INCOMPLETE DRAFT NOTES:

(Part of CoSy deliverable DR.2.1:

Requirements study for representation)

(Last revised October 9, 2005)

Towards an ontology for factual information for a playful robot

Aaron Sloman
University of Birmingham, UK
http://www.cs.bham.ac.uk/~axs*

October 9, 2005

---

* With much help from the CoSy Project Team (See URL http://www.cognitivesystems.org/) in Birmingham (See URL http://www.cs.bham.ac.uk/research/projects/cosy/) and elsewhere.

Contents

1 Introduction: the need for ontologies

One of the long term objectives of the CoSy project is understanding how to achieve, within a human-like robot, integration of multiple capabilities that exist in humans, and to some extent in other animals. This includes perceiving, acting, learning, reasoning, and communicating in a complex environment, such as a family home. This requires the robot to understand both general facts about its environment, such as what sorts of things can exist or occur in it and how various sorts of things normally behave, and also particular facts about what actually exists and what actually happens.

Various kinds and levels of *self-understanding* or more generally, *understanding of information-processing systems, including other agents*, will also be important, in order both to explain and debug one's own actions, decisions and reasoning and also to think about corresponding mental states in others. In particular, being able to think about mental states and processes in others (i.e. states and processes with semantic content in others) is required in order to predict or understand their actions, or to help them. Just as it can be useful to know about size, shape, location, contents, strength, functions, etc. of physical objects, so can it also be useful to know how another individual sees things, thinks about things, understands things, or what the individual wants, likes, dislikes, is afraid of, etc., in teaching situations Sloman (1992), and in situations where advice or help is given to someone at an earlier stage of development or who lacks some information because it is out of sight for him.

In all these situations, the robot (let's call it Fido, as in the CoSy work-plan, page 11) will need to deploy an ontology, or more precisely a collection of ontologies. In this paper we begin to analyse some of the requirements for its ontologies, restricting our attention for now to physical environments excluding other intelligent agents. Later work will need to address extensions to the ontology to include animate individuals, which for our purposes can be characterised as things that acquire, transform, use, or communicate information.

Note

> The word "ontology" is now often used in AI and software engineering circles, though the concept is much older within philosophy, as explained below. As a result of these different contexts of use, the word has acquired a number of different though related meanings. Appendix A provides a brief history and compares different uses of the word. In an engineering context we can distinguish a *design-ontology*, used by designers to produce some system, which may or may not be an intelligent system, and an *application-ontology* which is used by an intelligent system for whatever its tasks are, which need not include designing anything. This paper is primarily concerned with the "application ontology" for a type of robot that manipulates 3-D objects on a table top. As such it is also a subset of the design ontology for people designing such robots. For now we ignore questions about what is explicitly programmed into the robot and what is learnt through interactions with the environment, with or without a teacher.

2 Background, and relation to question-ontology paper

This is one of two papers on ontology developed within CoSy work-package 2 (Representation of space, objects and context), both submitted as part of deliverable DR.2.1 (Requirement study for representations[1]). The companion paper *'Towards an ontology for factual questions'* is concerned with a largely unnoticed requirement for an intelligent agent, namely the ability to discover gaps

---

[1] Available at http://www.cs.bham.ac.uk/research/projects/cosy/papers

in its information, so as to be able to take steps to fill the gaps. Doing that requires a means of representing those information gaps, which in ordinary parlance amounts to a means of expressing questions. So the parallel paper presents a first draft ontology for questions. Identifying information gaps and taking steps to fill them need not involve communication with other individuals (e.g. asking questions) but is part of the information-processing *within* an agent. However it may be a pre-requisite for linguistic competence that involves dialogues in which questions are asked and answered.

The ontology for questions presupposes that the questioner already has an ontology for the *contents* of questions and answers to questions, namely an ontology for factual information. This paper presents a first draft incomplete ontology for factual information for a CoSy robot (or perhaps a child, or some other animal interacting with objects in the environment). We ask what sorts of information our robots might need, and summarise in a very general way some of the representational requirements for expressing such information. The topic is potentially vast, but has been constrained here by the needs of the CoSy robot scenarios, initially focusing mainly on a subset of the PlayMate Scenario.[2] This scenario is based on a robot able to perceive, manipulate and talk about 3-D objects on a table top, including eventually assembling complex objects from components (hopefully going beyond Freddy[3] the Edinburgh robot which was able to assemble a few toy objects about 30 years ago, e.g. making a toy wooden car from a body with holes, two axles and four wheels).

The ontological requirements of the PlayMate scenario are related to but different from the requirements for the Explorer. There is a lot of ongoing work in AI on mobile robotics on which the Explorer scenario can build, but relatively little on manipulation of 3-D objects, a kind of task which was accomplished relatively late in biological evolution and which seems to have had a crucial influence on human cognitive evolution. This document focuses mainly on requirements for the PlayMate, but will later be extended to include the Explorer. Insofar as the two requirements overlap, it already addresses the latter.

Another restriction is that the ontology presented here does not include other intelligent agents in the environment. So it is concerned only with objects, events and processes involving inanimate physical objects, except insofar as the agent thinking about such objects may act so as to cause, modify or prevent physical processes in the environment.

3 This is about pre-linguistic competence

Although the PlayMate scenario ultimately has to include linguistic competence, it is likely that such competence will build on more fundamental and general forms of competence shared by preverbal human children and other animals that manipulate objects such as New Caledonian crows.[4] So this document is a first draft incomplete investigation into some of the kinds of information that a pre-linguistic robot with perceptual and manipulative skills may need to cope with, especially one which, unlike Freddy, knows what it is doing and why.

---

[2] The CoSy PlayMate scenario is concerned with a tiny subset of the requirements for Fido!
See URL http://www.cs.bham.ac.uk/research/projects/cosy/PlayMate-start.html

[3] See URL http://www.aiai.ed.ac.uk/project/freddy/

[4] Betty, the New Caledonian crow in the Oxford Zoology lab made headlines in 2002 when she demonstrated that she could make hooks out of wire in order to retrieve food in a basket from the bottom of a tube. See URL http://users.ox.ac.uk/.kgroup/tools/tools.main.shtml

4 Propositional components for a physically embedded information user

There may not be any point trying to cover all possible propositional forms and question forms in the CoSy project. E.g. a subset of the ones a two or three year old child can understand and generate may be plenty, at least for the next few years. It is probably wise to start with mechanisms and forms of representation that do not presuppose the ability to use an external language, since those evolved first and are present first in young children, so it is very likely that they are used by the language-using mechanisms when they develop. In any case it seems that the variety of pre-linguistically comprehensible propositions is very large, and worth understanding in order to investigate how robots might work, independently of what is true of humans and other animals.

An animal, child, or robot may have a complex architecture with many different components using information of different sorts in performing different tasks, often concurrently, e.g. controlling eye vergence, controlling posture, controlling breathing, controlling direction of gaze, causing linguistic input to be parsed and interpreted, controlling the digestion of food, the insertion of hormones into the blood, the pumping of blood, etc. Different subsystems will use various specialised representations tailored to different sub-ontologies and different subtasks. Even information from the same sensory source, e.g. visual information, may be transmitted to different subsystems that use it in different ways to derive different information, represented in different formats, for instance in controlling current actions and in planning future ones, or predicting what will happen next in the environment. For now we ignore all those differences and focus only on subsystems that can use factual information that might usefully be expressed, at least partly, in a propositional form (including non-linguistic internal indexical referring devices, as mentioned previously). Later we return to non-propositional, e.g. spatial, forms of representation.

We consider, as an example, kinds of propositional structures that might be relevant to a robot perceiving and manipulating objects on a table top. (Simpler versions of the robot will cope with only a small subset of ontology presented here.)[5]

4.1 Types of entities that can be referred to

4.1.1 Physical object types

For now ‘object’ will not be defined, though most of the things called objects will be enduring, spatially-bounded, possibly moving, entities, which may have fixed or changing attributes. A more general notion of object is definable as ‘anything that can be referred to’, which amounts to the same as Quine’s notion of whatever can be a value of a variable. This would include such things as times, numbers, colours, shapes, strategies, styles, theories, proofs, explanations, problems, functions of objects, etc. For now we focus on an ill-defined subset of spatio-temporally located objects (what many philosophers would call ‘particulars’, as opposed to universals, like shapes, colours, numbers and proofs).

Any object will be an instance of one or more (usually a whole hierarchy) of object types. Some of the object types considered here are fairly abstract, but they all have instances with a

---

[5] Some aspects of this investigation overlap with the ‘Naive physics’ project of Pat Hayes and colleagues. See Hayes (1985), though that does not focus on requirements for action. For a commentary with some history see Barry Smith and Roberto Casati (1994), ‘Naive Physics: An Essay in Ontology’ *Philosophical Psychology*, 7/2 225-244. http://ontology.buffalo.edu/smith//articles/naivephysics.html

spatial location and the possibility of relationships to other objects. Examples would be types of physical object for which we have names:

- Object types categorised on the basis of function and shape: box, tray, cup, mug, saucer, lid, bowl, ball (?)
- Object types categorised only on the basis of shape: cube, sphere, cylinder, ball (?)
- Biological object types: lemon, apple, dog, ...
- Body parts: arm, hand, finger, fingernail, fingertip
- Parts of objects defined by shape, function, ...: Handle, keyhole, rim, base, ...
- Generic object parts: surface, face, edge, corner (2-D or 3-D), hole, crack,

  NOTE: We could also have typical parts for other things, e.g. plants, pieces of fruit, various animals, various utensils. various kinds of rooms, various kinds of buildings, various kinds of towns, etc.
- Abstract entities that exist by virtue of relations between other entities, gap, opening, mouth, passage, enclosure, corridor, ...
- various entities that can occur on or in a surface:
  - 2-D surface features: marks, texture boundaries, colour boundaries, shadows, edges of those, etc...
  - 3-D (shape-determining) surface features: furrows, indentations, protrusions, curvature extremes, saddle points, edges (where two surfaces meet in a well defined curved or straight line)
- Entities that exist relative to a viewpoint, e.g. visible portion of a surface or object, occluding edge of a surface (e.g. of a sphere viewed from a certain direction – the spherical surface itself contains no surface edges)
- thing (catch all ? Or meta-level concept, like a variable)

[The above is to be extended. Some additional relevant ideas about types of spatial entities can be found in CYC[6].
Some first draft thoughts about requirements are here.[7]]

In a different environment there could be different object types including various plant and animal types, types of furniture, parts of buildings (walls, doors, windows), and various types of out-door objects, such as rocks, trees, clouds, lawns, roads, etc. We ignore those complications for now.

There is no requirement that an animal or robot should be able to see *only* things for which it has names, or which it can recognize. It is clear that we are perfectly capable of seeing, and thinking about a complex structure that we have never seen before, but on parts of which we are capable of performing many actions (grasping, prodding, pressing, pushing, pulling, twisting, bending etc.)

---

[6] See URL http://www.cyc.com/cycdoc/vocab/spatial-vocab.html

[7] See URL http://www.cs.bham.ac.uk/˜axs/misc/ontology-for-a-manipulator.txt

The ability to see named, or nameable things almost certainly rests on this more basic ability shared with many other animals that act in an environment about which they cannot talk.

A pre-linguistic animal or robot may, however, use some object types that correspond to recognition states of some internal pattern recogniser without being able to use any external labels for those types even if there are internal labels of some kind. (A draft document discussing, among other things, the evolution of internal labels in biological organisms is available at

http://www.cs.bham.ac.uk/research/cogaff/vis-affordances.pdf[8])

4.1.2 'Stuff' types

The object types have well defined instances with boundaries between individuals so that they can be counted. The table-top environment might also include kinds of stuff (labelled in natural languages by 'mass nouns'). If X is a kind of stuff (kind of material), such as water, sand, mud, bread, dough, cotton, etc., then one cannot have two Xs though one can have two pieces, lumps, pools, spoonfuls, (in some cases) squirts, or stretches of X, where the additional (chunker) noun, 'piece', 'lump', 'stretch'. etc, is used to refer to a bounded portion of space filled with X.

For a stuff type X it is not possible to ask how many Xs there are without adding a chunker noun, but it is possible to ask how much X there is, since adding more X to a portion of X, or removing some X from it does not stop what is there being X, it merely alters the amount of X (which is the size, volume, weight, etc. of the portion of X). In a table-top scenario some subset of the following stuff types might occur:

- Rigid Types of stuff: wood, metal, plastic, glass, (stiff card?)
- Non-rigid Types of stuff:
  - Dry cohesive: paper, string, cotton wool, wire, foil, plastic (e.g. film), plasticine, various plant materials (non-dry twigs, leaves),
  - Dry particulate (pourable, stirrable, spillable): sand, sugar, salt, pepper, ... (Compare piles of smaller and smaller marbles, or cubes: type boundaries may not be sharp, or may be task dependent.)
  - Liquid (e.g. pourable, stirrable, spillable): water, paint (could be more or less viscous, sticky, etc. type boundaries may not be sharp, or may be task dependent.)
  - Gaseous: Steam, smoke, wind, clouds.... maybe in another project

Note that many attributes that an object has, as described below (e.g. rigidity, hardness), will be inherited from properties of the stuff of which it is composed. Understanding this relationship may or may not be required for a particular level of child or animal competence. For particular purposes more sub-divisions may be required, e.g. between breakable and non-breakable rigid objects. But such things could come later.

Insofar as surfaces are important for the PlayMate robot, it should be noted that the perceptual and manipulable qualities of surfaces will be heavily dependent on the kind of stuff involved. Picking up a portion of liquid, a piece of thread, a lump of plasticine, a wooden cube all involve different percepts and skills.

---

[8] See URL http://www.cs.bham.ac.uk/research/cogaff/vis-affordances.pdf

4.1.3 Location types

Physical objects (including chunks of stuff) have locations: they occupy space. Locations of objects are important for many purposes, including grasping them, putting things in them, avoiding them, throwing things at them, identifying them when communicating with others, working out where to go in order to see them, etc.

In addition to objects, events and processes can also occupy or occur at a location. Some locations are not filled by the object (or event) but moved through. For instance a pea in a constantly shaken small box moves around in the volume that is enclosed by the box, and that volume moves around in a larger space as the box is shaken. If a ball is thrown the location through which it moves can be variously identified as the room through which it moves the volume above the table over which it is thrown, the 'tube' which forms the space enclosing all the volumes it occupies during its motion, and no doubt many more.

Besides having spatial locations events and processes can also have temporal locations. As with spatial locations there are different sorts of temporal locations.

Not all types of locations are relevant to all contexts. For example, spatial locations on a large farm may include fields, boundaries, paths, ponds, hills, hillsides, hilltops, valleys, passes, etc. none of which occur on a table-top (although sometimes toy versions do).

A significant part of the history of mathematics has been concerned with space and its properties. Euclidean geometry and more recently topology can be seen as attempts to generalise many of the sorts of notions referred to above, while both abstracting away from features of particular physical objects and processes, and also exploring limiting cases, e.g. as locations get smaller indefinitely, or routes get thinner indefinitely, or lines get longer indefinitely, or structures get more and more symmetrical indefinitely. (Straightness, circularity, perpendicularity, are all examples of symmetry in the limit.) It may be that the human ability to understand such mathematics depends on our evolutionarily older ability to understand, perceive, act on or in, and reason about the less abstract, usually physically instantiated, spatial structures, even though it is possible in principle to produce axiomatic specifications of the abstract limiting cases and reason about them using only general logical capabilities.

There has been an enormous amount of research on ways of referring to or describing spatial locations and regions of various kinds (e.g. by Tony Cohn and collaborators at Leeds University), although much of this research has been done without regard to what the information about locations is to be used for. This paper will not attempt to summarise or reinvent that work, but merely gives some simple examples of what might be useful for a robot that manipulates and converses about objects on a table top, or perhaps a young child playing with toys, or even a nest-building bird.

Spatial locations that an object can occupy or move through, or at which an event or process can occur, that might be relevant to our 'play-and-manipulate' context include at least the following, where the word 'generalised' is used to indicate that we are not restricted to the corresponding mathematical notions.

- generalised points
  (the location at which an object is, or to which it is being moved, or which it has come from in a movement: where the sort of location that is relevant will depend on the sort of object)

- generalised regions of surfaces
  (surfaces can be carved up into regions in many different ways for many different purposes, including possible areas where something could be placed, where something could move,

where an individual can reach, or see, or in some contexts regions owned or controlled by someone or something).

- generalised volumes of space
  (Like generalised regions, only 3-dimensional).
- routes of various kinds
  (including 2-D and 3-D routes, both relative to fixed objects or relative to parts of moving objects – e.g. the route of a teardrop down one's face).
- locations on or in other objects, or locations.
  These include regions of surfaces, e.g. 'on his left cheek', or portions of volumes, e.g. 'in his mouth', 'around his left kidney', and also parts of locations and routes, e.g. near the centre or boundary of R, where R is a region, at the beginning of R, where R is a route.

A more complete discussion would also specify different kinds of temporal locations that can be referred to in propositions, including absolute locations and relative locations,

It is worth noting that the forms of representation for locations, surfaces, routes, etc. that a mathematician might first think of are not necessarily the most appropriate for a robot or animal. In particular mathematical specifications will normally be very precise so that when what is scene is unclear or ambiguous a large variety of mathematical expressions may be needed to express all the possibilities, and subsequent reasoning or planning may have to handle a large disjunction, which can lead to combinatorial search. It may be possible to simplify thinking, reasoning, planning and control of actions, if, instead of a precise specification a a less precise form of description can be used, which inherently covers a range of different possibilities (like specifying a time interval rather than a time point, or a spatial region rather than a precise location, or referring to a polyhedron rather than a cube, or tetrahedron, etc.)

CONJECTURE: the move towards less precise concepts makes it possible to avoid complex mechanisms for dealing with uncertainty.[9]

4.1.4 Other types [to be completed]

E.g. agents, animate objects, their mental sates and processes, information-processing systems, mechanisms, etc.

4.2 Attributes (of objects, locations, events, etc.)

All sorts of things can have attributes (and also relationships, discussed later). E.g. physical objects can have size, shape, colour, location (and many other attributes); locations can have extent (length, area, volume, width), shape, or location in a larger location; events can have locations, durations, start or end times, participants, causes, effects, etc.

In the sort of domain we are considering, each sort of entity has a collection of attributes which belong to different types. Each attribute type has a set of possible values, e.g. the type colour can have values red, blue, green, and possibly many others, the attribute height may have either chunked qualitative values, such as small, large, huge, or continuous measure values, e.g. 5cm, 5.1cm. 5.11 cm, etc. In some cases the values are linearly ordered (e.g. height) in others not

---

[9] As discussed in http://www.cs.bham.ac.uk/~axs/misc/ontology-for-a-manipulator.txt

(e.g. shape, discussed further below). When the values are ordered there may be some direction in the ordering that is naturally construed as *increasing* the value whereas the opposite direction is a *decreasing* direction.

Note: insofar as objects grow, develop, learn, become ill, get damaged, etc. their attributes are not fixed. It is not just that the values can change (e.g. size getting larger, colour getting grayer), it is also possible to acquire attributes in a space that was previously totally absent in the individual. E.g. when a child has learnt to do mental arithmetic it makes sense to ask about the speed at which he multiplies two numbers, whereas at an earlier stage there is no speed because he cannot multiply. Likewise if someone has no arms you cannot ask about the extent of his reach or the strength of his grip. In our initial domain we can ignore growth, learning, development, etc., though we may need to consider objects that get damaged.

Some attributes will have possible values that form ordered sets (e.g. height) whereas others will not (e.g. shape, material, uses, or species in the case of living things). Some may be ordered in several dimensions, especially functional attributes such as usefulness for a task, which may have dimensions like ease of use, quality of result of use, difficulty in learning to use, cost of use, speed of operation, etc. (The multi-dimensionality of ‘better’ is discussed further here.[10])

Here are examples of types of attributes that objects in a simple robot-manipulation domain might have

- colour
- height
- width
- overall size
- weight
- shape
- volume (e.g. of a container)
- hardness
- rigidity/flexibility
- smoothness
- graspability
- stability
- wetness
- stickiness
- warmth
- material
- function

That is not supposed to be an exhaustive list. In fact for an agent that can act on objects it is a seriously incomplete list, for there will also be a range of important object attribute types that are not definable in terms of intrinsic features of the object (e.g. as size, weight, shape and material are) but depend on what the agent can and cannot do to the object and the consequences of doing, or trying to do those things.

---

[10] See URL http://www.cs.bham.ac.uk/research/cogaff/sloman.better.html

4.3 Affordance-based object properties

In fact some of the attributes listed can be interpreted that way. For instance, instead of hardness being an objective (or intrinsic) attribute of objects (or surfaces of objects), it can be interpreted as relative to a particular agent, depending on how much resistance the surface offers to pressure that the agent can apply. Thus a surface that is hard for a weak, small agent might be soft for a much larger stronger agent. Similarly, smoothness instead of being objective may be relative to the amount and kind of resistance to motion when an agent attempts to slide one of its body parts along the surface. What smoothness amounts to for such an agent will depend on what sorts of sensors are activated by such sliding actions.

Even properties like weight and volume instead of being represented in absolute terms such as cubic centimeters and grams might, for some animals and some robots be represented in terms of kind of effort required to lift or throw them, in the case of weight, and in the case of volume in terms of what one can put in the object, e.g. a finger, a whole fist, one's whole body, or in terms of the kind of grasp or posture required to hold the object. That sort of affordance-based notion of volume will of course be closely related to shape, and will not have the kinds of invariance properties that the physicist's or mathematician's notion of volume has. (The use of an affordance-based concepts of volume or 'amount' may account for Piaget's discovery that young children do not understand conservation of volume across transformations that do not preserve shape.)

4.4 Use of predicates vs attribute-value pairs

In many cases each of the possible values of an attribute can be applied as a predicate, producing a proposition that is in some sense an abbreviation of a proposition specifying both the attribute and the value. E.g. 'The block is red and square' could be regarded as an abbreviation for 'The block has colour red and shape square'. The abbreviations lose no information because the name of the attribute value unambiguously (at least in some contexts) identifies the relevant attribute.

However when forming questions if we simply start from the unexpanded abbreviation and create a gap by removing the predicate, the result could be too uninformative as regards what sort of gap filler is required. Thus going from the proposition 'the block is red' to the question 'what is the block?' (or what gap-filler makes the proposition 'the block is ...' true?) does not specify precisely enough the intended question better expressed as 'what colour is the block?', which imposes much stronger limits on what counts as an appropriate answer, insofar as it is equivalent to asking what gap-filler makes true 'The colour of the block is ...'.

Note that the ability to refer explicitly to the attribute by some sort of label (internal or external) need not occur first in development or in evolution. There could be simpler stages where the values of particular attributes are implicitly identified as being of a certain type because of how they are produced by sensory mechanisms and how they are used. Thus the output of a temperature detector fed into some sort of neural control system need not be labelled as being a value of temperature, because the output is connected only to parts of the system that make use of temperature information. So in that context, simple predication without any possibility of expansion to the attribute value form is all that can occur and it is all that is needed. Only when more sophisticated architectures evolve, or develop, performing more complex functions, is it necessary or possible to use explicit attribute labels.
(**NOTE:** Explicitness here has nothing to do with consciousness: it is merely a matter of whether some structure or process exists which has a certain sort of function in the system.)

4.5 Object relations

Physical objects, locations, routes and other things can have relations to all sorts of things, of the same type or different type. The relations differ in ‘arity’, depending on how many things are involved. E.g. they may be binary, ternary, or n-ary for any integer n.

A major distinction can be made between relations based on attribute values (e.g. X is taller than Y, means the value of X’s height attribute is larger than the value of Y’s height attribute), and likewise for other relations of ordering, e.g. X is between Y and Z in height, relations involving measures, e.g. X is 14 cm taller than Y, etc. The types of attribute-based relations involving attribute A that are possible will depend on the structure of the set of possible values for attribute A, e.g. whether it is totally or partially ordered, whether it is continuous, dense, discrete, etc., whether the values are themselves structured (e.g. vectors or trees or nets), and so on. For example, one object can be the mirror image of another if the shape of the first and the shape of the second are essentially the same except for a single reflection, so that together the objects can form a symmetric structure. Another kind of relationship that can hold between X and Y depending on shape is X being isomorphic with a part of Y, e.g. if X is a sphere and Y consists of two spheres joined by a cylindrical rod.

There are other relations that are more complex and are not derivable from or definable in terms of the intrinsic properties of the objects but depend on where they are located in space and or time. We’ll call those extrinsic relationships, and return to them later.

4.6 Intrinsic relations: Relations based on attribute values

Some of the relations involve comparisons of attribute values. There are two main sub-cases, namely equality and inequality of values, and the inequality sub-cases can be further sub-divided, depending on the structure of the set of possible values, e.g. whether it is totally ordered, whether the total ordering has a kind of asymmetry supporting a notion of more or less, whether it is a continuum of values or not, whether the values vary in different dimensions so that they form some sort of vector space, etc.

For example, if attribute A has possible values Av1, Av2, ... then propositions can be expressed regarding two objects X and Y saying

- ‘X is the same A as Y’ (e.g. X is the same colour/height/size as Y),
  meaning something like: the value that fills the gap in ‘The A of X is ...’ and the value that fills the gap in ‘the A of Y is ...’ are the same thing.

  Variants of this use a relation weaker than having identical values, and instead specify values that are close in the ordering, which could be expressed as ‘X is like Y in respect of A’, or using phrases like ‘similar to’ ‘close to’, etc.

- ‘X is more A than Y’, or ‘X’s A is exceeds Y’s A’ (e.g. X’s size, height, hardness, smoothness, exceeds Y’s)
  meaning something like: the value that fills the gap in ‘The A of X is ...’ is ‘bigger’ (or ‘higher’) than the value that fills the gap in ‘the A of Y is ...’.

In many natural languages there are alternative constructs for expressing this sort of idea, some of which use the attribute name with a modifier, e.g. ‘more height’, ‘less weight’, some of which use a comparative form of an attribute value, e.g. bigger, weightier, smaller. As usual we are not concerned with the precise syntax but with what might be expressed.

Although most of the examples given involve relations between physical objects there are also relations between events and processes that are intrinsic. For instance one of the attributes of a process is its duration. So one process can have a longer or shorter duration than another, or a duration that differs by 2 seconds or 2 days, etc. One process may involve something happening faster than another, e.g. rotation, or colour changes, or speed of motion.

The examples given so far of relations based on attribute values are all binary. However insofar as as attribute values are ordered, and many different relations can exist between items in an ordered set, there will be derived non-binary relations between objects with those attributes. For instance

- Ternary relations
  For example: X is between Y and Z in A (e.g. in height, in colour, in size) means something like the attribute of type A of X is between the attributes of type A of Y and Z (in the ordering of attribute values of type A). This could be true if A is height, and the attribute values of X, Y and Z are respectively 6cm, 3cm and 22cm, or, 6cm, 22cm and 3cm (since the proposition says nothing about the relative position of Y and Z). There are many other Ternary relations, such as that X and Y differ from Z in attribute A by the same amount.

- Quaternary relations
  As with ternary relations many quaternary relations can be formulated on the basis of attribute values, for example the proposition that W differs from X in A more than/the same amount as/less than Y differs from Z in A.

Exactly which sorts of n-ary relations can exist between objects with attributes of type A will depend on the structure of the space of values of A. If there is a small discrete set of values, e.g. small, medium and large, then far fewer relationships will be possible than if there are many values, or if the values form a continuum (or dense set).

If the set of attribute values has some sort of distance metric, then comparisons of distances can be used to construct yet more attribute-based relationships, which are not necessarily elegantly expressible in English. Are some languages more suited to saying this sort of thing than others? Examples are:

- X is more tall than Y by a larger amount than Z is more tall than Y.

- W is more red than X by the same amount as Y is more red than Z.

In the case of some qualitative attributes, such as colour, taste and smell, we find it useful to describe things in terms of closeness to particular important or interesting attributes, for instance describing X as more red than Y (redder) or more sweet than Y (sweeter). In contrast we don't say of X's height that it is more 10cm than Y, as a way of saying that X's height is closer to 10cm than Y's height. However in principle there is no reason why a particular language should not have such syntax added to it, if that would prove useful, e.g. because some heights have particular social, religious or economic importance, for certain objects.

All the types of propositions expressing relations derived from ordering or other relations between attribute values allow the removal of components leaving gaps on which question-forming transformations can operate. For example, introducing different gaps in the proposition *'X is between Y and Z'* in A, could produce such questions as

- What is between Y and Z in colour? (one gap)
- In what respect is X between Y and Z? (one gap)
- What is X between in colour? (two gaps)

4.7 Relations based on shape

The examples so far have used attributes whose values are essentially points in a space of possible values, where the points themselves (e.g. heights, widths, colours(?)) have no structure, though the space may be wholly or partially ordered, discrete or not, finite or not, and so on. One particular set of attributes has elements which instead of being mere points in a space (or even vectors in a fixed dimensional space) are structures of varying complexity, namely shapes. Two objects can have shapes that in turn can have attributes of varying complexity, for instance, number of edges, number of holes, relative lengths of holes and edges, being convex or concave all over or in parts of the surface, having grooves, having dents and many more.

The precise set of shapes that any individual can think about and ask questions about will vary from species to species as well as from individual to individual within a species. Of particular interest in robotics and (human or animal) psychology are the relationships between shapes of objects and the actions that individuals can perform. Whatever set of names we have for shapes will never be enough because shapes can be made more and more complex indefinitely by adding components, replacing components, joining two or more shapes in various ways to make a more complex shape, adding grooves, hollows, bumps, holes, and so on. The fact that such transformations produce shapes that cannot be recognised does not imply that they cannot be seen. On the contrary, seeing shapes needs to come before recognising them when they are re-encountered.

The point of all that is that it is a mistake to think of perceptual systems as producing only propositions about features and relationships of recognised (or named) whole objects. Rather there will be many object fragments and surface fragments that are seen, and relations between them that are seen, including not only relations within images, which don't concern us for the moment, but also relations within the 3-D environment.

Thus if there are propositions expressing what is seen they may be in part like parse trees summarising a network of relationships between simple or complex fragments, which may or may not be recognised fragments.

The process of question formation through gap manipulation then will have to operate on these structures, and the results may not be easily expressible in familiar human language. For example it may be possible to wonder whether a partly visible object or fragment has a shape for which there is no label but is (internally) identified by the perceiver as 'that shape', referring to the shape of a wholly visible component of the scene.

Similar comments apply to perceived events and processes, such as the motion of an object. For instance a perceiver may wonder what kind of action that it can produce could cause X to move as Y is seen to move, where there is no name or prior label for the latter type of motion. (E.g. think of a dancing student wondering how to replicate a teacher's actions.)

4.8 Extrinsic relations: Spatio-temporal relations

So far we have considered intrinsic relations based on attribute values, including relations like being more or less A, where A is an attribute, or having an attribute value between other attribute values. Among these are relations like being the same shape or having a shape with fewer concavities, that depend on the shapes of the objects standing in the relations.

There are also *extrinsic* relations that depend not only on the attributes of the objects but also where they are, and how they are oriented in space, and in the case of events and processes some some of the relations depend on temporal location, i.e. when they occur, when they start, when they end, etc.

There are also very many relations that depend on spatial location, orientation, and shapes of objects. Some of these involve metrical relationships, such as being similar in shape, or being a certain distance apart, or being adjacent, whereas others can be thought of as purely topological, such as being inside, or being linked (as two rings can be). Others are a mixture.

For instance when a hook and a ring are joined up as in a toy train where one truck pulls another, it is possible to think of the relation between hook and ring that makes pulling happen as metrical rather than topological because bending the hook so that it is lifted out of the ring and then makes pulling impossible will change only metrical properties, not topological properties. However if we think of the hook as part of a ring with a missing part then we can think of the hook and its missing part as linked to the ring, a topological relationship.[11]

4.9 Multi-strand relations and multi-relation facts

Although two point-like objects may have a small set of relations of distance and direction (which can be combined into one relation vector) when two objects have complex structures they can have multi-faceted relations because the relations between the two objects are composed of many relations between their parts. For example if a small cube (SC) rests on a large cube (LC) we can have relations like the following between the objects and their parts:

- SC is immediately above LC
- bottom face of SC and upper face of LC face in opposite directions
- bottom face of SC is in contact with upper face of LC
- Edge E3 on bottom face of SC coincides with part of Edge E17 of top face of LC
- Vertex V5 of SC coincides with vertex V6 of LC
- Face F4 of SC meets face F20 of LC perpendicularly

and many more. Thus a vision system that can merely report that SC is above LC is inadequate for tasks where multi-strand relationships are important, as in many assembly and manipulation tasks.

Many of the objects in the environment are complex objects composed of parts. It is often thought that complex objects are best described in terms of hierarchical decomposition yielding a tree-structured description where nodes are mostly linked by a ‘part-of’ relation. But what we actually see and what we often need to know is far more than the part-whole decomposition. A partial acknowledgement of this point is in this online ontology for bicycle parts, which besides the ‘part-of’ relation includes *constraints* between parts:

http://www.ksl.stanford.edu/htw/dme/thermal-kb-tour/bike-domain.lisp.html

Moreover, a *process* involving two structured objects moving in space will involve a collection of sub-processes where several relationships all change in a systematically constrained way. A robot observing and producing such processes needs to be able to represent and reason about *multi-strand changes in relations*. We could call these ‘multi-strand processes’. Multi-strand processes are, of course, familiar in physics and engineering, where the most highly developed form of representation is a collection of differential equations relating such things as positions, distances, volumes, currents, voltages, pressures, tensions, torsions, and their rates of change and possibly also higher order derivatives. But that form of representation does not capture the collection of

---

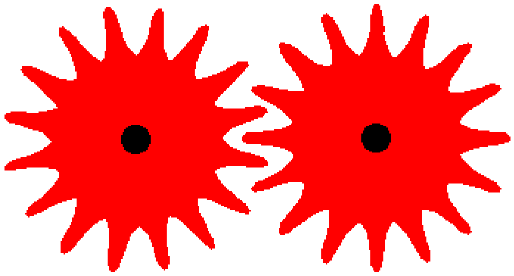
[11] The need to move between more or less metrical and topological relationships in solving problems is discussed in a little more detail in Sloman (1998) and in A.G. Cohn, S.M. Hazarika, ‘Qualitative Spatial Representation and Reasoning: An Overview’, (2001). http://www.comp.leeds.ac.uk/qsr/pub/funinfreview.ps.gz

changing relationships that are typically of interest in an environment in which an animal or robot perceives and acts. Verbal descriptions are also often clumsy and inadequate, which is one reason why we often prefer diagrams, working models, or complex gestures or demonstrations, for instance in teaching someone how to tie shoe-laces, or how to dress a child. Logical descriptions can handle arbitrary collections of relationships and changes of relationships but it is not clear what would be best for a robot involved in domestic manipulative tasks.

4.10 Multi-strand relationships and causality

This is deeply connected with our everyday notion of causation, especially the (Kantian) deterministic notion of causation, rather than the (Humean) correlational, probabilistic notion of causation that has recently received a great deal of attention in AI and psychology (e.g. in connection with Bayes nets).

For example, part of the notion of an object being *rigid* is that its parts always move in such a way as to maintain a collection of metrical relationships including distance, direction, parallelism and similarity. Understanding what this involves, combined with the notion of impenetrability gives a way of thinking about what happens if two rigid bodies are close together, constrained in some way, and one of them starts to move, such as two meshing gear wheels each constrained to rotate around an axle at its centre, as in the figure, where most people will easily see what must happen to the wheel on the right if the wheel on the left rotates clockwise, even though this is a complex collection of multi-strand processes in which new relationships come into being and then later cease to exist (e.g. contact between teeth of the gears).



Depending on the shapes of the wheels and their teeth there may be situations where rotation is seem to be impossible unless the wheels move apart. More subtle features of the changing relationship can give clues as to whether there will be friction and wear during the rotations, an example of longer-term causation. It is not clear whether existing forms of representation used in computers can easily be applied to the task supporting such causal reasoning.

There is a long history of attempts to develop symbolic forms of representation suited to multi-strand relationships, including semantic nets in the 1960s, Minsky's Frame Systems (Minsky, 1978), formalisms such as KL-ONE, object-oriented extensions to AI programming languages (such as FLAVORS, LOOPS and CLOS), and more recently systems such as CLASSIC using description logics[12]. These forms of representation are essentially all discrete and suited to reasoning about discrete relationships, differences, and changes, although in principle some of the numerical attributes represented may have real values. In parallel with these developments many people have attempted to understand the significance of visual forms of representation (see the references in

---

[12] Online tutorials on description logics are listed here http://dl.kr.org/courses.html

DR.02.01).

A complete survey of spatial relations, spatial structure, spatial changes, causation in spatial changes, and ways in which they can be represented in intelligent agents would require a book at least. However a few points are worth making about some of the kinds of variation a perceiver may need to take account of.

One of the requirements for a visual system in a robot observing complex motions of complex objects is that it should be able to represent combined translations and rotations of rigid, jointed and flexible objects whose parts move relative to one another while the whole object moves, including for instance a lion seeing a deer during a chase. Birds building nests by inserting one twig at a time in such a way as to form a new complex rigid object have related requirements. The processes a human-like domestic robot will have to be able to cope with, like those already coped with by many animals, involve a mixture of *continuous* and *discrete* changes (e.g. a cup approaches the rack in a dishwasher through empty space and then is in contact with the rack and constrained from moving downwards or horizontally, or rotating). Although there has been much psychological research on human abilities to visualise continuous changes, as far as I know there are no forms of representation used in AI that meet all these requirements, though it is possible that special cases can be handled by a mixture of discrete data-structures and techniques based on differential equations or spreading activation mechanisms in neural nets. It is possible that a survey of techniques used in systems for generating dynamic displays in CAD systems and computer games will reveal something relevant.[13]

One form of representation that looks particularly promising involves the use of *simulations*, where a collection of data-structures and programs provides a virtual machine running in the computer with structural relations that correspond in a useful way to those in the objects represented, and where the programmed constraints on the changes in the data-structures correspond to constraints in the physical environment. However, for many purposes a very detailed and precise simulation would be both impractical and overkill, e.g. for high level planning of actions and their consequences. It seems that investigating the use of varieties of simulations at different levels of abstraction could yield promising insights into ways of giving a robot something like human abilities to perceive, act in and think about the sort of domestic environment the robot Fido will have to deal with. We could start doing this for very simple table-top manipulations. I conjecture that the use of this sort of simulation process will turn out to be the basis of many observations based on introspection about the role of visual or spatial representations in reasoning and problem-solving. The actual processes are not spatial but self-monitoring mechanisms interpret them as spatial because that is part of the mechanism whereby they are usefully applied to reasoning about spatial processes.

4.11 Local vs global spaces .... to be extended

Some relations are relative to an object's intrinsic frame of reference, e.g. in front of, on the left side of, behind, when applied to something like a car, a horse, or a person. For objects that don't have an intrinsic front and back, such as a ball, a box or a banana, the very same words can be used to specify a relationship based on the perceiver's viewpoint, so that 'in front of' is interpreted

---

[13] A useful collection of ideas is summarised in Cognitive Models of Dynamic Geographic Phenomena and Their Representations: Report of a Specialist Meeting held under the auspices of the Varenius Project (October 28-31, 1998). Eds Stephen Hirtle, and Alan MacEachren. Online at http://www.ncgia.ucsb.edu/Publications/Varenius_Reports/Cognitive_Models.pdf

as equivalent to something like ‘nearer to me’. It is also possible to refer to front-back relations between things that don’t have a front-back distinction, by referring to another object that does, e.g.

‘The ball was in front of the horse and the box in front of the ball’

For both the PlayMate and the Explorer it is important that the visibility of objects can change without their existence being accepted (Piaget on object conservation.) E.g. things on the table that are temporarily out of view because they are occluded, or because you are looking in a different direction can still be remembered. A video of a three year old child playing on the floor with a toy train set that surrounds him illustrates this very well

http://www.cs.bham.ac.uk/˜axs/fig/josh_tunnel.mpg [5MB]
http://www.cs.bham.ac.uk/˜axs/fig/josh_tunnel_big.mpg [15MB]

The retinoid mechanism, and related mechanisms described in Trehub (1991) were designed to explain this, along with other phenomena related to the fact that visual information persists across saccades, though it is not clear that it is adequate to the task.

Moreover, there are big differences in the representation of local space (i.e. immediately surrounding the viewer) and the representation of more global terrain, such as one’s house, village, or country. The local space seems to be at least partly based on a viewer-centered form or representation (which therefore has to change as the agent moves around), and is rich in 3-D information whereas the global space is more persistent and seems to be (at least for humans, though perhaps not for flying animals) largely 2-D, as is the case for most computer models of way-finding and learning about locations.

Some of the local and global spatial representations make use of frames of reference defined by specific objects, including the agent’s own body, another agent’s body, part of a body (e.g. an arm), the room, the table, the toy aeroplane on the table, etc. Some objects may simultaneously be represented as occupying locations within different reference frames. E.g. fingers have locations relative to the hand, the arm, and the whole animal, as well as locations in the room. Which ones are important and need to be made explicit can change dynamically according to task requirements. For instance relations between finger and nose are normally not relevant, yet they become relevant when someone is seen as scratching his nose, or an agent needs to scratch its nose.

This argues against assuming that there is one hierarchical part-whole model for each complex object from which all information is derived as needed (e.g. the Marr/Nishihara model based on connected generalised cylinders, and Hogg’s and Mackworth’s PhD theses.(???))

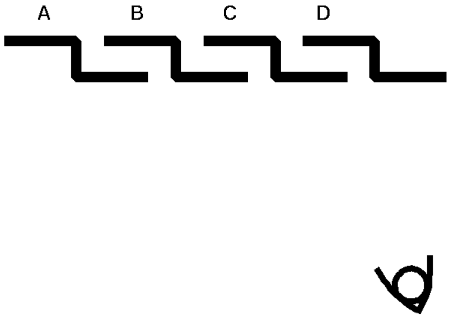
4.12 Integration, zooming, etc.

If a robot uses multiple forms of representation, that raises questions about how they are integrated. In principle there be only a collection of completely separate task-specific representations used by different sub-modules in the architecture which perform different functions. (This is one of the themes of Minsky (1987).) However, it is clear that humans (at least adult humans in our culture) have an ontology that somehow combines the local 3-D and large scale 2-D information within larger scale 3-D (or more accurately large scale 4-D space-time) — information structures, that allow zooming in and out according to the needs of the task. Recently such zooming operations, which in the past we could perform only in our heads, have been made available on computers, e.g. in CAD systems, but at present such systems do not share our understanding of what they are doing, and the representations are useful only for generating images or controlling specific actions, such as parts manufacture or checking design constraints. E.g. they do not, as far as I know, support understanding of how a small scale 3-D structure (e.g. a toy house) can provide information about a larger structure (the real house) that might be useful for route planning.[14]

4.13 Indeterminacy of spatial concepts

Many of the relationships for which we have names, like the ones mentioned above also have a kind of indeterminacy as to what regions are involved: e.g. if you look at a box in the middle of the table, and consider under what conditions you would describe a ball placed anywhere on the table as being to the left of the box, there may be considerable regions of indeterminacy. It is sometimes thought that such cases reflect statistical concepts with a probability distribution. An alternative view is that they are actually higher order concepts with an implicit reference to the *purpose* or *task* for which the relationship is being described. If there is no specific purpose or task the question about where the boundaries between the regions (left, right, front, back, etc.) are is essentially pointless, like asking whether a specific rock is or is not big without having any criterion for things being big enough. If forced to answer such pointless questions (e.g. in a psychology experiment asking which locations are to the left of that box) people may then use a default rule that covers many of the normal reasons for mentioning the 'left of' relation.

Instead of postulating a probability distribution it may be more accurate to regard this as a case where there is no answer, and therefore the statement 'X is to the left of Y' in such a case has no truth value, like the statement 'My lawnmower is better than my car' taken out of any context specifying the relevant respects of comparison.



Many spatial relations involve an order which depends on location or a combination of location and other attributes. For instance things may be ordered according to their distance from some reference location (e.g. the viewer) or according to what partially occludes what. 'X occludes Y' often implies 'Y is further away than X' but need not where complex shapes are involved. For instance, when the figure is viewed from somewhere below D, A occludes B, B occludes C, C occludes D, but A may be furthest from the viewer and D nearest.

4.14 Affordance-based relations – and embodiment

For a perceiver that is capable of manipulating objects, many spatial relations are concerned with what motions are possible or constrained. For example saying that there is a gap between X and Y could in some contexts refer to the possibility of some object currently under consideration being able to move between X and Y. There will also be many relations for which we don't have verbal

---

[14] In that respect they are like the two-year old children described in a recent article 'Mindful of Symbols', Author Judy S DeLoache Scientific American, August 2005, pp 60–65.
http://www.sciam.com/article.cfm?articleID=000ACE3F-007E-12DC-807E83414B7F0000&sc=I100322

names but which an expert manipulator learns to perceive and think about as relations that facilitate or obstruct actions, or which change during actions.

For example, when a child lifts a cut-out picture of a car from its ‘home’ in a flat sheet of wood, and then attempts to replace it, she may be unaware of the requirement to line up the boundary of the car shape with the boundary of the hole it left behind, and merely think of putting it back in the same general location. Such a child can be seen to try to get the piece back into its hole merely by pressing hard on it.

At a later stage the child has learnt about the significance of the additional spatial relationship, and appreciates the need to do some sliding and rotating as well as pressing, in order to get the shape back in place. At that stage the child need not have any *name* for either the relationship of alignment or coincidence between corresponding parts of the outline of the car and of the hole, nor for the changing relationships that occur while it is being moved around in order to get it into position to drop into place. Yet the child may understand the relationships, may be able to think about what sort of misalignment exists (e.g. where there is a boundary mismatch) may be aware of information gaps, may change its viewpoint, or move its hand out of the line of sight, in order to obtain missing information and then use the newly acquired information in order to complete the task. Thus there is an intermediate stage in which a question is considered, for which no linguistic expression is available (at least not to the child), and an answer found, for which there is also no linguistic expression available.

Similar remarks can be made about tool manipulation and manipulation of items of food or nest-building materials in other animals, e.g. squirrels, monkeys, chimpanzees, crows, etc., none of which can use anything like a human language.

4.15 Affordance-based and mathematical relations

In ways that remain unexplained, a learner seems to be able eventually to understand that there are spatial relations that may have originally been learnt about through their relevance to action but which can be described in a manner that is independent of how they are perceived or used by acting animals or robots but can be expressed in an abstract manner in terms of their own intrinsic properties and relations. The invention of Euclidean geometry as axiomatised first informally a few thousand years ago, and then more formally in the 20th century, was perhaps one of the most important such discoveries in the history of human thought, and it is not clear whether any other animals have such an affordance-free conception of space and spatial relations.

One of the tasks for CoSy and related projects in the longer term is to explain how that conceptual change can occur within an individual: what sorts of forms of representation, mechanisms, architectures are needed.

4.16 Self-knowledge

Nothing said here about needing, acquiring and using information of a particular type implies that the child or the animal *knows* what it is doing. That kind of self knowledge requires additional architectural resources and meta-semantic capabilities supporting reflective self-description. It is possible that many animals that are very good at thinking about and solving problems are totally incapable of thinking about or even being aware of their own problem-solving. This is probably true of much of what happens in young children: the meta-management architectural layer required for self-understanding does not seem to develop until some time after many other capabilities have developed. In general, knowing X, and knowing that you know X require quite different mechanisms, including different forms of representation whose semantic contents presuppose very different ontologies. The latter requires a meta-semantic ontology which includes things that manipulate information, whereas the former does not.

---

5 Further details [to be reorganised]

Some comparative relations can be used as ternary relations, e.g. ‘X is further from A than from B’, ‘X is further from A than Y is’

Other things to sort out

- further
- moreleft
- moreright
- moreinfront
- morebehind
- derived (parametrised) comparative relations: closerin (colour, height, size, distance, ...)

Similar kids of ternary relations can be generated from features that allow ‘more’ or less’, e.g more red, more high, more fast

more kinds of relations

- QUATERNARY
  - W is as far from (close to, higher than) X, as Y is (to) Z
- GEOMETRICAL RELATIONS
  These can involve arbitrarily many objects. I.e. the relations can be
  - unary (is a triangle),
  - binary (contains)
  - ternary
    - form an equilateral triangle
    - form an isosceles triangle
    - are collinear
    - form a acute/obtuse/right angle (X Y Z)

  etc.

OTHER PROPOSITIONAL FORMS

- x is P
- x is at X
- x is a K
- morethan (X’s Y’s)

- Existentially quantified
- Universally quantified
- Most (Xs are Ys)
- More (Xs are P than Ys are)
- Numberof Xs is N
- Numberof Xs has some numerical property.
  (E.g. even, graspable simultaneously, beyond our budget, affordable, ...)

6 What about a non-linguistic (pre-linguistic) agent?

A significant subset (what subset?) of the above proposition types and question types may be capable of being considered by animals without human language and pre-linguistic children. The questions will not be posed externally, but internally, and will determine goals for cognitive and other processes.

So that raises the question of what sort of formalism could do these things *within* a cognitive agent, as opposed to *between* speakers?

It may be that answering that question will give us deeper insights into what goes on in speakers, for they will presumably, to some extent, build on the pre-verbal capabilities that evolved earlier.

Later on the internal processes were certainly expanded by the availability of linguistic constructs – e.g. things like

> 'By how much is the distance from A to B bigger than the distance from C to D?'

Could a pre-linguistic animal (or child) wonder about such a question?

Some of these issues are posed in relation to images of a cup, saucer and spoon in this short presentation: http://www.cs.bham.ac.uk/research/cogaff/challenge.pdf [15]

7 Representations in reactive systems: speed, and fluency using implicit representations

In addition to the pre-linguistic cognitive mechanisms we need to refer to what are sometimes referred to as *sub-cognitive mechanisms*, a notion that overlaps with, or may be equivalent to some uses of, the notion of *reactive mechanisms*.

The spatial competence of an octopus is amazing. Many animals, including insects and and other invertebrates, and perhaps many vertebrates, such as fish, amphibians, reptiles, many birds and some of the less intelligent mammals all seem to have considerable spatial competence yet probably lack the forms of representation and the ontologies described here. (It's hard to be absolutely sure about exactly what forms of information-processing they do and do not have, but let's ignore that question for now.)

For example bees land on flowers suitably positioned in order to get nectar and pollen, they find their way back to the hive, they can later return to the same location, and they can even communicate direction and distance to other bees by doing their famous dance. Spiders make intricate webs and when insects get caught in the web a spider can take appropriate action. Many animals find and eat

---

[15] Elaborated in CoSy Deliverable DR.02.01 on Requirements for representation

food which may be animate and try to escape or inanimate and have to be removed from whatever it is attached to, e.g. grass, berries, leaves, nuts, fruit, etc. Many animals make nests and termites make elaborate cathedrals. There are also complex spatial actions involved in mating, and egg-laying.

For many kinds of actions behaviour may be controlled by information that is implicit in the changing patterns of activation of various parts of a control system that includes sensors, effectors, various kinds of feedback and feed-forward mechanisms, and may even be hierarchically layers. In such a system actions and the consequences of actions are represented only while subsystems are active in producing and controlling those actions. Likewise the contents of sensors are represented only while the sensors are being stimulated. Past actions and percepts are not representable, though there may be gradual changes produced by hysteresis and and adaptation mechanisms, as in neural nets that change synaptic weights over time.

These *implicit* means of encoding information in the states of dynamical systems can be very powerful and are probably the only form of representation available to the vast majority of biological species. However[16] such forms of representation have major limitations and evolution seems to have discovered, in at least some species, additional more explicit forms of representation that support such competences as thinking about unobserved objects or spatial regions, predicting sequences of future events, explaining observed phenomena in terms of unobserved phenomena, making structured plans, having motives that refer to the remote future, and even thinking about thinking. These are often thought to be what distinguishes cognitive systems (though some proponents of dynamical systems claim these are just a special case, which is either trivially true because everything that works is a dynamical system, or false if dynamical systems are characterised on the basis of models from 19th century physics based on differential equations).

However, for the purposes of CoSy, if robots are to be capable of performing fast fluent actions such as rapidly and smoothly picking up a cup of tea and passing it to a person, then many of the details of such a process will probably need to be controlled by an insect-like reactive mechanism with many feedback loops and continuous (analog) online control systems, unlike a typical plan execution mechanism.

It appears that in humans, and possibly other animals, there are both sorts of systems and they are closely integrated in complex ways that are not yet understood, but one of them could involve the more symbolic, cognitive, deliberative system making use of explicit representations *training* a reactive, dynamical system that uses only, or mostly, implicit information structures. Repeating actions under the control of a deliberative mechanism seems to be a common means of developing speed, accuracy and fluency in many human activities, including running, jumping, grasping, throwing, many athletic skills, musical skills, linguistic skills and skills at operating complex machines, including driving bicycles and cars. What is not so generally appreciated is that similar extensions of competence can occur in low-level perceptual mechanism when trained in various repetitive tasks under the control of deliberative mechanisms. Examples include learning to read text fast, learning to read music, learning to 'read' the behaviour of opponents in various sporting activities, and perhaps learning to read facial expression, posture and gestures in many forms of human interaction.

Another kind of relation between the two sorts of mechanisms can be the development of reactive mechanisms that generate new motives or even alarm signals for the deliberative system so that its focus of activity can be switched or its actions modulated when something important has changed in the environment (or in the body).

---

[16] As discussed in an online draft, incomplete paper, http://www.cs.bham.ac.uk/research/cogaff/sloman-vis-affordances.pdf

If all that is right, then it may also be a long term requirement for CoSy robots that they too should have both reactive sub-symbolic and deliberative explicitly symbolic mechanisms if they are to be able to function as domestic helpers for the aged or infirm, for example.

(An alternative possibility is that speeds of computation will increase so fast that extremely fast cognitive mechanisms constantly planning and re-planning at speeds that are impossible in brains will be able to take over all the functions that in humans, apes, etc. are performed by evolutionary very old reactive mechanisms. There are reasons by physical impossibilities, including power-weight and other constraints, may rule that out, but the issue will not be discussed further here. At least we seem to have an existence proof of the effectiveness of a hybrid system, even if we have no proof of uniqueness or optimality.)

An example of an attempt to produce a hybrid symbolic/neural system meeting some of these requirements can be found in Ron Sun's CLARION project http://www.cogsci.rpi.edu/˜rsun/clarion.html Discussion of how well it meets the requirements specified here will be left to another occasion. It is discussed briefly Sloman and Chrisley (2005)[17].

8 Nature Nurture Issues

Suppose it is true that at least some natural language users are capable of understanding the full variety of types of propositions and questions sketched here, in at least some sense of "understanding" involving knowing what sort of thing counts as a correct answer, without necessarily knowing the answer or even known how to find or verify the answer – because that might require technology that has not yet been invented, or in some cases might require use of a theory that has not yet been discovered). This raises the question how such understanding arises?

One possible answer is that it arises out of the use of the sorts of mechanisms sketched in a paper presented at IJCAI05[18] discussing the precocial/altricial spectrum, where altricial species (or altricial skills) instead of being restricted to

(a) learning by the use of positive and negative reinforcement mechanisms that gradually transform weights in some sort of statistical mechanism (e.g. a neural net) can also be learnt by
(b) a combination of more 'symbolic' mechanisms:

- exploring (randomly or otherwise) effects of many kinds of distinct sensory patterns and action types (starting with innate ones)
- noticing, categorising, and inventing (internal) labels for 'interesting' cases
- using some (innate? learnt?) 'syntactic' mechanism for internally recombining the labels in various ways to generate new, larger perceptual and action structures, which are also explored and if found interesting stored, labelled, and made available for re-use

Such mechanisms can support discrete, creative learning and discovery steps producing quite large changes in competence.
(The idea goes back at least to work by Oliver Selfridge, as reported in this toy demonstration http://www.cs.bham.ac.uk/research/poplog/teach/finger[19] which can be run using the poplog system.[20]

---

[17] Available online here http://www.cs.bham.ac.uk/research/cogaff/04.html#cogsys

[18] Co-authored with biology Jackie Chappell, http://www.cs.bham.ac.uk/research/cogaff/altricial-precocial.pdf 'The altricial precocial spectrum for robots'

[19] See URL http://www.cs.bham.ac.uk/research/poplog/teach/finger

[20] See URL http://www.cs.bham.ac.uk/research/poplog/freepoplog.html

Defining that hypothesis about altricial mechanisms properly will have to wait for another occasion. For now it is important only that these mechanisms have implications both for the innate ontology of the robot or animal and for how the ontology develops over time.

9 Provisional Concluding Remarks

The study of ontologies is not merely a theoretical exercise for philosophers, even though it may first have been done by philosophers, as explained in the Appendix, below. Neither is its relevance to AI and software engineering restricted to processes of reasoning and communicating, the purposes for which ontologies have most often been sudied in AI. In the context of our research ontologies are relevant to perception and action insofar as what a robot or animal can perceive and what it can do are both limited by the ontology (or ontologies if there are several) available to it. It may be that an ontology is implicit in the forms of representation, kinds of variation, and kinds of use to which its information structures can be put. Or it may be explicit on some meta-semantic information structure of the kind mentioned in Section 4.16. In both cases the "application-ontology" used by the animal or machineis not the same thing as the "design-ontology" used by designers and scientists who think about it, though the two may be closely related.

By analysing in great detail various more or less ambitious scenarios involving exploration of the environment and manipulation of objects[21] we can work out very specific requirements for kinds of information content that will be required, analyse the ways in which the information will be used in various sub-systems within the whole architecture, and from there proceed to investigate options and tradeoffs between different solutions. Different forms of representation for the same information may be needed for different purposes, especially in different parts of the architecture. For instance the task of thinking hypothetically about what would happen if something were done is different from the task of making it happen and the requirement for the latter to produce signals controlling muscles or motors is different from the task of analysing presuppositions, implications, costs and benefits of an action. For some of the detailed sub-tasks we may have to invent forms of representation that are more powerful than any currently in use.

A conjecture arising from our analysis, especially our analysis in section 4.9 of the need to represent multi-strand processes in which there are continuous and discrete changes in multi-strand relationships, is that a perceptual system may need to use representations of changing spatial structures of various kinds (including structures with multi-strand relationships), within *simulations of multi-strand processes in the environment*, at different levels of abstraction, running concurrently in registration with one another and with sensory input (and in some cases with motor output). Some of the processes need to be continuous (with different resolution levels) and some discrete. It is possible that new rich planning formalisms for representing multi-agent plans with conditionals and loops can be interpreted as programs for running discrete high level simulations. All of this can be seen as a generalisation of a notion proposed by the late Max Clowes that "perception is controlled hallucination", which we now expand as "multiple, concurrent, synchronised hallucination".

This is not an entirely new idea: it builds on ideas also found, for example in Craik (1943) and more recently in Berthoz (2000) and Grush (2004), among many others. We are not aware of any earlier presentation of the "multiple-concurrent simulation hypothesis", however, which links together many past problems and proposed solutions.

---

[21] As explained in http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0503 and in http://www.cs.bham.ac.uk/research/cogaff/gc/targets.html

We conjecture that if such mechanisms can be made to work, they will add powerful new capabilities for use in perceiving, reasoning about, and acting in a richly structured environment, including both reasoning about the actions of others and planning or analysing the robot's own actions. A simulation mechanism with different levels of abstraction is potentially a powerful basis for discovering affordances, which are concerned with 'what would happen if'.

Implementations of such ideas will be very difficult, and will ultimately stress both neuroscience and AI. Such implementations, if they are ever produced, can be tested by showing how, as more and more features of the design are added, more and more complex scenarios become demonstrable, in a partially ordered collection of scenarios, ordered by various kinds of difficulty[22]. There are also bound to be empirical conjectures for biology, neuroscience, psychology and linguistics that come out of such theories, and which might be testable empirically in those fields.

---

[22] The proposed methodology based on a partially ordered network of scenarios is described in the notes for the IJCAI05 tutorial http://www.cs.bham.ac.uk/research/projects/cosy/papers#tr0503

10 Some References on ontologies

These references were found with the help of Google. The list is neither complete nor authoritative nor representative. It's just what I found in a fairly short time spent searching, with some help from Push Singh.

Pylyshyn, Z.W., 'The role of location indexes in spatial perception: A sketch of the FINST spatial-index model,' *Cognition* 32, 65-97, 1989.
http://ruccs.rutgers.edu/ftp/pub/papers/cognit89.pdf[23]

See OpenCyc on 'Spatial Relations'
http://www.cyc.com/cycdoc/vocab/spatial-vocab.html[24]

> This document describes collections, predicates and other Cyc constants that are used to represent spatial objects and relations. See also documents for Groups, Quantities, Movement, Paths & Trajectories, Parts Of Objects, and Geography.

The 'Simple bicycle design ontology' referenced in Section 4.9 is part of the project 'Model-Based Support of Distributed Collaborative Design' (Richard Fikes, Edward Feigenbaum, Sheila McIlraith, Robert Engelmore Todd Neller, Liang Zhu. Yumi Iwasaki, James Rice) at Stanford, described here: http://www.ksl.stanford.edu/htw/

A.G. Cohn, S.M. Hazarika, Qualitative Spatial Representation and Reasoning: An Overview (2001), http://citeseer.ist.psu.edu/cohn01qualitative.html

Brandon Bennett, Modal Logics for Qualitative Spatial Reasoning (1996), http://citeseer.ist.psu.edu/bennett96m

Jochen Renz, Bernhard Nebel, Efficient Methods for Qualitative Spatial Reasoning (2001), http://citeseer.ist.psu.edu/renz01efficient.html

References

Berthoz, A. (2000). *The Brain's sense of movement*. Perspectives in Cognitive Science. Harvard University Press, London, UK.

Brachman, R. and Levesque, H., editors (1985). *Readings in knowledge representation*. Morgan Kaufmann, Los Altos, California.

Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press, London, New York.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27:377–442.

Hayes, P. (1985). The second naive physics manifesto. pages 1–36. Ablex, Norwood, NJ. Also in Brachman and Levesque (1985), pp. 468–485.

Minsky, M. L. (1978). A framework for representing knowledge. In Winston, P. H., editor, *The psychology of computer vision*, pages 211–277. McGraw-Hill, New York.

Minsky, M. L. (1987). *The Society of Mind*. William Heinemann Ltd., London.

---

[23] See URL http://ruccs.rutgers.edu/ftp/pub/papers/cognit89.pdf

[24] See URL http://www.cyc.com/cycdoc/vocab/spatial-vocab.html

Sloman, A. (1992). Prolegomena to a theory of communication and affect. In Ortony, A., Slack, J., and Stock, O., editors, *Communication from an Artificial Intelligence Perspective: Theoretical and Applied Issues*, pages 229–260. Springer, Heidelberg, Germany.

Sloman, A. (1998). Diagrams in the mind. In *in Proceedings TwD98 (Thinking with Diagrams: Is there a Science of Diagrams?)*, pages 1–9, Aberystwyth.

Sloman, A. and Chrisley, R. L. (2005). More things than are dreamt of in your biology: Information-processing in biologically-inspired robots. *Cognitive Systems Research*, 6(2):145–174.

Trehub, A. (1991). *The Cognitive Brain.* MIT Press, Cambridge, MA.

11 Appendix A: Meanings of ‘ontology’ and some history

(These notes arose out of email discussions in the CoSy project during the first half of 2005).

The multiplicity of the meanings of the word ‘ontology’ can generate confusion. There’s no *right* meaning. This appendix presents a ‘potted’ history (very much over-simplified) to draw out a few main types of use of the word ‘ontology’, showing (crudely) their conceptual and historical relationship. I end with the question whether CoSy needs ontology tools.

11.1 Meanings of the word ‘ontology’

A: The oldest sense of the word, going back to Aristotle and perhaps earlier, as indicated by the suffix ‘ology’, is as a name for a subject of study or expertise or discussion (Greek: ‘logos’ = ‘word’, ‘meaning’, ‘thought’....). In that sense ontology is a branch of philosophy close to metaphysics (in the same way as these are names of subjects of study or areas of expertise: zoology, biology, geology, philology, oncology etc.)

In that sense ‘ontology’ refers to an investigation of the nature of being, what exists, why there is anything rather than nothing, constraints on possible worlds (e.g. could causation exist without space and time?), and perhaps whether what exists could be improved on or is the best of all possible worlds, etc. etc. Our project will not need to spend much time on doing ontology in that old sense.

Note that in that sense ‘ontology’ is not a count noun (e.g. it has no plural and you can’t easily refer to ‘an ontology’ or ‘different ontologies’).

However, like other ‘ologies’ the word developed different uses including uses as a count noun. E.g. ‘geology’ can be a name for an area of investigation, but can also refer to features of a part of the world, e.g. the Alps and the Himalayas can have different geologies. The word ‘ecology’ has also developed that way quite recently (damaging the ecology of X isn’t damaging the science but damaging the features of X described by ecology the science).

B: Recent philosophical usage of the word ‘ontology’ changed, to refer to specific conceptual frameworks. (I think this was a 20th century change, but I may be misinformed.)

This change was partly inspired by Strawson's 1959 book *Individuals: an essay in descriptive metaphysics*, and also by the discovery by anthropologists and others (developmental psychologists?) that different people have different views about what exists or can exist e.g. some do and some don't include souls of dead ancestors, tree spirits, transfinite ordinals, uncountable infinities, quarks, etc.). I.e. they have different ontologies. Strawson distinguished

- *Descriptive metaphysics*: the task of expounding the conceptual/metaphysical framework *actually* in use by some community or by various communities (we could add, or various individuals at different stages of development, or various species)

  from

- *Revisionary metaphysics*: the task of arguing about which is the *correct* framework, usually including the claim that we've got it wrong so far.

The products of descriptive metaphysics are offered as accounts of ontologies actually in use, not as accounts of how things have to be.

In this new usage (sense B) the word 'ontology' became a count noun referring (roughly) to the most general conceptual assumptions underlying a specific system of beliefs.

In that sense there could be *different* ontologies, e.g. the ontology of the ancient Greeks, the ontology of Buddhists, a modern scientific ontology that includes genes, ecosystems, quarks, economic inflation, etc., an intuitionist or a platonist mathematical ontology, the capitalist economic ontology, etc.

This sort of ontology can be a complex structure including several sub-ontologies (as our current scientific ontology does).

This second sense (B) is the one that I have mostly been using recently in talking about the kind of ontology required for CoSy (as described in my previous message.)

C: Software-engineering uses of the word 'ontology' are much more recent, as a result of people in software engineering and AI coming to realise the following:

(a) Engineers designing complex systems need to think clearly about what sorts of things they are designing and what sorts of things their machines have to interact with, prevent, produce, maintain, etc., so that they need to think about and if possible make explicit the ontology they use as designers – the *design ontology*. This became increasingly important as engineering moved beyond systems that can be characterised by sets of differential equations and the like.

(b) Insofar as these complex systems process information, designers need to specify *what* kinds of information a system can be expected to process, which semantic contents it may have, how the information will be represented, how it will be manipulated and used, etc. So they started using 'ontology' to refer to the ontology used by their products, the *application ontology*

Because developing either kind of ontology (design or application ontology) can be a difficult and in some cases very complex process, it was soon realised that designers need tools, techniques, and formalisms to specify *what* kinds of information a system can be expected to process, which semantic contents it may have, how the information will be represented, how it will be manipulated etc. Some of the tools might also be used by the systems produced, e.g. ones that don't have a fixed ontology, but go on extending their ontology.

(Outside AI this was connected with the growth of object-oriented programming languages and so-called object-oriented design – both of which generated much confusion as well as useful programming techniques).

So, engineers started moving away from informal descriptions and started producing formally specified ontologies for themselves to use (i.e. design ontologies) and for their machines to use (i.e. application ontologies), where the two could, of course, overlap in some cases.

D: Ontologies as formal structures

Ontologies of both the above two types are often expressed informally using familiar human forms of representation such as natural language, diagrams of various sorts, tables, etc. But as part of the drive towards theoretical clarity and elegance, and the need for more automated tools for producing working systems there was also a move towards requiring ontologies, even when produced by humans, to use a specific formalism with well defined conventions.

So there is now a sub-community of researchers whose only experience of the word 'ontology' is in the context of that kind of science and engineering, so for them the word usually refers to a formal structure specifying an ontology. (E.g. they may say 'My ontology uses XML and requires a megabyte of filestore'.)

In parallel with all those developments, some people started producing more or less formal ontologies specific to their field of study or research, e.g. an ontology to describe the various stages, techniques, formalisms, products, processes, etc. involved in doing software engineering, or the ontology of biology. This might be called a *discipline meta-ontology*. Such meta-ontologies can be very important for teaching a discipline, or for preventing reinvention of wheels among researchers, etc.

That contrasted with producing ontologies for specific application domains, e.g. banking, weather forecasting, process control, etc.

11.2 Form vs content of some part of the world

Some ontologies are referred to as 'meta-models', because people sharing such an ontology (a very general set of concepts) can use it to produce different competing models of the same domain: the ontology merely specifies types of things that are conceptually possible, leaving open alternative theories specifying what actually exists and happens. This distinction (which is sometimes described as a distinction between the form of the world and its content) is very important in science, and often ignored in philosophy of science.[25]

It may be that one way to view the kind of ontological development that occurs within an individual of an altricial species learning about the environment through play and exploration is that it starts with a genetically determined meta-ontology which is used by mechanisms for generating hypothetical ontologies that can be tested by using them in playful exploration and experimentation. Of course some of those ontologies will be mistaken and that is one of the reasons why this process of learning and development needs to happen under the watchful gaze of more experienced individuals.

---

[25] This point is developed at greater length in Chapter 2 of *The computer revolution in philosophy* http://www.cs.bham.ac.uk/research/cogaff/crp/chap2.html

11.3 Ontology Tools

Because the tasks can be very difficult, especially for people not well trained in philosophy, software engineers and AI researchers have started developing (and re-inventing) more and more sophisticated methodologies and tools to aid the process of ontology development especially collaborative development of domain ontologies.

However, most of them (as far as I know) tend to assume a uniform representation (e.g. some kind of logic) and, as pointed out in our email discussions, that may be too restrictive for a project like CoSy.

If we don't find suitably general tools that already exist we may need to develop our own tools for producing, maintaining and using ontologies suited to multi-level multi-functional robot architectures. In part this should also be a by-product of designing new kinds of learning mechanisms that develop or extend ontologies.

11.4 Further information

A lot more information relevant to this discussion can be found by giving search engines phrases like

ontology "software engineering"
or
ontology engineering