

Possibilities between form and function

Or between shape and affordances.

Aaron Sloman

School of Computer Science, University of Birmingham

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

These PDF slides are available in my 'talks' directory:

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#dag09>

This is part of a collection of presentations on related topics about requirements and designs for intelligent systems, natural and artificial, available here:

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

Most of them are updated from time to time.

Feel free to suggest improvements or send me criticisms: A.Sloman@cs.bham.ac.uk

Overview

The Dagstuhl seminar overview states:

Our hypothesis is that the form and shape of objects is a key factor deciding upon actions that can be performed with an object.

While not disagreeing with that, I argue that:

There is a large, mostly unnoticed, gap between form and function, which is also a gap between form and affordances.

The gap includes proto-affordances; possibilities for change, of many sorts.

There are also meta-proto-affordances, epistemic affordances, and various kinds of vicarious affordances.

These slides give an introduction to some of the variety of types of proto-affordances that exist in various kinds of environment and introduces some of the problems of representing them.

There are closely related problems of how to represent **processes** and how to represent various kinds of scene fragments that can occur in processes.

While much research is concerned with how a machine or animal can learn to predict what **will** happen in various situations, or what **the consequences** of various happenings will be, that presupposes a solution to a more basic problem of how to represent the **possibilities** from which predictions need to be selected.

Being able to represent and think about **what is possible** is a precondition for being able to think about **what will happen** or **probabilities of possible happenings**.

See Chapter 2 of *The Computer Revolution in Philosophy* (1978)

<http://www.cs.bham.ac.uk/research/projects/cogaff/crp/chap2.html>

Much modern research assumes configurations, changes, processes and probabilities can all be represented numerically, often using global coordinate frames, ignoring the possible alternatives to global coordinate frames and numerical representations.

Three key points

- For an organism or robot to be able to perceive, think about, or make use of affordances in a situation, or the functions of some object, it needs to be able to represent and predict consequences of possible actions that could be performed involving that situation or object.
- Being able to predict consequences of possible actions requires being able to represent the **actions that are possible** in that situation or involving the object, and also being able to represent **possible future processes** involving that object or situation.
- This is a special case of a much more general ability to think about **what is possible in a situation**, e.g. how things can change in a situation, including which relationships between objects and parts of objects can change and which properties, locations, orientations and shapes of objects can change.

Actions are a special case of the more general category of **processes**.

Thinking about or perceiving actions requires the ability to think about and perceive processes.

See also

Actual Possibilities (1996)

<http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15>,

Ontologies for robots and animals

It is obvious that humans, and some other animals, and also future human-like robots, need to use ontologies that include:

- (a) **configurations** of physical objects made of many kinds of physical matter, with many different spatial structures and relationships.

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

Less obvious needs are for ontologies to include also:

- (b) **processes** involving changes in everything mentioned in (a), and **interactions between processes**;
- (c) **possibilities for change and interaction** (possible processes), including possibilities that are not realised; (the vast majority of possibilities are not realised).
- (d) **influences on possibilities**, including enabling, hindering, preventing, and modulating what is possible;
- (e) **information** and **sources of information** about contents of the environment (a) to (f);
epistemic affordances are concerned with what information is available.
- (f) **information-users** that can acquire, manipulate, derive, store, communicate, and use information.

Representational requirements for handling (a) to (f) differ.

(E.g. meta-semantic competences are needed to represent instances of (f).)

I call possibilities and constraints on possibilities in items (c) and (d) “proto-affordances”.

The epistemic affordances in (e) include information available about proto-affordances and causal relations involving them, e.g. enabling, changing, preserving.

Many proto-affordances depend on empty spaces, whose existence, shape, size, location, etc., enable or constrain possible movements of things occupying space.

Hard problems about forms of representation

Saying that some things, e.g. processes, need to be represented is one thing.

Specifying **how** they should be represented is another matter.

- E.g. some people find it plausible to suppose that **processes** could be represented by something like **video** recordings of instances of processes.
- But that form of representation cannot explain what is involved in understanding a video, since if understanding a video required creating another video (which is understood) then that could generate a requirement for infinitely many videos in order to watch one process.
- In any case re-runnable videos (or 3-D simulations, as in physics-engines) do not provide the information required (a) for thinking about features of processes that can be varied and (b) for reasoning about the consequences of variation.
- Re-runnable videos or simulations will also not support considering new combinations of structures and processes, e.g. when designing a new kind of escapement for a clock, or a new steering mechanism for a car.
- We need **modular** forms of representation of processes that include information about sub-processes and how those are related to portions of the objects involved, in a systematic fashion, that allows re-arrangement and recombination of parts to be used in reasoning about new kinds of process – with new arrangements of process-components: **Grammars for various sorts of process?**

Possible changes of varying complexity

Here are some of the ways a simple line (on the left) can change, continuously or discretely, e.g. into any of the other forms shown:



It can grow longer, grow thicker, change colour, become curved, acquire additional inflection points, change orientation, or change in several ways at once: changing any combination of length, thickness, colour, curvature, number of inflections, orientation, etc.

Such **combinations of processes** in our environment are sometimes indications of **liquids**, while others indicate life forms (in some cases the processes are very slow).

A non-linear object, such as a blob, can also be seen to change in many ways:



It can grow, change colour, change texture, acquire dents and protrusions, be rotated, split into disconnected blobs, merge with another blob, and many more.

Processes possible for 3-D objects are much richer.

One of the complications involving 3-D objects is **changes of visibility** of portions of surfaces, as objects rotate, or pass through or behind other objects, producing and removing occlusions and self-occlusions.

Some constraints on motion may also be invisible, such as constraints on motion of a screw relative to a bolt through which it passes: rotation about its long axis is impossible without translation along the axis.

“Possible”, vs “probable”, vs “will happen”

It is important to distinguish the ability to perceive and think about the **possibility** of X occurring as distinct from any ability to predict that X **will occur** or **has a certain probability of occurring**, with or without conditions (discussed later).

We can consider possibilities without being able to assign any probabilities to them or to order their likelihood or probability.

The ability to represent and reason about what is **possible** is clearly more basic than the ability to find out what is **likely** or **probable** in a situation, since all the things that have (non-zero) probabilities, whether high or low, must be treated as possible.

Even the ability to perceive some type of object when presented with an instance depends on the pre-existing ability to represent objects of that type: learning about actualities presupposes implicit knowledge of possibilities.

However:

In some cases there may be very fast development, triggered by presentation of a new kind of object, which results in immediate extension of representational powers to cope with objects of that sort.

In that case the learner must initially have been very close to having the competence.

You cannot learn something general by being taught unless you were very close to learning it unaided.

Specific, e.g. historical, geographical, facts are exceptions.

Why proto-affordances are more general than affordances and functions

James Gibson's notion of [affordances](#), for an agent in a situation, is explained in

The Ecological Approach to Visual Perception, (1979)

which developed ideas presented earlier in

The Senses Considered as Perceptual Systems (1966)

Affordances include only the [subset of proto-affordances](#) that

- (i) are concerned with actions possible for the agent and
- (ii) are relevant to purposes, desires, dislikes, needs, preferences, etc. of the agent.

I use “Proto-affordance” as a more primitive and more general concept, lacking any reference to an agent, desires, dislikes, preferences, etc.

Some proto-affordances, but not all, provide affordances for the perceiver.

Others have nothing to do with the perceiver or other agents, e.g. proto-affordances include the possibility of treetops waving in a breeze, or the possibility of boulder rolling down a slope.

There are also [vicarious affordances](#), relevant to needs, intentions, etc. of [another](#) individual: e.g. seeing proto-affordances relevant to children, other animals, enemies, prey, collaborators, etc., or oneself in a future possible situation – “auto-vicarious affordances”.

Attempts to study affordances and functions of objects without taking account of the underlying proto-affordances, and causal relations between them, are bound to prove shallow and inadequate, and will not explain creative use of novel situations (e.g. combining proto-affordances), in humans, other animals or future robots.

[Compare trying to learn a language by learning only whole sentences.](#)

.... and also more general than probabilities

Proto-affordances are more general than probabilities.

- Discussing the probability of something happening requires a prior grasp of alternative possibilities.
- If you don't know anything about what can happen you have no way of talking about (relative or absolute) probabilities concerning different alternative things that can happen.
- So thinking and reasoning about probabilities **presupposes** an ontology in which there are things that are possible, and a form of representation that allows alternative possibilities to be expressed.
- However, you can think about things being possible without having any idea of what the relative probabilities are or even any way of finding out.
- You can think about
 - the possibility of the sun exploding in the year 2053
 - a tornado destroying the tower of London at some future date
 - somebody discovering a proof of Goldbach's conjecture
 - humans landing on Mars in the 22nd century
 - all forms of cancer being curable by the end of the 22nd century

without being able to assign probabilities, or relative probabilities to those possibilities.

Permissive vs constraining proto-affordances

Gibson's affordances for an agent could be positive or negative: positive affordances enable or support actions of the agent, while negative affordances prevent or constrain actions. Likewise there are **negative proto-affordances**.

- If a proto-affordance is the possibility of a rock rolling down a hill along a particular path, a negative proto-affordance could be the effect of narrowness of the gap between two trees, or the effect of a ditch across the path, either of which has the power to prevent the rock continuing down that path.
- In those cases configurations in the environment would **cause** rolling to stop.
- Some negative proto-affordances prevent objects remaining still: e.g. it is possible for a stone to remain where it is on a rock (a static proto-affordance), but that possibility would be eliminated by a process in which a larger fast-moving rock hits the stone.
- In other cases causes prevent motion, e.g. where an obstacle makes a certain movement impossible.
- Where friction prevents motion, adding a lubricant or releasing pressure, can allow it.
- In general, causes are positive or negative **conditional** proto-affordances: if C were to occur that would create or remove a proto affordance (possibility of change or non-change).
- We can think of causes as **meta(or second-order)-proto-affordances**: they are possibilities which if realised produce or remove other proto-affordances.

Different sources of constraints on possibility

The previous examples show that what is enabled or made impossible by existence of some process or configuration has at least two very different aspects:

- Some aspects of what is and is not possible are due to spatial structures and relationships (e.g. the gap between two trees being wider than a rolling rock, or a wall being in the space between two objects)
- Other aspects are due to the materials of which things are made: in particular whether they are rigid, flexible, easily broken, strong, transparent, opaque, and their density or total weight.

It is important not to confuse

- questions about what a particular agent in the situation, knows or believes about the entities involved in some interaction and
- questions about what is the case.

An agent may not know whether interacting entities in a process are completely rigid, or completely impenetrable, yet be able to reason about the consequences if they are and if they have certain shapes and relationships (as on the next slide).

Moreover, **uncertainty is often qualitative, not quantitative (probabilistic)**: I know that if two gear wheels are rigid, impenetrable and unbreakable, then if meshed they must rotate in opposite directions; if not, I don't know what will happen, and have no probabilities.

Conditional possibilities and impossibilities

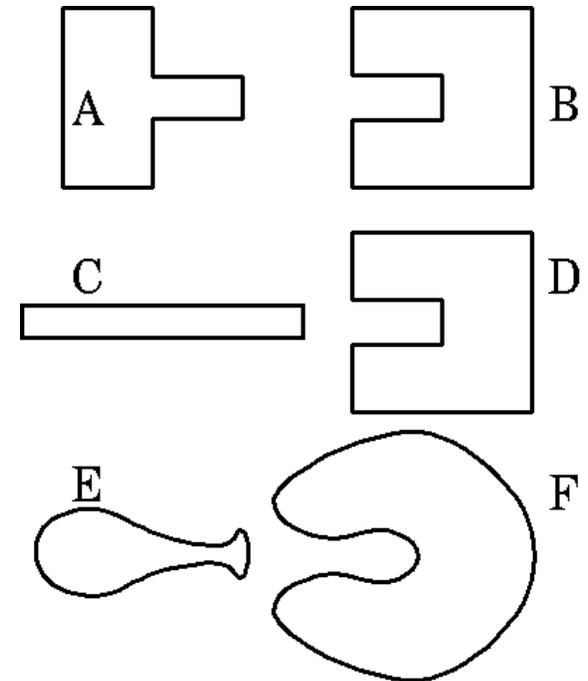
Changes of relations can alter constraints

Suppose you are looking down at six flat plates (A to F) lying on a flat smooth table, all made of rigid impenetrable material, where the items on the right (B, D, and F) are fixed to the table.

A, C, and E can slide and rotate freely in any direction on the supporting surface, but some of their possible movements will alter subsequent possibilities.

You can probably see that:

- If A moves right from the position shown, after a while its ability to rotate and to move up or down will be constrained. Eventually, with B, it will form a composite object with a rectangular outline, and the only possibility for motion of A will be moving left, away from B (if the fit is tight).
- If C moves right its ability to move up and down and its ability to rotate will be constrained: when fully to the right if an upward or downward force is applied at its left end, depending on the material of which it and D are made, either nothing will happen, or the whole configuration will rotate, or C will bend or break, or D will be deformed – or some combination!
- If E moves right its motion will be obstructed by F, and then E will be able to move left or move slightly up or down while rotating. If it is lifted off the table, moved further right then dropped, a part of E will be trapped by F, preventing movement to the left. E will be able to rotate slightly and to move left or right or up or down slightly. Depending on the materials used, various forces applied to E could deform or break E or F or both.



How could an AI system produce something like the above descriptions of possibilities and impossibilities? (Compare Sloman IJCAI 1971)

Matter matters

Reasons why it is important to use information about different kinds of matter include

- the properties of matter affect the changes/processes that are possible
e.g. whether things can be bent, stretched, dented, twisted, folded or otherwise deformed;
- the properties of matter affect the **propagation of consequences** of sub-processes and the consequences of variations in a process.
 - For example, if an object is rigid and impenetrable, then, as one part of it moves, that has consequences for how other parts move, and how other things in contact with it move.
E.g. think of rotating gear wheels with their teeth meshed, or a nut rotating on a threaded bolt.
 - A particular feature of rigidity is that when **rotations** occur, the effects change scale with distance from the centre of rotation.
 - An important class of substances has fairly high tensile strength, and high flexibility with respect to bending but no stretchability (e.g. some strings, shoe-laces).
Such a thing can be used to transmit a force from one place to another place where the route between the two places includes a number of obstacles.
E.g. pulling down the end of a string hanging over the edge of a table can cause something tied to the other end of the string to move horizontally along the table.
 - Incompressibility combined with fluidity can cause a pushing force to be transmitted round corners (using pistons connected by rigid or fixed-width tubes).

Towards an ontology for a science of affordances and functions

The world is full of things that can change. Some possible changes occur and are perceived, while others do not occur, but their possibility is perceived, along with constraints that limit the possibilities.

- Some possibilities are **topological** – changes of contact, containment, order, connectivity.
- Some are **geometrical** changes: length, distance, angle, shape, speed, etc.
- Some are **physical**, e.g. forces, pressures, strains, chemical composition, temperature, phase, density, strength, colour, energy, ...
- Some are changes in **information available** at a location. or from a direction (e.g. through vision, touch, smell, etc.)

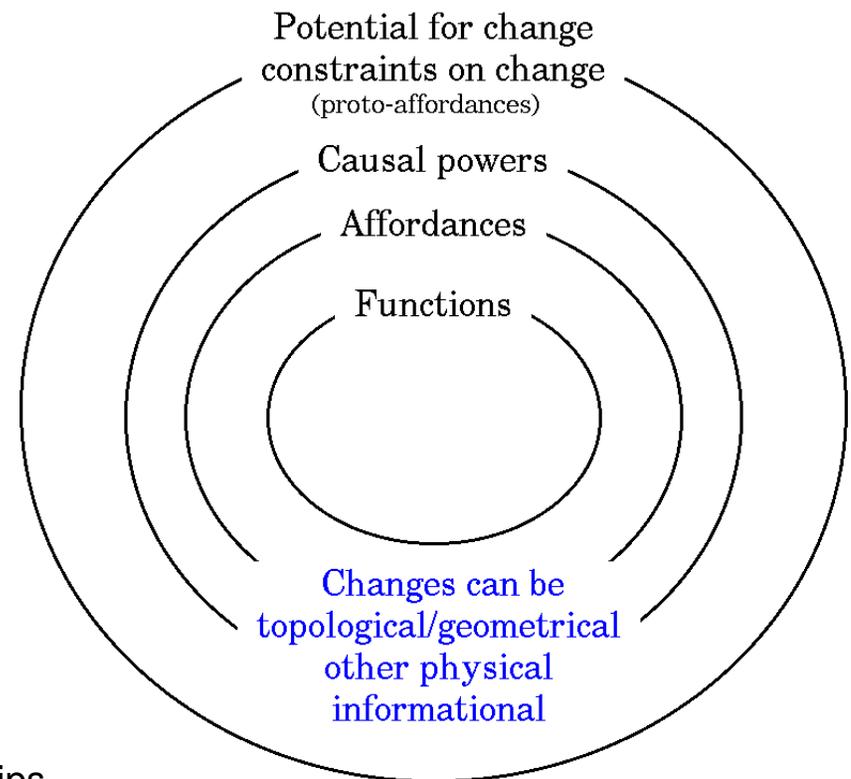
Proto-affordances are possibilities for, and constraints on, change in a situation, including changes in relationships.

Causal powers are powers, based on both form and matter, related to producing or constraining changes.

Affordances (for an individual perceiver) are the causal powers relevant to enabling or constraining actions by **that** individual in ways relevant to **that** individual's goals, needs, preferences,...

Functions are causal powers either intentionally added to, or designed into, or selected for, or regularly used in some object or situation. (By evolution, by an individual's intention, by a social practice.)

Thanks to Danica Kragic for the idea of the diagram



Searching for possible constrained movements

Seeing **simple** proto-affordances involves seeing what simple rigid or non-rigid processes (translations, rotations, bending, etc.) are and are not possible in a situation.

Seeing **compound** proto-affordances involves seeing (serial or parallel) **combinations** of processes that are possible and seeing what constraints exist at different stages in those combinations.

Sometimes this requires search.

(The following is based on an idea by Jeremy Wyatt.)

In each configuration, consider these questions:

- Can the rod with a blue end slide from the “start” position to the “finish” position within the rectangle, given that the grey portion is made of rigid, impenetrable material?
- In cases where it is possible, how many significantly different ways of doing it are there?

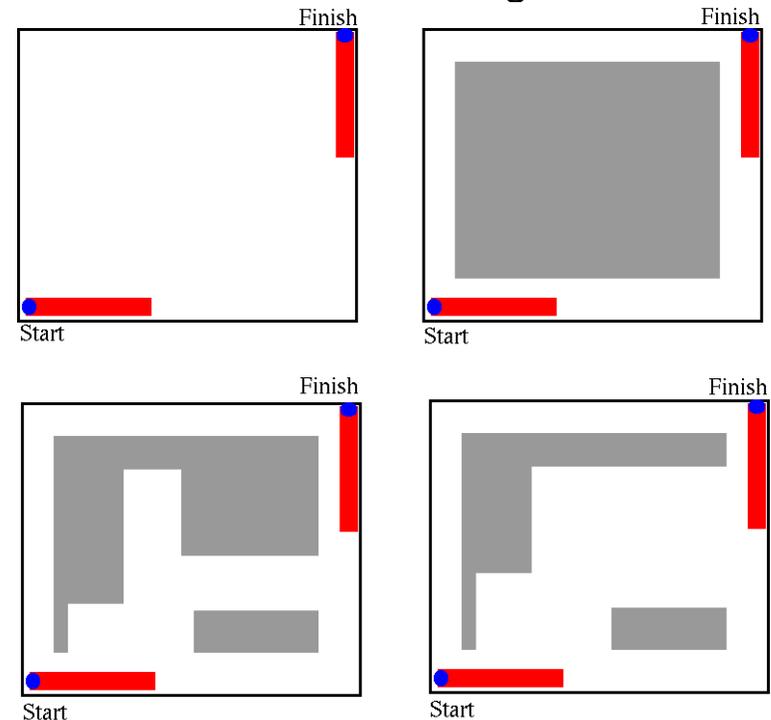
Answering that requires a way of “chunking” possible alternative processes.

These tasks use abilities to detect the **possibility** of movements that are not happening, and the

constraints limiting such movements, and also to **visualise** combinations of such movements, while **inspecting** them for consequences: **Using what brain mechanisms?**

Consider a ball just large enough to fit through the narrowest gaps, moving from bottom right to top left.

Could it move concurrently with the red and blue rod in all configurations?



Carrying a chair through a door

3-D Process fragments (**proto-affordances**) can be combined, in sequence or in parallel, in action or in hypothetical reasoning, to form new complex processes (actual or possible).

Affordances can interact in complex ways when combined, because of changing spatial relationships of the objects involved during the processes of performing the actions.

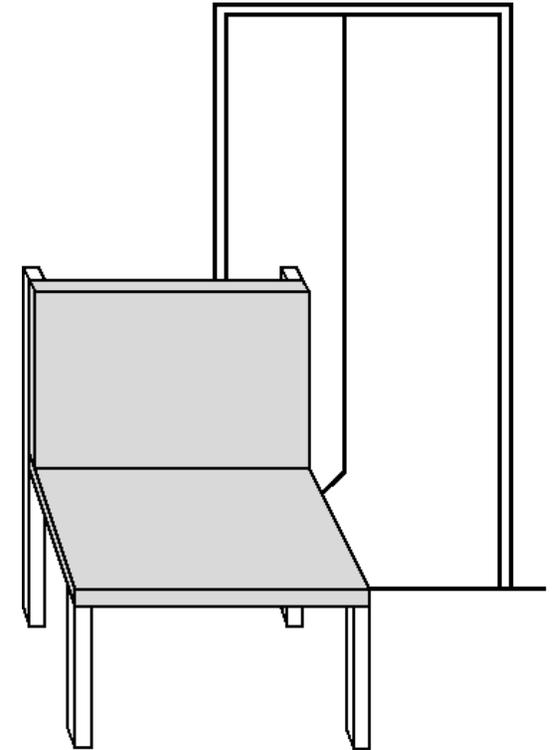
A large chair may afford lifting and carrying from one place to another, and a doorway may afford passage from one room to another, but attempts to combine the two affordances by lifting and carrying the chair to the next room may fail when the plan is tried.

A very young child may not be able to do anything about that, but an older child who has learnt to perceive the possibility of rotation of a 3-D object, may realise that a combination of small rotations about different axes combined with small translations some done in parallel, some in sequence. can form a compound process that results in the chair getting through the doorway.

Is any other type of animal is capable of understanding that?

Even the very familiar process of grasping an object is a complex combination made of various successive sub-processes, and some concurrent processes, with concurrently changing relationships between different parts of the surface of the object and different parts of the grasping hand.

Problem: **What needs to be added to traditional AI planners to enable them to construct plans involving such continuous, concurrent, interacting processes?**



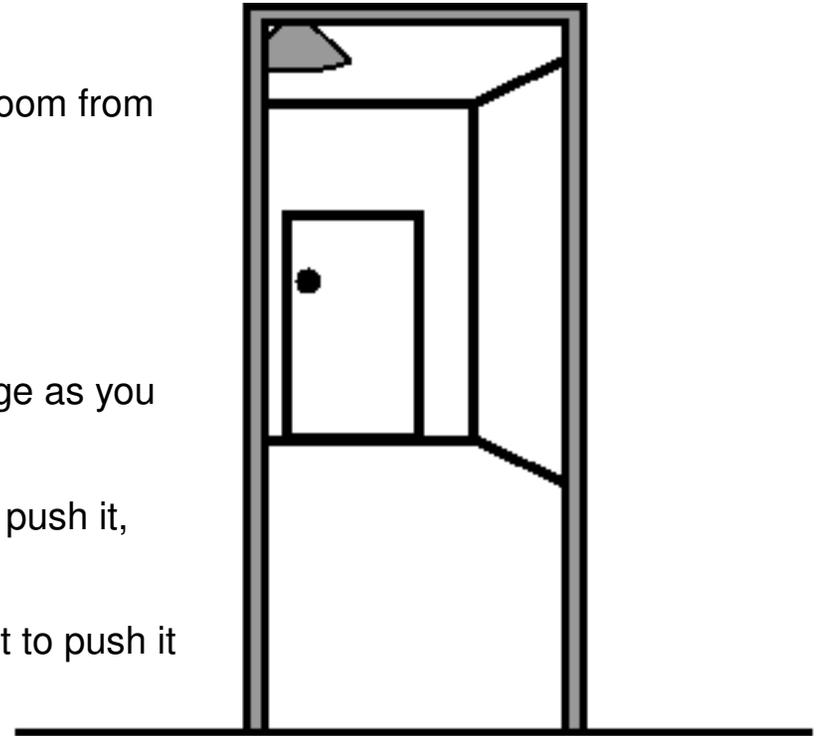
Getting information about the world from the world

An **action** affordance concerns what can and cannot be **done** by the perceiver, whereas an **epistemic** affordance concerns what information is and is not available in the environment.

Actions can change both action affordances and epistemic affordances.

Things you probably know:

- You can get more information about the contents of a room from outside an open doorway
 - (a) if you move closer to the doorway,
 - (b) if you keep your distance but move sideways.Why do those procedures work? How do they differ?
- Why do perceived aspect-ratios of visible objects change as you change your viewpoint?
- In order to shut a door, why do you sometimes need to push it, sometimes to pull it?
- Why do you need a handle to pull the door shut, but not to push it shut?
- Why do you see different parts of an object as you move round it?
- When can you avoid bumping into the **left** doorpost while going through a doorway by aiming further to the **right** – and what problem does that raise?



Clues about human and animal abilities

There are so many and such varied possibilities (proto-affordances) adult humans can see and think about, that it is very hard to find ways of imposing structure on that variety.

Finding partial orderings in evolution, development, or learning may give clues as to how various competences work.

Contrast research on earliest ages at which young humans can do various things, or research on which human abilities exist in other species – these may yield no explanatory information.

It may be possible to produce a partial dependency ordering by analysing examples of possibilities older children can see or do or think about, that they could not do earlier.

- Careful observation may reveal intermediate states.
- A learner may fail to see how to: open something, shut something, put something inside something else, get something out of a container, get something into a hole or recess that requires alignment, find the next smallest cup to stack, transfer food from plate or bowl onto fork or spoon; and many more.
- A child will be able to see a process occur yet not be able to replicate it – until some time later.

What changes?

Imitation is not an **explanation** of learning, it is something that **needs to be explained**, e.g. in terms of ontologies, forms of representation, prior knowledge, algorithms, architectures required, etc.

- Discovering how competences in other animals **co-occur** or **develop** can also suggest dependencies.
- Finding partial orderings in such competences may provide clues for research on the forms of representation, mechanisms, architectures that can grow in an individual.
- Attempting to design and implement working systems, can help us learn more about dependencies between requirements, competences, representations, mechanisms, architectures, etc., if we look!

Clues about forms of representation

Thinking only about human adult competences can tempt us to make strong assumptions about the forms of representation used, e.g. because adult humans can use propositional forms including logical connectives and quantifiers.

Competences of other animals (e.g. nest-building birds, hunting mammals, primates, elephants), and competences of pre-verbal children, can suggest representational requirements to be met in a wider range of animals than adult humans.

All have requirements for perception, motivation, intention formation, plan formation, action control (etc) that need forms of representation supporting:

- structural variability (i.e. not just quantitative variation)
- recombining of parts (generative powers)
- compositional semantics (with context sensitivity)

These requirements are met by human communicative languages, and various formal languages (e.g. predicate calculus, programming languages).

But they may also be met by other forms of representation, including **analogical representations** used internally. (Sloman, 1971)

What forms of representation of information are capable of meeting these requirements is an open question.

For more on this see the discussion of generalised languages (GLs) in <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang>

Generative representations for processes

It would not be possible in one lifetime to learn separately for each type or shape of object, or for each pair, triple, etc. which possible changes/processes it can be involved in.

So we need some widely applicable **default** ways of generating representations of possible processes, so that the possibilities can be reasoned about, used for deriving affordances, making predictions, etc.

- A generally applicable set of ways of generating possible processes is to consider how to represent changing geometrical or topological relationships.
 - Make gaps between separate surfaces larger or smaller.
 - Alter orientations, or orientation-relations, of parts or whole objects.
 - Alter alignments, of parts or whole objects.
 - Alter curvature of some object parts or surface features.
 - Alter thickness, length or shape of flexible objects.
 - Alter topological relations, e.g. contact, overlap, containment, betweenness, etc.
 - Alter relative velocities, directions of motion, etc.
 - Things becoming attached, or separated/broken.
- Since relations between objects can include multiple relationships between different parts, i.e. **multi-strand relations**, processes can involve simultaneous changes in different relationships i.e. **multi-strand processes**.
 - E.g. alignments and distances between edges and faces of two blocks can change simultaneously.
- Less widely applicable characterisations of change include changes of colour, kind of matter, and other physical properties.

Learning in infants and toddlers

A great deal of the earliest learning by humans (and possibly many other species) seems to be concerned with;

- finding out what changes are possible in the environment (learning proto-affordances),
some produced by the learner, some not
- learning about consequences of those changes in various situations,
Including learning about:
 - effects of combining processes both sequentially (in different orders) and concurrently (with different spatial relations, e.g. contiguity – with or without pressure, etc.)
 - **meta-*proto-affordances***: possibilities which if realised will alter other possibilities, e.g.
possible changes in position of A that alter possibilities for motion of B
inserting C between A and B prevents certain processes bringing A and B into contact.
(I.e. second-order proto-affordances.)
- learning which of the agent's possibilities for action are meta-*proto-affordances*,
which actions can enable, initiate, maintain, hinder, modulate or terminate, various processes
- learning qualitative statistics: discovering among things that are possible in various conditions which are more likely and whether likelihood differences have detectable conditions.

The implication is that causal learning can be either deterministic or non-deterministic. I.e. it can be about what is brought about or prevented, or about making things more or less likely, or simply about making things possible (e.g. unbolting a cage door).

Learning based on play and exploration

Such learning can be produced by play and exploration driven by innate motivators whose outcome is increasing breadth and depth of knowledge about

- what things can occur and which possibilities exist in different situations (proto-affordances)
- how new proto-affordances can be composed from old ones by considering combinations of processes that can occur sequentially or in parallel
- The various kinds of matter in the environment and the ways in which they can differ and how those differences affect proto-affordances.
- which side-effects can occur as a result of various proto-affordances being realised (including both absolute and conditional side-effects).
- which proto-affordances are within the individual's powers, e.g. which could be affordances for the individual.
- how the actions the individual performs affect the environment (including affordances related to the individual's motives, preferences, etc).
- statistical information regarding correlations between occurrences of different proto-affordances. (These need not be numerical probabilities, e.g. if they are only qualitative orderings of probabilities,
- such curiosity based processes can cause the individual to develop a host of concepts of types of shape, types of material, types of process, types of interaction between process, etc. Probably many thousands are learnt by typical human children.

Notes on teaching/learning concepts

Concepts form **families** of different sorts.

Some concepts are inherently functional and therefore **cannot** be taught simply by presenting examples and uttering words, since examples do not determine a common function:

E.g. knife, kitchen, corridor, window, cornflakes

Also various action concepts: fight, hammer, avoid, drill, glue, eat, capture, cook, steer, ...

Concepts that belong to a family **cannot** be learnt without learning about the family

(e.g. learning how its members are related, and why, and what that implies)

- concepts of places, routes, regions, doorways, doors, tunnels, are all concerned with regions of space that
 - (a) are all related to one another in various ways (e.g. see work of Tony Cohn, Pat Hayes, and others)
 - (b) may or may not be connected (temporarily) for a given agent or mode of locomotion
 - (c) may be continuously or sequentially occupied by various entities
 - concepts of parts that can be assembled in various ways to make a class of entities
 - (a) parts of houses and other buildings (bricks, walls, floors, windows...)
 - (b) parts of electronic devices
 - (c) parts of operating systems (scheduler, interrupt handlers, file-system, ...)
 - (d) parts of sentences (phonemes, syllables, words, phrases, ...)
 - (e) parts of processes (causes, actors, patients, roles, enablers, means, obstacles, ...)
- And many more such families

The concept “room” presupposes a theory

If a learner L, e.g. a child or robot, is taken into some rooms by a person P who says in each one “this is a room”, should L thereby acquire the information that:

- P can make noises?
- P is trying to communicate with L?
- P is trying to teach L that “room” is a label for a concept?
- The concept is a type concept (a common noun)?
- Instances of the concept are spaces?
- Instances of the concept are spaces occupied by people?
- Instances of the concept are spaces bounded by plane surfaces?
- Instances are spaces bounded by vertical and horizontal plane surfaces?
- Instances of the concept are created by humans for a purpose?
- Instances of the concept have openings through which things like P and L can move?

What are the conceptual competences that could allow L to learn any of those?

How could L find out which hypothesis about P’s intentions is correct?

AI learning systems are typically incapable of reaching any of those conclusions: the “teaching” is just a form of programming that produces ritualised responses to a range of tests – but nothing like the normal human understanding of the concepts involved.

Human concept learning builds on a rich collection of other competences and knowledge.

The pointless 2D/3D debate: both can be useful

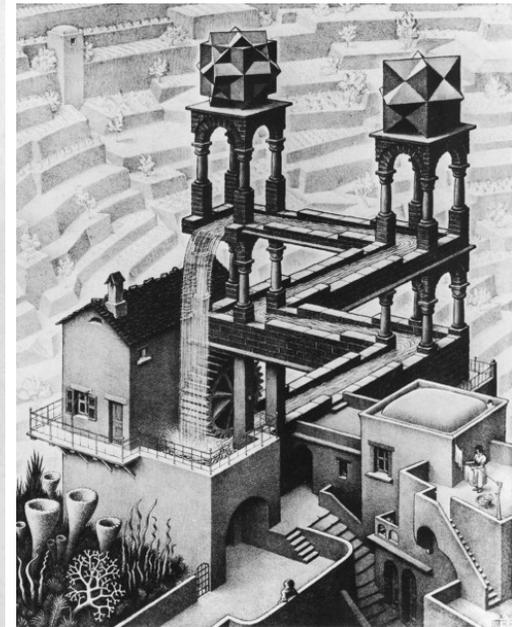
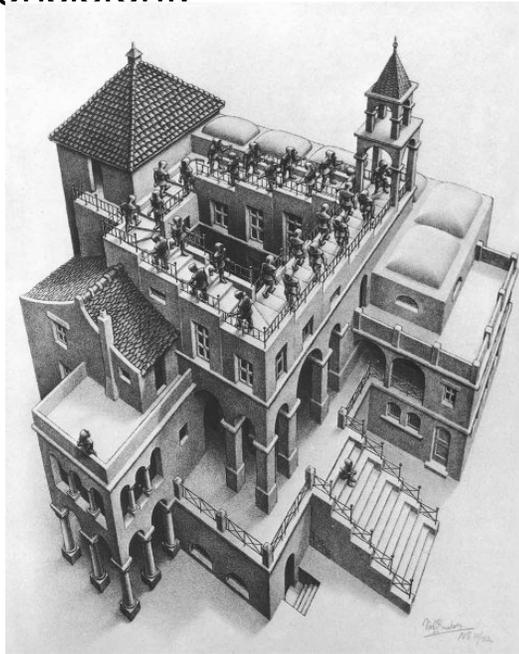
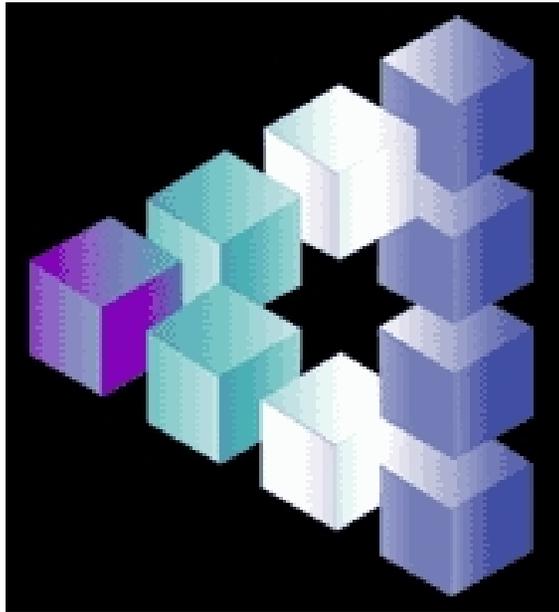
Some researchers in vision and robotics attempt to restrict all information to a **somatic** ontology based entirely on whatever can be found in, or derived by data-mining from, sensory and motor signals and patterns therein.

Others include an **exosomatic** ontology that includes 3-D entities, relationships and processes in the environment.

- That somatic/exosomatic contrast is sometimes wrongly linked to a contrast between 2-D and 3-D structures in vision.
- The environment includes some inherently 2-D structures, for example, the information structured in the visual optic array (Gibson) associated with a particular viewpoint and viewing direction (whereas a laser range-finder has a 3-D “optic array”).
- 2-D retinal image projections (or projections in the primary visual cortex) are not the same thing, since the optic array is exosomatic whereas the projections are somatic (internal) information structures produced by **sampling** the optic array:
 - The internal structures can change during saccades across an unchanging optic array.
 - Visual tracking can cause changing external structures to map onto unchanging somatic structures.
- Gibson realised that patterns and changes in the optic array are a very rich 2-D source of information about the 3-D environment. (I think there’s more than he noticed.)
- Those patterns are particularly relevant to changes in **epistemic** affordances, since movements of the perceiver, or of objects in the environment, systematically vary the contents of the optic array. <http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#906>

Another temptation: treating 3-D interpretations as models

Pictures by Oscar Reutersvärd (1934) and M.C. Escher (1960,1961) show that 3-D scenes as perceived, need not be consistent, proving that perception is not model building, since models must be consistent.



Conjecture: 3-D perception involves construction of consistent fragments of information about space occupancy and proto-affordances, assembled in a manner that does not impose consistency, but supports usability (many kinds of usability).

The form of representation must support consideration of, and reasoning about, ways in which the scene might be different, i.e. [possibilities for change](#), or proto-affordances.

You can visualise objects being moved around in the above scenes, and other objects moving through the spaces (e.g. your hand grasping and moving a cube, or birds flying through gaps).

Vision research should not ignore processes

Very few vision researchers seem to pay attention to the importance of perception of **processes**, whether actual or possible.

The majority seem mostly to study perception of **static** structures (apart from 2-D tracking).

Usually even the study of static structures is limited to segmentation and recognition of whole nameable objects in various static configurations – ignoring object parts and the relationships between parts within and between objects, as well as ignoring changes in those parts and their relationships.

Many actions require perception of discontinuities and regions of varying curvature in surfaces: e.g. bumps, dents, ridges, grooves, holes, openings, protrusions, edges, corners, etc.

Moreover in dynamic scenes the properties and relationships of such features can change in parallel.

Changes in “multi-strand” relationships produce “multi-strand” processes.

There is no need to assume that all perceived features are describable in human language – human language evolved **after** animals were able to perceive, interact with, and use things in the environment.

It can be argued that the main precursor to evolution of human language, and development of language in human children, is a set of forms of representation used **internally** for perception, thinking, reasoning, motivation, planning, control of actions, learning etc. found also in other species.

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

This will include means of representing structures, processes, and proto-affordances of various kinds along with epistemic affordances, and vicarious affordances.

Seeing shapes without a global measurement frame

Most vision researchers assume that perceiving shapes, sizes, angles, velocities, curvatures all involve use of a global metric

so that any two shapes, lengths, angles, curvatures, etc., perceived simultaneously or at different times, will be comparable.

It is highly implausible that animal brains are capable of assigning such measures that are comparable across arbitrary spatial gaps or time intervals (unless using very low levels of precision).

However there are special cases, e.g.

- Two linear objects aligned with a pair of ends adjacent – if one object is longer the far end will protrude.
- Being able to “visualise” a translation or rotation that will bring one of two objects alongside the other so that their sizes are comparable.
- Being able to watch a process in which a rigid object is first placed alongside one object then translated to be alongside another.
- Being able to compare the curvature of a convex surface (or edge) with the curvature of a concave surface by attempting to place the former in the latter.
- Comparing plane angles by aligning one arm of each.

And very many more: we need a systematic study of ways in which “semi-metrical” (partially ordered) measures can be assigned

It should not be assumed that spatial structures and relationships or their changes and rates of change are represented using a global metric and a unified coordinate frame used for locations, orientations, curvatures, velocities, etc. See <http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#brown>

Perception is not recognition

Many researchers assume that perception is primarily recognition. In order to counter this and also for other reasons the “polyflap” domain was developed in the CoSy cognitive robotics project and is also being used in the successor project CogX <http://cogx.eu>.

A polyflap is defined to be a rigid object formed by taking any 2-D polygonal shape and producing a single fold, forming a 3-D object: that’s equivalent to joining two polygons at one edge of each.

More information about polyflaps can be found here:

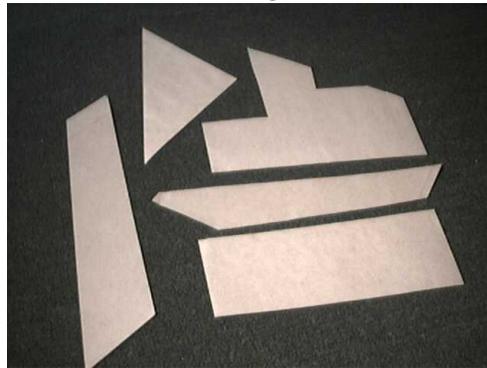
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0504>

Discussion note on the polyflap domain (to be explored by an ‘altricial’ robot)

<http://www.aaai.org/Fellows/Papers/Fellows16.pdf>

Polyflaps as a domain for perceiving, acting and learning in a 3-D world

Configurations of polyflaps can be perceived and manipulated (whether recognised or not), though current robots cannot do much with such things. Future robots could learn much by playing with polyflaps.



Polygons on left, Polyflaps on right. Shadows can help or hinder

Clearly we can see configurations of unfamiliar polyflaps, can rearrange them and can construct new configurations in which some are supported by others.

What forms of representation, of structures and processes, are needed, and what mechanisms?

Different types of vision research

Vision research in AI/psychology takes many forms, with different types of focus, some of which are:

- Type 1 research on vision: objects
- Type 2 research on vision: scene structures
- Type 3 research on vision: affordances & functions
- Type 4 research on vision: processes and process possibilities (proto-affordances)
- Type 5 research on vision: control (planning, servoing)
 - NOT “what” vs “where”
 - Compare “what” vs “how” in Sloman 1982:
<http://www.cs.bham.ac.uk/research/projects/cogaff/06.html#0604>
- Type 6 research on vision: recognition of affective states in others
 - This is typically very shallow, partly because based on the very shallow research on emotions found in psychology literature (most of which has nothing to say about information processing architectures).
- Other types (e.g. tracking, SLAM,)

Type 1 research on vision: objects

There is much research, both in psychology and AI/Robotics/Machine Vision that focuses on learning about object categories and recognition (and/or tracking) of instances of those categories.

Earlier slides indicate that the ability recognise object types (for which humans have names) is not a fundamental visual capability since much can be seen and acted on without recognition.

It is arguable that seeing involves recognition of various kinds of object fragment (e.g. surface patches, edges, corners, bumps, dents, etc.) and relationships between them.

This includes the ability to see such features smoothly combined to form a single complex (possibly novel) surface.

A major problem in vision research is to explain how spatial composition is perceived.

When that has been done, it may be possible to do recognition properly!

Type 2 research on vision: scene structures

There is also much research (going back to the 1960s and earlier) on perception of object and scene structure, including locations, orientations, curvature and other properties of objects, object parts and surfaces, and their relationships.

(That includes the “syntactic” approach to vision, inspired by work in linguistics on parsing, etc.)

E.g. S.Kanefff *Picture Language Machines* (1970?)

One of the good things about some of this work was its emphasis on generality: finding ways in which some relatively small collection of primitives can be recombined to generate a very wide range of scenes.

An example was research on polyhedral scenes: Clowes, Huffman, Waltz,
(Recently extended substantially by research at Cardiff University. REF)

Biederman also proposed primitives and modes of composition: based on “geons”.

I think his set of primitives is of the wrong sort: we need [surface fragments](#) to be combined not 3-D [solids](#), to get the right sort of generality – also process fragments involving moving surfaces.

Composition of whole objects can come later.
(Compare clock-making vs sculpture.)

Type 3 research on vision: affordances

More recently, many people follow J.J.Gibson in emphasising the role of perception of affordances in vision and other perceptual modalities (e.g. haptic perception of affordances).

For an individual X to perceive affordances involves several conditions:

- (1) X gains information about which actions are possible for X (positive affordances) or hindered (negative affordances), where
- (2) X is capable of performing those actions in some situations, and
- (3) performing those actions is able to contribute to achievement of X's goals or prevention of states or occurrences that are undesirable for X.

Much research on perception of affordances seems to be based on learning to recognise object categories (e.g. cup, chair, knife, hammer) and learning shallow associations with action categories (e.g. lifting, pouring, sitting, cutting, hitting, etc.)

For an individual to learn about and understand **new** affordances it needs the more basic ability to perceive and think about/ reason about possible motions – i.e. an understanding of proto-affordances provides a basis for investigating the special cases that are affordances.

Type 4 research on vision: processes and process possibilities (proto-affordances)

For some time I have been claiming that all three of the previous sorts of research (on objects, structures and affordances) ignore visual competences that lie **between** perception of structure and perception of affordances (between perception of form and perception of function), and are more general and more fundamental than either, namely:

Processes

- (a.1) perception of processes in the environment (whether produced by the perceiver, or another agent, or simply by physical forces, such as wind, gravity, water pressure, changing weather, or plant behaviours, etc.)
- (a.2) where the processes need not be relevant to the perceiver's goals or preferences and need not be ones that the perceiver is capable of producing or preventing, and

Possibilities of and constraints on processes

- (b.1) perception of possibilities for and constraints on processes of those kinds, independently of whether they are relevant to the perceiver's possible actions, goals, or preferences (i.e. perception of possibilities and constraints that are not affordances), and
- (b.2) independently of whether the perceiver is capable of producing, preventing, or constraining those possible processes.

See 'Architectural and Representational Requirements for Seeing Processes, Proto-affordances and Affordances' (2008).

<http://drops.dagstuhl.de/opus/volltexte/2008/1656>

Type 5 research on vision: control (planning, servoing)

- NOT “what” vs “where” as some have assumed.
- But control involves process
- We need to understand many ways of relating to processes: not just perceiving them.
- Not just recognising, classifying, categorising them...
- Somatic vs exosomatic ontologies

Teachability/learnability of concepts

Concepts that come in families defined by roles, and relationships, rather than by collections of features, **cannot** be taught simply by presenting instances:

it is necessary for the learner to acquire the relevant modes of reasoning as well – theories about what is involved in being an X, or part of an X, or a cause of Y, etc.

Learning to identify some enlarged photographs of pathogens will not teach you what pathogens are.

The same is true of types of botanical and zoological species.

For the vast majority of concepts (if not all), learning the concept involves

learning other concepts and some sort of theory linking the concepts

where the theory supports more forms of inference than just class membership or set inclusion inferences:

The theory may support

- predictions

- explanations

- evaluations

- decision-making

Hypotheses (1)

The seminar proposal states:

Our hypothesis is that the form and shape of objects is a key factor deciding upon actions that can be performed with an object.

“The proto-affordance hypothesis”

Yes, but...

Trying to go direct from form and shape to affordances/functions is a bit like trying to map sequences of phonemes directly to meanings, without using intermediate structures, e.g. words, phrases, clauses.

Understanding how to go from form/shape to affordances/functions, requires use of an ontology that includes:

- (a) the form and shape of objects, parts of objects and relations between them
- (b) a rich variety of phenomena that need to be perceived, represented and reasoned about:
 - processes,
 - possible processes,
 - interactions between processes, and
 - constraints on processes
- (c) functions and affordances of objects and locations
- (d) A variety of **epistemic affordances**, which can be changed by processes.

We need to understand how structures, processes and causation can be perceived, understood and used, **before** we try to understand the special cases of affordances and functions.

Why?

Why do we need to understand how structures, processes and causation can be perceived, understood and used, **before** we try to understand the special cases of affordances and functions.

In order to deal with novelty, and to understand why objects and configurations have particular functions and affordances.

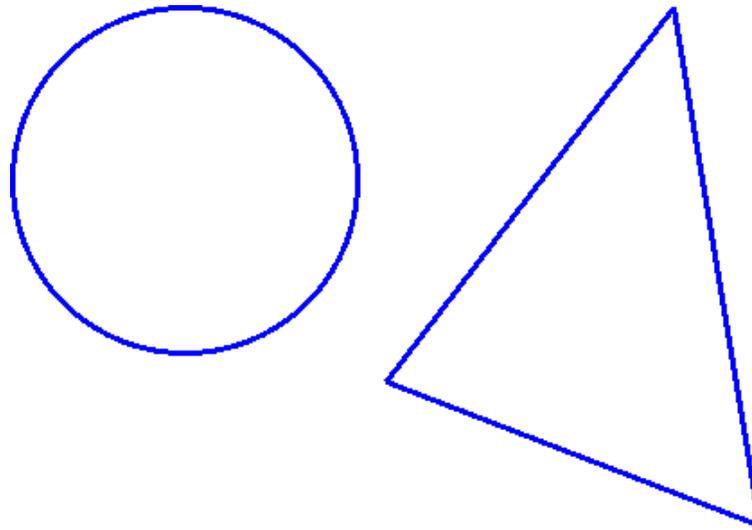
Also needed for repairing broken items, for designing new examples.

Hypotheses (2)

Any bit of the world that allows different kinds of **information** to be acquired and used includes **epistemic affordances**.

Including epistemic affordances concerning information about epistemic affordances.

Possibilities and constraints on possibilities



How many possible points of intersection if these two shapes move around, changing position, size orientation?

Under what conditions can a robot think about this and get the right answer without having had to move the objects around?

What enables you to avoid having to do empirical research?

See these slides on “toddler theorems”

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

Disclaimer

I am not claiming that any individual can perceive ALL such possibilities and constraints in the environment.

A major form of learning is learning to see NEW types of possible process and NEW types of constraint on processes.

Some of the concepts involved in perceiving, characterising, and reasoning about actually occurring and possible processes make use of concepts of 'kinds of stuff',

i.e. different sorts of matter with different dispositional properties, including different sorts of rigidity, breakability, flexibility, elasticity, fluidity, stickiness, etc.

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

Proto-affordances

I call the possibilities and constraints in the environment that are not affordances ‘proto-affordances’ since they are the kinds of primitive modal facts about the environment that underlie affordances.

Two senses of “modal”:

1. Facts are “modal” in the logician’s sense (contrasted with “categorical”) if they are not about what is or is not the case, but about what can or cannot or must or may be the case.
2. Facts can be logically modal yet “a-modal” in the psychologists’ sense, i.e. not concerned with any particular sensory or motor modality, though possibly relevant to several.

Amodal concepts in the second sense can be either “somatic”, i.e. concerned with what’s going on inside the agent, or “exosomatic”, concerned with what is happening in the environment (e.g. 3-D surfaces changing their relationships).

(Theories about mirror neurons seem to me to fail to take account of the requirement for this kind of abstraction.)

Types of vicarious affordances

In some cases where affordances are perceived they are not affordances for the perceiver, but for some other individual, e.g. a child, or prey, or predator, or collaborator.

I call those 'vicarious affordances'.

In some cases where affordances are perceived they are affordances for the perceiver, but for some other occasion, some other possible location, some future possible situation.

I call those 'autovicarious affordances'.

Epistemic affordances

Besides processes in which physical structures, properties, and relationships change there are also processes in which information is acquired, e.g. through vision, touch, hearing, or other senses, or in some cases with the use of instruments.

Any particular configuration of a portion of the world supports and hinders the acquisition of particular sorts of information.

The possibility of acquiring information can be called a **positive epistemic affordance**.

Obstacles or limitations on acquiring information (of many kinds) can be called **negative epistemic affordances**.

There are many complex relationships between proto-affordances, action-affordances and epistemic affordances.

In particular, intelligent agents need to work out which epistemic affordances are relevant to their goals in order not to waste effort attempting to process information of the wrong kind, e.g. wasting effort trying to cope with noise and uncertainty where both can be ignored because there is enough unambiguous information at a higher level of abstraction.

See

<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0702>

Predicting Affordance Changes (Alternatives ways to deal with uncertainty),

More on affordances for others

- The above characterisation of X being able to perceive affordances can be generalised to include seeing or learning what can and cannot be done by individuals similar to X (i.e. with similar action capabilities and similar kinds of goals, needs, interests, etc.) who are not necessarily present at the time of perception.
- This is very close to perception of 'vicarious affordances', except that the relevant other agent need not be present.

It can also be related to perceiving affordances that X would have if located in a different place relative to the objects with those affordances (auto-vicarious affordances).

That can generate an explosive set of possibilities, which can be partly defused by abstracting away from many of the details of particular actions, e.g. representing a wide variety of types of grasping simply in terms of two opposing surfaces coming together with an object between them.

- A further generalisation would include abstracting away from various details of X, e.g. seeing graspability in terms of a wide range of possible grasping devices, not just devices similar to the perceivers' hands, claws etc. This shades into perception of proto-affordances (See Type 4 research)

Warning

Don't assume that the main function) of vision is to apply categories from our language (e.g. object, form, or function labels)?

The labels (and subdivisions) are cultural artifacts that arrived after millions of years of evolution.

What were animal vision systems doing before that?

Contributing to:

Eating

Avoiding

pushing out of the way

Learning (playing)

Building, repairing, breaking...

Understanding causation

What do we (and other animals) use vision for, apart from applying labels or helping us to act on things?

Think of answers involving

mates, offspring, food (prey), predators, conspecifics

other moving things (inanimate, animate)

shelters, aids to transportation, weapons, tools, routes, layout of terrain, ...

Gibson started something

But he did not realise how much more there was

Gibson's contrast

- vision (perception) is not acquiring information about what is out there (e.g. as Marr and others suggested)
- it meets a biological need
 - related to the perceiver's possible goals/needs
 - related to the perceiver's possible actions

But that's a special case.

Of what?

- It's a special case of what I have been calling [proto-affordances](#).
- And there are other special cases, including various kinds of vicarious affordances and various kinds of epistemic affordances, and also meta-affordances (e.g. affordances which if used can produce or prevent other affordances coming into existence, or enduring).

Gibson's idea may have been the beginning: but it was a small step on a long, still mostly untravelled research road.

Early organisms may have used perception only for Gibson's affordances, but later evolution produced visual systems with more powerful uses.