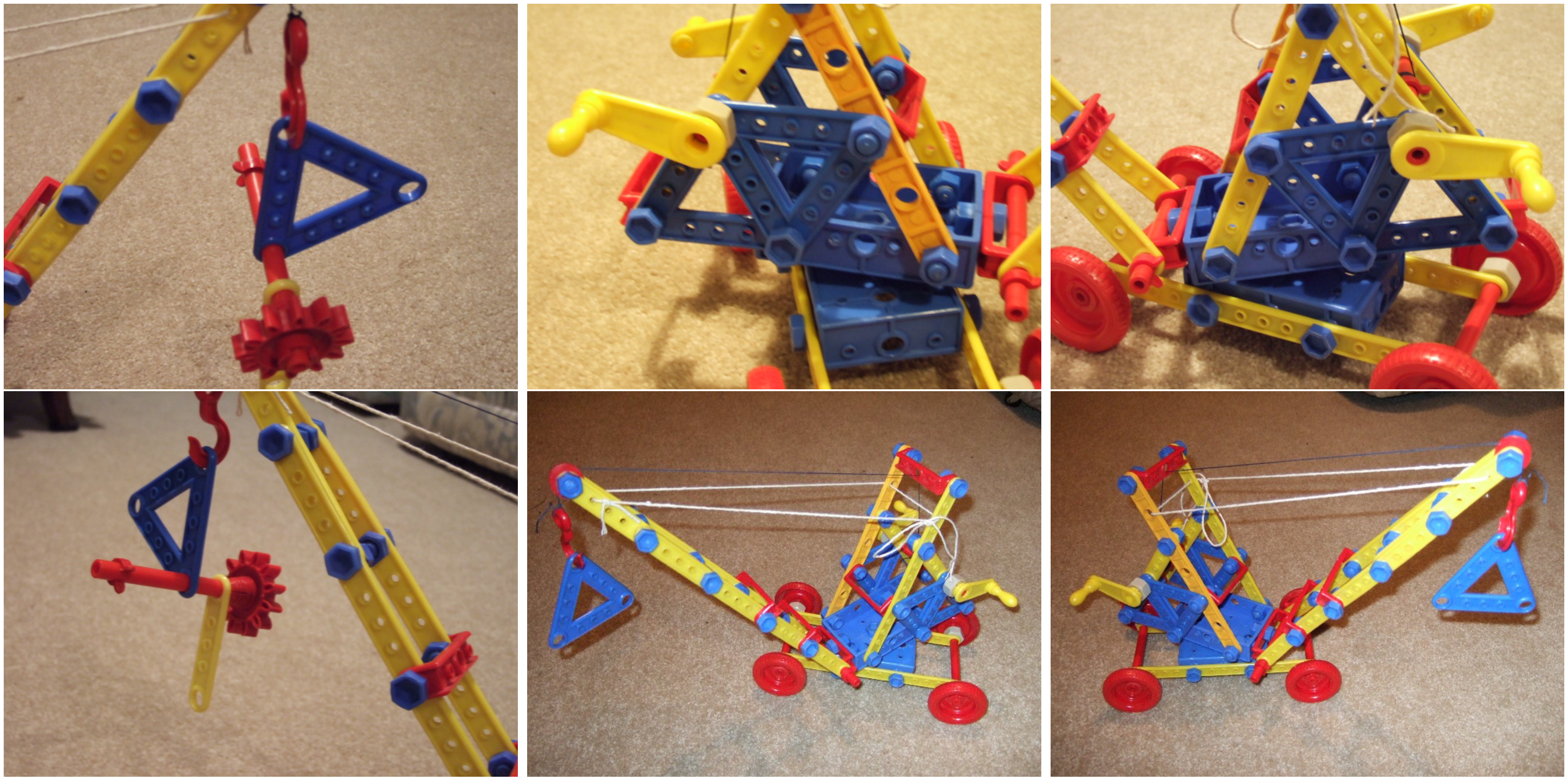
Compare John McCarthy "The Well-Designed Child" *AI Journal* December 2008
`http://www-formal.stanford.edu/jmc/child.html`

Warning: Simply investigating and attempting to replicate biological mechanisms is not a good way to be biologically inspired, if you don't know what the problems were that evolution solved.

Some of what current systems cannot do

Familiarity with roles of low level pictorial cues in representing 3-D edges, orientation, curvature of surfaces, joins between two objects or surfaces, etc., allows you to use compositional semantics to see 3-D structure, and some causal and functional relationships, in pictures (even static, monocular pictures) never previously seen.

How many features, relationships (topological, semi-metrical, metrical, causal) can you see in these?



`http://www.cs.bham.ac.uk/research/projects/cosy/photos/crane/`

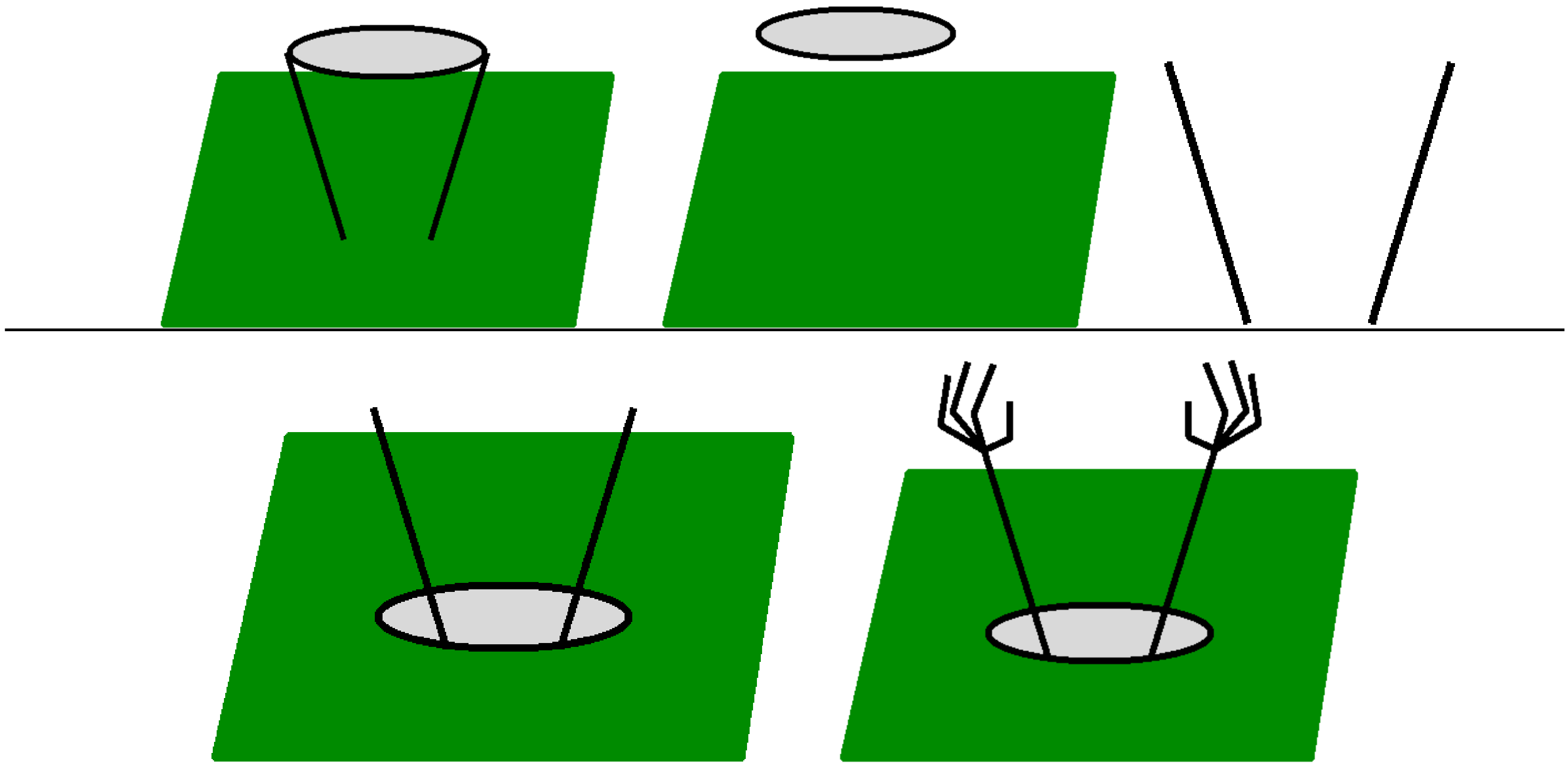
No AI vision system comes close to being able to do that – yet.

Different combinations of the same elements

What do you see in these pictures? Only 2-D configurations?

Notice how context can influence interpretation of parts.

Perceptual compositional semantics is highly context-sensitive.



Words can add more context: Strong worm catches early bird?

What about: Shark-infested sewer?

How can we make progress?

We need a shift of attention – to thinking about what it is in the **environment** that a new-born infant or robot needs to be able to represent, perceive and think about.

CONJECTURE:
The answer should at least include (amodal) representations of the following:

- bits of stuff (of various kinds) that can occur in the environment
- bits of process (of various kinds) that can occur in the environment
- ways of combining them to construct larger structures and processes in the environment (not necessarily with global consistency)
- at various levels of abstraction: metrical, semi-metrical, topological, causal, functional....

  Semi-metrical representations include things like: "W is further from X than Y is from Z", orderings with gap descriptions, symmetries and partial symmetries. (And other things, still to be determined.)

  Semi-metrical distance and angle measures could include comparisons between distances and angles instead of use of global units, like 'cm' or 'degrees'.

  Instead of items in the environment being located relative to a single global coordinate frame, they could be embedded in (changing) networks of more or less local relations of the above types.

  Powerful constraint-propagation mechanisms will need to be available for vision, haptic perception, reasoning, planning, predicting, etc. to work.

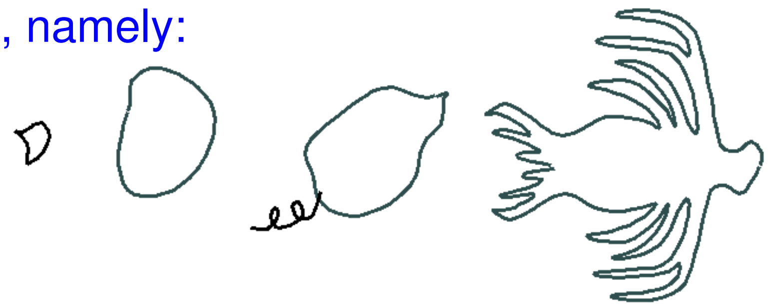
Life is information processing – of many kinds

The world contains: matter, energy, information

Organisms acquire and use information,
in order to control how they use matter and energy
(in order to acquire more matter, energy and information,
and also reproduce, repair, defend against intruders, dispose of waste products...).

Somehow evolution produced more and more sophisticated information processors.

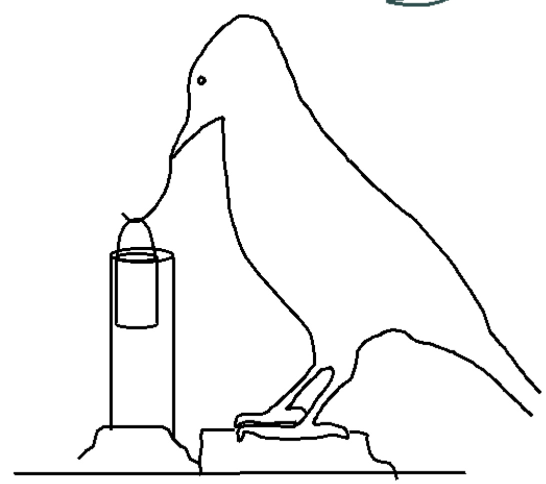
These pose challenges for science and engineering, namely:



- To understand that process.
- To understand the products.
- To replicate various aspects of the products.

We need to understand

- the structure of design space
- the structure of niche space
- the many design tradeoffs linking them
- the possible trajectories in design space,
- the possible trajectories in niche space,
- the many complex feedback loops linking both.



Development of environment and cognition 1

The cognitive system, including sensory mechanisms, motor control systems, learning systems, motivational mechanisms, memory, forms of representation, forms of reasoning, etc. that an organism (or robot) needs will depend both on

- what is in the environment
  and
- what the physical structure and capabilities of the organism are.

For a micro-organism swimming in an ever changing chemical soup it may suffice to have hill-climbing mechanisms that sense and follow chemical gradients, perhaps choosing different chemical gradients according to the current needs of the organism.

As the environment becomes more structured, more differentiated with more enduring objects and features (e.g. obstacles, food sources, dangers, shelters, manipulable entities) and the organisms become more articulated, with more complex changing needs, the information-processing requirements become increasingly more demanding.

As more complex information processing capabilities develop, the opportunities to observe, modify and combine them in new ways also develop.

See: Diversity of Developmental Trajectories in Natural and Artificial Intelligence
AAAI07 Fall Symposium
`http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0704`

Development of environment and cognition 2

The cognitive system, including sensory mechanisms, motor control systems, learning systems, motivational mechanisms, memory, forms of representation, forms of reasoning, etc. that an organism (or robot) needs will depend both on

- what is in the environment

  and

- what the physical structure and capabilities of the organism are.

Many researchers who emphasise the importance of embodiment of animals and robots make a mistaken assumption:

> they claim that embodiment and physical morphology solve the problems and reduce the burdens on cognition, by producing required results "for free" when movements occur.

However, the point I am making is that

> **As bodies become more complex, with more parts that can be moved independently to cooperate with one another in performing complex actions on complex, changeable structures in the environment, the cognitive demands (for perception, learning, planning, reasoning, and motor control) increase substantially, requiring more powerful forms of representation and more complex information-processing architectures.**

For more on this see `http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0804`

A. Sloman, "Some Requirements for Human-like Robots: Why the recent over-emphasis on embodiment has held up progress". in *Creating Brain-like Intelligence*
Eds. B. Sendhoff and E. Koerner and O. Sporns and H. Ritter and K. Doya, 2009, Springer-Verlag

Turing's mistake?

A major challenge for such an investigation is

- to understand the variety of possible starting points
- for an individual born or hatched in a particular sort of environment,
- after millions of years of evolution of the species

In his 1950 *Mind* article, "Computing machinery and intelligence", Turing wrote:

> "Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child brain is something like a notebook as one buys it from the stationer's. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child brain that something like it can be easily programmed."

On this point (little mechanism and much space), Turing was uncharacteristically badly wrong, like all the AI researchers who try to find a small number (some hope **one** will suffice) of powerful, general, learning mechanisms that can learn from arbitrary data.

Evolution did not produce general-purpose data-miners.

- Most species produced by evolution start off with almost all the information they will ever need, leaving only scope for minor adjustments of parameters, e.g. for calibration and minor adaptations.
- A few species learn a lot using mechanisms that evolved to learn in a 3-D world of static and changing configurations of objects, including other intelligent agents:
  they start with powerful special-purpose mechanisms.

Evolution is a general-purpose data-miner, changing what it mines
But it needs something like a planet-sized laboratory, and millions of years, to produce things like us.

McCarthy does not agree with Turing

In “The Well-Designed Child” John McCarthy wrote:

“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?

“Animal behavior, including human intelligence, evolved to survive and succeed in this complex, partially observable and very slightly controllable world. The main features of this world have existed for several billion years and should not have to be learned anew by each person or animal.”

`http://www-formal.stanford.edu/jmc/child.html`

To be published in the AI Journal (December 2008)

McCarthy grasped an important point missed by Turing (and by many AI researchers).

McCarthy’s own theories about requirements for a neonate are tempered by his goal of attempting to see how much could be achieved using logic.

We need to keep an open mind as to which forms of representation and modes of syntactic composition and transformation may be required, or may be useful at times.

As argued in 1971 in: Interactions between philosophy and AI: The role of intuition and non-logical reasoning in intelligence.

`http://www.cs.bham.ac.uk/research/cogaff/04.html#200407`

I am not arguing **against** the use of logic, but **for** a search for additional (new) forms of representation.

Developmental psychologists vs Designers

Many developmental psychologists investigate what is and is not innate in newborn humans, and other animals.

Examples studying humans include (among many more):
E. Spelke, P. Rochat, E. Gibson & D. Pick, A. Karmiloff-Smith, and much earlier J. Piaget,

and studying animals:
N. Tinbergen, K. Lorenz, J. Goodall, W. Köhler, E.C. Tolman, I. Pepperberg, M. Hauser, A. Kacelnik (and colleagues), N. Clayton, S. Healey, F. Warneken, M. Tomasello,

Unfortunately not enough of these researchers have learnt to look at something done by a child, chimp, or chick and ask

How could **that** work? What else can the mechanisms do? **How** do they do it?

Instead most of them ask questions like

- Under what conditions does this happen?
- How can the task be made easier or more difficult for species X?
- Is this innate or learnt?
- If it is learnt what triggers the learning?
- Which other animals can and cannot do it?
- How early does it happen?
- Which additional tests can I perform to detect these and similar competences?

They don't adopt what McCarthy calls "the designer stance".

Liz Spelke said she could not see any connection between her work and robotics when she gave her address on "Core Knowledge of Number and Geometry" at the euCognition conference in Nice (slides available online at `http://www.eucognition.org/inaugural.htm`)

In order to think like a designer you need to think about requirements as well as about mechanisms, formalisms, architectures, algorithms, etc.

Many of the requirements
for animals and robots,
even for their cognitive systems,
come from features of the environment.

We need to understand the environment

Many of the requirements for biological information-processing (including cognitive systems) arise from features of the environment.

Ulric Neisser (the psychologist) wrote about psychologists: “We may have been lavishing too much effort on hypothetical models of the mind and not enough on analyzing the environment that the mind has been shaped to meet.”

Cognition and Reality, 1976, San Francisco: W. H. Freeman.

Exactly the same comment can be made about roboticists and vision researchers.

It is tempting to assume that the environment is made up of the things we think about and talk about, such as named types of objects with named properties and relationships (trees, chairs, clouds, doors, people, table-legs, rooms, corridors, and properties like size, colour, hardness, relations like touching, being on, supporting, actions like grasping, throwing, pushing, lifting, and functions like being used for cutting or eating in, etc.)

However, the human abilities to use those categories and labels build on much older biological capabilities shared with other animals that do not have our language and concepts but can do many of the things we do, including moving around, looking, touching, prodding, climbing over, pushing out of the way, picking up, peeling, eating, hanging onto .... the cognitive competences required for combining several such capabilities in a single individual have never been replicated in AI.

In order to look at, prod, lean on, climb over, eat, or dismantle X, you don’t need to have a notion of what sort of label humans might attach to X. Neither does a self-motivated robot.

Perceiving and acting on do not presuppose **recognising** everything perceived or acted on, or **using categories used in some linguistic community.**

Relevant features of the environment

Many of the requirements for biological information-processing (including cognitive systems) arise from features of the environment.

The features include (see also McCarthy's paper):

- Topological, geometrical, and physical complexity (generic),
- Detailed (particular) structures of the environment,
- The processes that can and cannot occur in the environment
- The causal powers of various portions of the environment.

  J.J.Gibson's affordances are a special case:

  there are also proto-affordances, vicarious affordances, epistemic affordances, etc.
- The kinds of stuff (kinds of matter) that produce such causal powers
- The kinds of information available from the environment in various circumstances.
- The presence of other information-processors in the environment.

  Understanding and interacting with them requires meta-semantic competences:

  including coping with referential opacity.

Don't assume that well-known, widely used, mathematical formalisms like vectors of numbers, matrices, differential equations, probability distributions, will necessarily suffice to represent either the problems or the solutions.

There are other ways to divide up and represent the world.
Probably many more are used unconsciously than researchers have thought of so far.

What should an initial meta-ontology contain?

A starting state would include a meta-ontology assuming that there are different kinds of material and processes involving materials, in the environment, some of whose properties can be discovered by exploring the environment, while others have to be postulated when constructing explanatory theories.

(A Kantian approach?)

I suspect that the perceptual system will need an ontology that includes

- fragments of 3-D surfaces (boundaries of bits of stuff) of various types, including types analogous to
  - texture, kinds of curvature, hardness, squishiness, etc.
- 3-D features where surface properties vary or different surfaces meet,
  - e.g. features analogous to edges, cusps, dents, cracks, etc.;
- many process primitives e.g.
  - surface fragments translating, rotating, flexing, pulsating, altering curvature, and also
  - changing relationships between 3-D surface fragments, e.g. changes analogous to approaching, receding, enclosing, coming into contact, moving past, etc. etc.

Of course, a baby, or baby robot, need not be aware that it has any such meta-ontology, or that it has an ontology: such meta-semantic competence seems to grow later, possibly because it needs to have a rich platform to play with before it starts growing.

See Chappell and Sloman 2007; `http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609`

The initial ontology for the environment

Besides the initial meta-ontology, that provides a basis for exploring the environment and extending the ontology, there will also need to be some initial ontology – perhaps different initial ontologies used in different parts of the system (e.g. the baby's sucking subsystem may have a lot of features not shared with the rest of the baby's mechanisms).

The initial ontology of 3-D structure fragments and 3-D process fragments will have to be very much richer (at least in its generative power) than anything found in current AI vision systems, or current robots

Probably most will NOT correspond to familiar words or phrases in our language.

If you think you can describe in words what a neonate's percepts are then you've probably got it wrong

But features of the initial ontology and meta-ontology may help to shape deep features of the language for communication that develops later: for the baby will need to communicate about information (or some of the information) that is already important for it.

See
`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang`
Evolution of minds and languages. What evolved first and develops first in children:
Languages for communicating, or languages for thinking (Generalised Languages: GLs)?

Initial forms of representation and composition

How should instances of the initial ontology (e.g. 3-D surface fragments, or 3-D process fragments) be represented so that

- the information can be used in various ways, e.g. controlling actions, explaining what happens
- composition is possible, e.g. for perceiving novel configurations, planning new action sequences, predicting what will happen
- representation of composition allows for spatial interactions,
- the need to add novel structure fragments and process fragments to the ontology can be detected
- the ontology is extended usefully
- mistaken ontology changes can be corrected

It seems clear from the competences of precocial species like chickens, deer and others that evolution is capable of producing visual systems that are very powerful from birth or hatching (even if some of that power depends not just on the genome but also on the typical environment provided in developmental processes in the womb or egg).

So it is unlikely that the neonates of altricial species like humans, apes, hunting mammals have none of those visual competences, even though they don't appear to be using such competences initially.

If human infants are using their initial visual and action capabilities primarily in order to identify and repair bugs and gaps in those capabilities, what sorts of laboratory experiments and observations could reveal what is going on?

Should information users be in the initial meta-ontology?

A very important aspect of the environment for adult humans and for some predator animals and some of their prey is that the environment contains active individuals that acquire and use information.

If a baby robot is to acquire knowledge about other information users what initial competences will it need?

Possible things the initial meta-ontology may or may not allow for:

- Varieties of ways of getting information from the environment

  (seeing, hearing, feeling, ...)

- Varieties of actions that can be produced by such an entity

- Varieties of internal states and processes mediating the relationship between perception and action

  (percepts, beliefs, competences, preferences, motives, enjoyment, plans theories, ...)

Most people believe human infants treat other humans differently from inanimate objects, though the evidence is not clear.

However young chicks and ducklings certainly imprint on something with appropriate characteristics and behaviours.

Many infants react to adults as sources of food (e.g. begging, seeking a nipple and sucking, etc.)

It is not obvious whether a baby robot could somehow generate for itself concepts like “belief”, “desire”, etc. and if so whether that would be different from generating concepts like rigidity, electric charge, ...

Tempting mistaken questions

Much research in developmental psychology attempts to discover at what age infants or toddlers first acquire some knowledge or competence.

That pursuit can be misguided for various reasons

- There need not be any particular age, if different individuals (possibly in different physical environments and different cultures) learn things in different sequences:

  human competences are more likely to form a partial ordering than a total ordering, as shown in part by their resilience in the face of various physical and other abnormalities, lengthy illnesses, etc.

- Often the labels used refer not to a single unitary competence but a complex cluster of related competences whose detailed structure has never been unravelled

  and which develop at different times in varying orders.

- The experiments typically probe what's going on in the child in a very superficial way

  e.g. use of eye-movements in infants (because that's the only method available)

- The research rarely adopts "the designer stance":

  (a) researchers don't ask questions about forms of representation, mechanisms and architectures for acquiring, processing and using different sorts of information, what the detailed epigenetic mechanisms and processes are (e.g. growth of an information-processing architecture)

  (b) they don't formulate their theories in such a way that a robot designer could use them to design any of the mechanisms needed for a robot to demonstrate the competences being studied.

  (c) worse – many of them don't even notice they are not doing either (a) or (b) because they have no understanding of what those entail – as I have sometimes discovered with dismay when asking questions in seminars.

Should the learner be in its own initial meta-ontology?

It is very likely that the vast majority of organisms do not represent themselves: this could be true of organisms whose knowledge of the world takes only a sensorimotor form, i.e. patterns in and relations between sensor and motor signals.

A microbe probably does not need to know that there is something doing the sensing and acting. Perhaps some insects do need to know this?

There's no obvious reason why evolution should not provide information about a self/non-self distinction as part of the initial ontology or meta-ontology.

If such a thing is not innate in some form, what kind of learning mechanism could discover the need for it?

I suspect it may be a precondition of some important kinds of learning – that depend on knowing the differences between things that are merely observed to happen and things that the observer makes happen.

But many philosophical discussions about "the self" are disguised nonsense. See

`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/the-self.html`
"The self" – a bogus concept.

Implicit self-models

The example below of the baby on a rug looking round and reaching over, and many other examples, suggest a need for even very young infants to regard the environment

- not just as containing space, stuff, things, processes, and actions,
- but also as a source of information of various kinds.

Insofar as the infant has, or acquires, and uses information about epistemic affordances, and how they can be changed by the infant's actions (e.g. what you can see changes as you move, or rotate, or manipulate something) there is an implicit model of "self" in the mechanisms.

- But it need not have any explicit representation of the fact that it is doing those things, or that it is distinct from the environment.
- The precise nature and function of that model needs to be specified in more detail.

  That may be difficult, especially if it is only implicit – i.e. compiled into architectures and algorithms.
- Does a bee that adjusts its orientation, velocity, posture, etc., as it lands on a flower have an implicit self-model?

There is a lot of nonsense talked about "the self" as if there were a special sort of non-material entity referred to by that phrase. I am not using that sort of notion. See
`http://www.cs.bham.ac.uk/research/projects/cogaff/misc/the-self.html`

What forms of action representation are appropriate?

A robot needs a way to represent its own behaviours if it controls them.

What aspects of its actions must it have innate ways to represent?

Most current roboticists tend to think of actions in terms of vectors, matrices, differential equations, trajectories in a global 3-D euclidean space, and probabilistic processes.

Is that the only way to proceed?

I suspect a different class of mathematical representations and methods will be required to support innate vision and action in a human infant, depending far more on networks of more or less local relations

including topological and semi-metrical relations, e.g. W is further from X than Y is from Z, and symmetries or partial symmetries, rather than euclidean coordinates.

Powerful constraint-propagation mechanisms will need to be available for vision to work.

This is still very vague but the implication is that contrary to Turing's assumption, and the assumptions of most robotic researchers that I have encountered, a robot that is to develop human-like competences will have to start with very much richer initial competences than anyone currently knows how to program.

Those initial endowments are unlikely to be achieved by evolutionary experiments in computer labs because they cannot easily replicate all the problems that evolution had to solve in our evolution.

Both the environments used and the fitness functions used in such experiments are far too simple and restrictive.

Ontologies for neonate physical agents

If all the above is right, we need to think very creatively about what the perceptual and other functions are, guided not by how adult humans describe the world, but by

- how young animals of various species (including human infants and toddlers) interact with the world,
- how their interactions provide the foundations for growth of more sophisticated and successful interactions.

The existence of "precocial" species whose young can perform complex actions in the environment soon after birth, including finding the mother's nipple, running on rough terrain, pecking for food, etc., demonstrates that evolution can produce powerful innate visual systems.

The apparent incompetence of human neonates may be deeply misleading.

What sorts of capabilities can such systems have?
What sorts are required as a basis for developing human-like competences?

Conjecture:

- **for a neonate the world is primarily made of bits of stuff**
- **and bits of process involving bits of stuff,**
- **all enmeshed in a web of spatio-temporal and causal relationships**
- **with topological and semi-metrical features and relations**
- **with local but not necessarily global consistency**
- **NOT anything expressible in a global coordinate system.**

Basic perception and action

CONJECTURES: Compare McCarthy's Well designed child.

A 3-D actor in a 3-D environment has perceptual and motor access to certain kinds of "primitive" constituents from which environments are composed: various kinds of **bits of stuff** and **process fragments** in various kinds of static and changing relationships of three main kinds

- spatial
- temporal
- causal

- Processes involving some bits of stuff (i.e. its own parts) are controllable by the agent.
- Many bits of stuff endure over time.
- Many processes involving bits of stuff endure over time.
- All the bits of stuff and processes exist in a spatio-temporal realm that extends beyond what the agent can access.

  spatio-temporal locations and regions, and bits of stuff and processes in them can exist prior to being accessed, can endure after being accessed and may exist without ever being accessed by the agent — though some of their effects may be accessed, and some of the effects of what the agent does may affect them.

- The information gained about the environment depends in part on the perceiver.
  It can select not only which portions to perceive but also what processing to do, e.g. what level of abstraction to use, whether to look closely or not, whether to look and feel or just one.

Do not make these tempting but false assumptions:

- That the perceptual competences listed here are too advanced for neonates:

  on the contrary, **some** species (precocial species, competent from birth or hatching) must have very sophisticated 3-D visual ontologies and action competences largely determined by their genes.

  Evolution had a lot of time and used it well.

- That the infant learner must start with some conception of a global coordinate frame in which it can represent coordinates of objects, their parts, their directions of motion, their mutual distances and directions.

  Instead, many details may be represented in collections of locally consistent semantic fragments, e.g. referring to parts of objects, and parts of actions.

- That everything the individual does is (at least initially) represented solely in terms of sensor signals and effector signals, and relations between them.

  "Symbol grounding theory" is seductive but leads to error and confusion, and needs to be resisted.

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

- That for a neonate the world must be "a blooming buzzing confusion" of signals with little structure or content (as supposed by William James).

  It is more likely that, for many species, the millions of years spent evolving produced genomes that start with a lot of very specific representational and investigative capabilities.

- That information gained visually is represented visually (in some image-like format).

  On the contrary amodal forms of representation will be more useful for some perceivers that can act in the world and use more than one form of perception, e.g. visual and haptic perception of 3-D structure.

Genome (+) at work on a rug - after about 6 months

Video available online:
http://www.cs.bham.ac.uk/research/projects/cosy/conferences/mofm-paris-07/sloman/vid/

The infant seems to produce a new combination of previous competences: looking round, identifying an edge, setting up a goal, rolling over, stretching arm, opening fingers, moving down, closing fingers, pulling.

This probably builds on a host of visual and other competences developed in previous months, building on (still unidentified) competences provided by the genome (plus common features of the human intrauterine development environment).

Human learning seems to progress in layers of competence, where new layers include components in old layers as "objects to play with" plus new layers in the ontology, not definable in terms of earlier ontologies.

http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609
Natural and artificial meta-configured altricial information-processing systems (Chappell & Sloman)

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang
What evolved first: Languages for communicating, or languages for thinking?

Prerequisites for being an information user (1)

- Need a form of representation (FOR) – with (unbounded?) generative power
  - medium
  - syntax (structural variability, context-based compositional semantics) (E.g. CCG ... and others)
  - mechanisms for manipulating representing structures
    (building, matching, decomposing, combining, modifying, parsing ....)
- Need a set of concepts to be expressed using a subset of the F.O.R

  NB: not just unitary predicates: also relations, functions, modalities,...
  NB: not necessarily all using a linguistic/logical form, e.g. spatial & other forms.
- Need to be able
  - to recognise need for ontology extension, and
  - to extend subset used.
- Need to be able to link a subset of instances of F.O.R to
  - sensory inputs
  - motor signals

  NOT **symbol grounding** but **theory tethering** is needed.

We can expand on all that, as follows:

Prerequisites for being an information user (2)

- Need a form of representation (FOR)
  - based on a medium with appropriate properties (e.g. structural variability, rapid changes, ...)
  - syntax (structural variability, context-based compositional semantics) (E.g. CCG ... and others)
  - supporting forms of "binding" of information that are appropriate to the nature and uses of the information (e.g. spatial juxtaposition as well as Fregean application of predicates, relations, functions)
  - with generative power so that novel configurations can be accommodated
    - In the case of humans generative power is unbounded: which other species have that, and how did it evolve?
    - How did the ability to notice that it is unbounded evolve?
    - How does that ability develop in individuals?

- Need to be able to use those representations for a variety of functions,

  including: perceiving; forming goals; forming plans; formulating and debugging theories; selecting, generating, controlling actions; learning regularities; episodic memories; identifying individuals; ...

- Need a set of concepts to be expressed using a subset of the F.O.R

  NB: not just unitary predicates: also relations, functions, modalities,...

  NB: not necessarily all using a linguistic/logical form, e.g. spatial & other forms.

Does the baby also have to "know" (implicitly) about epistemic affordances?

e.g. that swivelling head (and eyeballs) will select a new portion of the environment from which (exosomatic) information is available?

(Information that can be checked haptically and used by motor subsystems.)

Contrast that with the assumption that swivelling head (and eyeballs) will alter the visual sensory input (i.e. provide new modal, somatic information.).

Prerequisites for being an information user (3)

- Need a form of representation – with (unbounded?) generative power
  - medium
  - syntax (structural variability, context-based compositional semantics) (E.g. CCG ... and others)
- Need a set of concepts to be expressed using a subset of the F.O.R

  NB: not just unitary predicates: also relations, functions, modalities,...

  NB: not necessarily all using a linguistic/logical form, e.g. spatial & other forms.
- Need to be able to recognise need for ontology extension, and to extend subset used.
- Need to be able to link a subset of instances of F.O.R to
  - sensory inputs
  - motor signals

  NOT **symbol grounding** but **theory tethering** is needed.
- Need to use subsets of F.O.R to express
  - theories about type of world
  - facts about contents of parts of world (episodic memory) (immediate/remote)
  - goals
  - plans
  - questions
  - conjectures
  - possibilities
  - states of information users (self and others)
  - bugs, deficiencies, problems to be fixed
- Need an architecture to support all this. An extendable architecture.

Varieties of learning/development

To be extended.

See for example:

Eleanor J. Gibson and Anne D. Pick, (2000)
An Ecological Approach to Perceptual Learning and Development
Oxford University Press, New York

Even they have blind-spots.

E.g. they talk about two ways to learn new affordances:

- Empirically (by trial and error exploration)
- By imitation (though how the imitation process works is unspecified)

But they fail to mention what is probably a more and more important case as the individual gets older and more mature:

- By working things out Often using what I call "toddler theorems"
  See `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mkm08`
  The mechanisms used by an infant or toddler for working things out form the basis of mathematical competences that develop later.

Composition/Binding

These different aspects of reality can be composed/combined in many different ways.

Long before there was algebraic/functional/logical composition there was spatio-temporal composition.

Also auditory/temporal composition – music and many natural sounds.

We need to distinguish

- Composition in the spatio-temporal environment

  e.g. combining actions and things acted on, or sounds

- Composition in internal representations of things that can be spatio-temporally combined: i.e. composition in representations in virtual machines.

At present we have only a relatively small number of forms of information-composition that we can implement and use in computers.

By studying the environments of various sorts of intelligent systems very carefully we can derive new requirements for forms of representation and forms of composition and manipulation.

This may lead to the creation of new kinds of artificial information-processing systems.

Worrying about human uniqueness may be premature

In order to say anything sensible regarding what is unique about humans, we first need a deep theory of how animals in a wider variety of species do what they do.

There may be unsuspected aspects of human competence that are largely accounted for by design features common to humans and some other species.

Even aspects of our ability to learn human languages may depend on a larger shared evolutionary heritage than has been generally recognised.

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

It is also possible that some of the deep features common to the semantics and syntax of all human languages have their origin in requirements for meta-ontologies of a wider class of intelligent animals.

Some of those features will arise from requirements imposed by the complexity of the physical world, and some from requirements imposed by the need to live as social animals.

Only when whe have deeper, more complete theories about other intelligent animals will we know what needs to be explained about the uniqueness of humans.

The common biological basis of adult human intelligence

CONJECTURE 1

Almost all human adult cognitive competence builds on and grows out of

- very general biologically-based, culture-neutral competences
- developed through interaction with the environment in the first few years of life,
- driven by a powerful collection of genetically determined (meta-configured) learning systems that evolved specifically for learning in rich and changing 3-D environments.
- developed over millions of years by biological evolution
- whose functions and mechanisms have little connection with current AI/Robotic mechanisms,
- whose explanation is well beyond the current state of neuroscience

CONJECTURE 2

All attempts by AI developers to implement adult-like human competences without going via that route will be very brittle and severely bounded in scope.

CONJECTURE 3

The early learning of human infants and toddlers, and many other animals, starts with powerful mechanisms for using an ontology of perceivable, manipulable bits of stuff, in static and changing configurations (processes).

Spelke on growth of knowledge

Elizabeth Spelke asks (in her 'Six Suggestions' paper):

How do humans go beyond core knowledge and construct concepts and cognitive capacities that are unique to humans and variable across cultures?

She offers two possible answers

- 1. There are other, uniquely human systems of core knowledge.
  Examples:
  - a core system underlying communication and cultural learning (Tomasello; Gergely & Csibra)
  - a core system for reasoning about coalitions, cooperation and competition, social groups (Cosmides & Tooby; Dunbar)
- 2. There are uniquely human processes by which children go beyond the limits of core knowledge.
  Examples

  natural language may serve to combine representations from different core domains, both flexibly and productively.

  Language may provide a medium for combining information rapidly and productively, overcoming the limits of domain-specific, encapsulated core knowledge systems.

  She is apparently thinking of human communicative language, rather than the kind of generalised language (GL) that first evolved to support perception, thinking, planning, control of actions, and only later led to something usable for communication.

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

To Be Extended

THIS IS AN INCOMPLETE DRAFT

It may take several years (or more) to reach maturity.

The work of Peter Simons (Leeds University, about to move to Trinity College Dublin) is relevant e.g.

Real Wholes, Real Parts: Mereology without Algebra. *Journal of Philosophy* 103 (2006), 597-613.
http://www.journalofphilosophy.org/articles/103/103-12.htm (Unfortunately not freely available.)

Should I be looking at small-world networks?
http://en.wikipedia.org/wiki/Small-world_network

Am I ignoring social aspects of learning and development?

Much of what I am referring to occurs in non-human animals. So the specific features of human social interaction (e.g. child-parent interaction) cannot be a requirement for these developments, though they may facilitate them. In particular, the kinds of things parents provide in the physical environment of infants and toddlers, and the process of allowing them to interact with increasingly complex (and sometimes increasingly dangerous) physical environments may have a major role in how development occurs.

However, there does not have to be a totally ordered developmental sequence: different routes through a partially ordered network are possible, especially for children with different disabilities, e.g. deafness, blindness, missing or deformed limbs, control dysfunctions (cerebral palsy?) etc.

Related work

http://www.cs.bham.ac.uk/research/projects/cosy/papers/
CoSy papers and presentations

http://www.cs.bham.ac.uk/research/projects/cogaff/
The Birmingham CogAff project

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/
A collection of (mostly PDF) research and tutorial presentations (including this one).