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

Aaron Sloman
School of Computer Science, University of Birmingham UK
A.Sloman@cs.bham.ac.uk
`http://www.cs.bham.ac.uk/~axs/`

Abstract

Test domains for AI can have a deep impact on research. The polyflap domain is proposed for testing complex AI theories about architectures, mechanisms and forms of representation involved in features of human and animal intelligence that evolved to enable perception, action, and learning in diverse environments containing things that we can perceive and manipulate, and many complex processes involving objects that differ in shape, materials, causal properties, and relations to one another. We need a test environment that is rich enough to provide some of that variety of structures, processes and affordances, yet simple enough to be within reach of robotics research in the not too distant future.

AI as the General Science of Mind

Much AI research has engineering goals, but many AI researchers (including the founders) are also interested in AI as the *science* of how minds actually work in nature and how artificial minds of various kinds might work. This has to be a multi-disciplinary search for deep theories. We propose the polyflap domain[1] as an example of a testbed for such theories.

Whereas benchmarks for assessing progress can be derived from engineering goals, that is not so easy for the study of natural and artificial minds. E.g., it is far from obvious what competences humans and other animals actually have. They have many different capabilities integrated in a single functioning system, but specifying what those capabilities are and how they are integrated is hard (Sloman & Chrisley 2005). E.g. many vision researchers did not notice the importance of perception of affordances, until Gibson's work. Affordances can relate both to the perceiver and to others ('vicarious affordances') whose potential for action the perceiver needs to understand, e.g. children, predators, or prey. (See our paper in the AAAI'06 Cognitive Robotics workshop.) Even now it is unclear exactly what is involved in perceiving and using affordances, and many AI vision benchmarks have little to do with perceiving for action.

McCarthy (1996) discussed requirements for achieving 'human-level' intelligence, contrasting 'the bounded informatic situation' with 'the common sense informatic situation'. He suggests that many will find the list of problems to be overcome for the latter 'dismayingly large'. However the list grows even larger if, in addition to the ability to reason, plan, solve problems and answer questions, discussed by McCarthy, we also consider how children and many animal species, also learn to perceive and act in a world in which there are many simultaneously changing structures, relationships and causal interactions occurring at different levels of abstraction, some of them involving continuous changes that need to be perceived and controlled in real time, while others involve discrete (e.g. topological, functional) changes, including changes of positive and negative affordances.

This is not just a rehearsal of the 'new AI' fashion for dropping symbolic AI and using swarms, pheromone trails, neural nets, evolution, dynamical systems etc. Those techniques cannot solve our problems. My point is not to promote specific *solutions*, but to stress the need for much deeper research into *what the problems are* that we have to solve: i.e. requirements. This is one of the themes of the UK Computing Research Committee's Grand Challenge 5: 'Architecture of Brain and Mind' (GC5).[2]

Some researchers hope to by-pass the problem by attempting to model new-born infants on the assumption that they have no innate knowledge and merely use some powerful and general learning mechanism, that will learn as humans do if we behave to the robots as we do to children.

However knowledge-free learning may require evolutionary time-scales, because the space of concepts and facts that might be learnt is so vast. Moreover, different species learn different things in the same environment, suggesting that each species has specific knowledge and/or specific learning frameworks that, together with features of the environment, enable its members to learn.

A child of four may have learnt many things that none of its ancestors ever learnt, e.g. how to play computer games. But that learning may be based on deep and general learning about how to perceive, manipulate, reason about, assemble, disassemble, and interact in many ways with structured objects in a physical environment. In (Sloman & Chappell 2005) we indicated how the spectrum of precocial and altricial competences found in animals might also occur in robots built for a range of purposes.

[1] See http://www.cs.bham.ac.uk/~axs/polyflaps for full details

[2] See http://www.cs.bham.ac.uk/research/cogaff/gc/

Scenario-based Research on Requirements

Elsewhere (e.g. members poster session AAAI'06) I have promoted the idea of analysing requirements by designing and analysing a partially ordered network of scenarios, starting with ambitious distant scenarios and working backwards towards near-term tasks that could provide soundly derived early steps to the long term goals. Many existing AI challenges, including Robocup soccer and Robocup rescue can contribute to this, though the first includes limited affordances for manipulation and robocup rescue may be too difficult in the near future.

Another one that is closely related to GC5 is to try to develop a working model of a prelinguistic child (not a newborn infant) with a collection of competences such as might enable such a child to help a disabled person around the house, including learning to communicate. But solving all the problems of vision and manipulation is so far beyond the state of the art that considerable simplification will be needed (A.Sloman 2006). The polyflap domain was designed as a solution to that problem of combining structural richness with near term feasibility.

Polyflaps

A polyflap is a polygon (concave or convex) cut out of a flat sheet of some material (e.g. cardboard) and folded once (anywhere) to produce a 3-D object. Figure 1 shows some polyflaps before they have been folded. Figure 2 provides a moderately complex configuration of polyflaps, made by folding those in Figure 1, and stacking some on others.

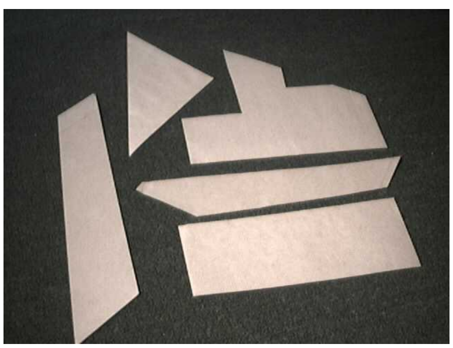

Figure 1: Polygons used for polyflaps

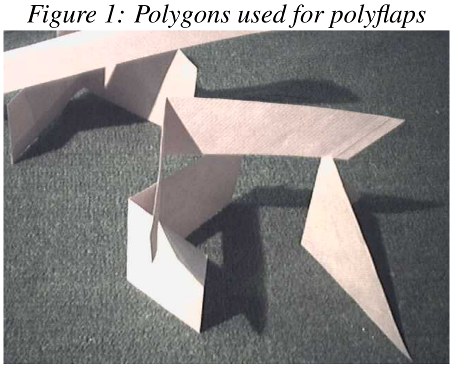

Figure 2: Shadows may help or hinder!

There are several important features of polyflaps. (a) Although they do not include curved surfaces and edges they raise many of the problems of perception of spatial structure, which could be addressed by continuing the work on 3-D vision that has mostly stalled since the 1970s (e.g. see (Barrow & Tenenbaum 1978)). (b) Unlike flat sheets the polyflaps are relatively graspable by a robot with only two fingers. (c) In general grasping cannot simply be achieved by opening the gripper wide, approaching an object and then closing the gripper (as with small balls and blocks): it will require understanding the spatial configuration to determine approach angle and the orientation of the gripper surfaces as well as gripping location. (d) Processes will generally involve multiple spatial and causal relations changing concurrently. (e) Both individual polyflaps and configurations can be made simple to start off then made increasingly complex indefinitely, allowing more and more complex challenges and more and more sophisticated learning about the domain to occur in a robot that learns through play and exploration. (f) Polyflaps could be made of various materials, e.g. differing in flexibility, weight, hardness, etc., so that a robot will have to learn that there is more to causation and affordances than spatial structure and motion. (g) Experiments with polyflaps can be done with children and other animals including learning about non-bayesian (Kantian) causation.

The topic is too rich to be exhausted here. Further discussion of goals and the polyflap domain can be found in

- http://www.cs.bham.ac.uk/~axs/polyflaps
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0505 A (Possibly) New Theory of Vision (PDF)
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0506 Two views of child as scientist: Humean and Kantian
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0601 Orthogonal Recombinable Competences Acquired by Altricial Species (Blankets, string, and plywood)
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0603 Sensorimotor vs objective contingencies
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604 Requirements for a Fully Deliberative Architecture

Acknowledgement

This work is being done as part of the CoSy project with Jackie Chappell, Tony Cohn. Nick Hawes, David Hogg Somboon Hongeng, Marek Kopicki and Jeremy Wyatt.

References

A.Sloman. 2006. Aiming for More Realistic Vision Systems. Technical Report COSY-TR-0603, University of Birmingham, School of Computer Science.

Barrow, H., and Tenenbaum, J. 1978. Recovering intrinsic scene characteristics from images. In Hanson, A., and Riseman, E., eds., *Computer Vision Systems*. New York: Academic Press.

Gibson, J. 1979. *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.

McCarthy, J. 1996. From here to human-level AI. Modified version of invited talk at KR'96, online at http://www-formal.stanford.edu/jmc/human.html.

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

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), http://www.cs.bham.ac.uk/research/cogaff/04.html#cogsys