CHAPTER 11

Conclusions and future directions

> *It's always worth having a few philosophers around the place. One minute it's all Is Truth Beauty and Is Beauty Truth, and Does a Falling Tree in the Forest Make a Sound if There's No one There to Hear It, and then just when you think they're going to start dribbling one of 'em says, "Incidentally, putting a thirty-foot parabolic reflector on a high place to shoot the rays of the sun at an enemy's ship would be a very interesting demonstration of optical principles."* (Pratchett, 1992b)

The work in this thesis was motivated by fallibility and transience. As humanoid robots become more sophisticated mechanically, it seems likely that their perceptual abilities will become a severe limiting factor on what they can do. In the absence of perfect perception, it will be important to have simple experimental methods that resolve ambiguity, and methods to derive information from clumsy actions so they are not repeated. Active segmentation is a good example. We may also expect that humanoid robots will need considerable flexibility – perhaps the task for the robot may change from day to day. It is best to build in adaptivity from very beginning. This thesis has made some steps towards building a perceptual system for a robot that can grow and develop through contact with the world. This is both a theoretical effort to show, for example, how adaptive modules can have persistent interfaces, and a practical effort of identifying and engineering opportunities for the robot to develop (see Figure 11-1). The areas explored are for the most part complementary to other work on development in robotics (Metta, 2000; Roy and Pentland, 2002; Weng et al., 2000b). This chapter summarizes the specific contributions made and identifies some important directions for future research.

11.1 Summary of significant contributions

- ▷ **active segmentation** – the thesis showed that passive methods for object segmentation can be augmented by physical exploration.
- ▷ **appearance catalog** – the appearance of small, low-level features can be almost exhaustively characterized, as this thesis showed for oriented features, given results that are competitive with classical model-driven approaches.
- ▷ **open object recognition** – for a robot it is important to integrate object recognition with a mechanism for enrolling new objects, since there are far more objects in the world than can

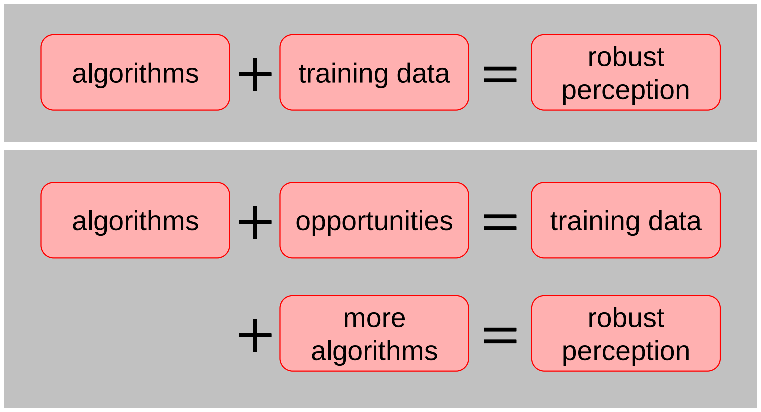

Figure 11-1: One route to autonomy is to switch the emphasis from collecting training data (top) to engineering methods to create and exploit opportunities for collecting and labeling such data autonomously.

reasonably be catalogued. This thesis showed that this makes false detections a less serious problem, because distracting objects can simply be enrolled and modelled explicitly, rather than having to come up with an accurate background model.

- **affordance recognition** – for a robot, it makes sense to switch from object-centric perception to recognizing action opportunities. A concrete example of this is given for rolling, an affordance that is of particular importance to a robot that needs to manipulate awkward objects. This work was a collaboration with Giorgio Metta.
- **open speech recognition** – in speech recognition, there is a trade-off between recognition accuracy and vocabulary size. This thesis assumes that at any time the vocabulary a robot needs is small and task-dependent. By creating explicit run-time mechanisms for vocabulary modification, the robot can quickly be given the vocabulary appropriate to the current task without needing a large pre-existing vocabulary.
- **virtuous circle of development** – familiar activities can be used to identify components used in roles within those activities. Then those components can be tracked out into unfamiliar activities, and used to discover the structure of those activities. These two processes dovetail to give a circle of development.

11.2 Grounding operational definitions

In this thesis, the appearances of objects and manipulators were characterized by using operational definitions. These are definitions which translate to measurements :-

> An operational definition is a procedure agreed upon for translation of a concept into measurement of some kind. – (Deming, 1993)

An effective procedure for finding objects, seen as physically coherent structures, is to poke around and see what moves together. An effective procedure for finding manipulators, defined as something that acts on objects, is to watch what pokes objects. Of course, both these procedures are not completely general, and they are worth generalizing. For example, active segmentation gives clear results for a rigid object that is free to move, but what happens for non-rigid objects and objects that are attached to other objects? Here the results of poking are likely to be more complicated to

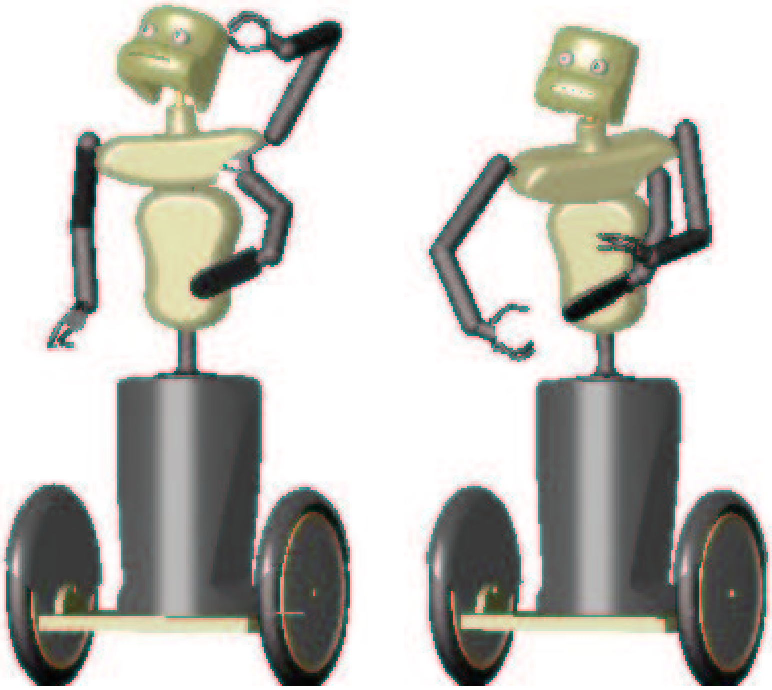

Figure 11-2: A segway-based robot (drafted by Jeff Weber).

interpret – but in a sense this is a good sign, since it is in just such cases that the idea of an object becomes less well-defined. Poking has the potential to offer an operational theory of “object-hood” that is more tractable than a vision-only approach might give, and which cleaves better to the true nature of physical assemblages.

11.3 Fully autonomous platform

Both Cog and Kismet are fixed platforms, and so have access to a very limited environment. A new project in the Humanoid Robotics Group at the AI lab seeks to remedy that (see Figure 11-2). This project combines several important threads: mobility (the robot will have a modified segway base), expressiveness (the robot head has a simple face and eyes), and dexterity (the robot has three arms). This combines all the elements needed for autonomous object exploration and activity learning.

11.4 Philosophy

This thesis has focused primarily on learning perception, and is complementary to other work on Cog that addresses learning action (Marjanović, 2003). In animals and robots, perception is fundamentally for action:

> Perceptions and ideas are found, upon analysis, to have their factual significance and meaning in terms ultimately of overt operation. Their meaning derives from the potential effect, that is, the difference they make or may make in behavior. In both its phylogenetic and ontogenetic histories, mental activity develops out of, and in reference to, overt action. – (Sperry, 1952)

This work is very much influenced by the "alternative essences of intelligence" enumerated in Brooks et al. (1998) – development, social interaction, embodiment, and integration. It attempts to bring all these threads together in one system, and show that they do in fact reinforce each other. Strong integration of perception and action provides the means for development to occur; social interaction and embodiment provide the opportunities. To those familiar with murder mysteries, the missing element is 'motive.' A complete treatment of this topic should examine how to move gracefully between training activity, where the goal is to learn, and actual performance of a task, where the goal is to achieve some more specific end.