Chapter 6: Conclusions

6.1 The effectiveness of focus of attention

The results described earlier in this thesis demonstrate that mechanisms for focusing attention can greatly decrease the amount of work required by a learning system. Both goal-independent and goal-dependent focus methods can be employed simultaneously, decreasing the work of learning (and hence increasing its speed) by taking advantage both of invariant characteristics of the world and of guidance provided by the set of goals that the agent must achieve.

Because these methods can decrease the rate of growth of a fundamentally $O(n^2)$ process, their effects only increase as the size of the problem or the number of known facts increases. This research showed combined improvements of over a factor of 50; longer runs would have shown even more.

In addition, it was demonstrated both qualitatively and quantitatively that the correctness and completeness of the learning performed in the systems studied was not impaired by these techniques. One pays a price for them, namely having to perform more experiments in the focused case to learn roughly comparable amounts of knowledge about the world, but this price is quite small compared to the increase in efficiency that results. It is possible that there are many other systems which can utilize similar techniques to achieve faster learning without substantially sacrificing correctness or relative completeness.

As mentioned in Section 3.3.1, on page 57, and in Section 4.4.1, on page 94, the results presented here are primarily from the infant/eyehand scenario. However, a smaller number of similar runs have been performed in the Hamsterdam scenario, with comparable results. Given the approximate similarity of sensor systems and action repertoires in the two scenarios, this is unsurprising, but reassuring. The Hamsterdam scenario additionally offers the potential for more interesting goal sets, due to the more dynamic world available—there are many opportunities available there for more sophisticated experiments involving goals and their control.

In short, using focus of attention is one of the many techniques that can and should be employed to allow autonomous agents to learn more about their environment with less computation. This can allow certain applications, which formerly ran too slowly to be practical, to be run at more reasonable speeds.

6.2 Future work

There are many ways in which this work might be improved or extended. A representative sampling of such ideas follows.

This is clearly far from an exhaustive list. Indeed, viewed in a larger context, the questions from [Maes 94] are still very much with us. Focus of attention cannot hope to address all of those questions, but many of them might be partially answerable by using more sophisticated focus mechanisms. This is an area deserving of future investigation.

6.2.1 Generalization and abstraction

One of the most frustrating aspects of the current schema system concerns its inability to generalize in certain ways. The synthetic item machinery allows one form of generalization, which is important for shifting to different levels of abstraction, but the implementation used in this research lacks composite actions, which severely limits the sort of generalizations that might be made. Even with this machinery in place, simple generalization across a category of input would be very useful;[1] currently, the learning system requires several examples at *each* point in the state space and therefore does not perform this sort of generalization. This research has not addressed the details of what would be required to function effectively in a learning system which employed such types of generalization.

The focus system as currently implemented does not explicitly address different levels of abstraction in the learning system. A system which created explicit categories at various levels of abstraction would require some form of support for focus; how to do this well is an open problem.

6.2.2 Filtering

In addition, the overall focus mechanism as implemented decides how to *filter* its perceptions, cognition, and actions based on characteristics of the domain and its current set of goals, analogously to certain ideas about attention as a system of limitations that were questioned in Chapter 5. It does not reason at a metalevel about the goals themselves,[2] nor does it use multiple concurrent learning mechanisms or multiple simultaneous processing pathways. As such, there are large opportunities for further work using multistrategy learning and recent, non-filter-based, multi-locus ideas from cognitive science. Such work could enable a more sophisticated action selection system.

1. An example would be automatically inferring that, if moving the eye right causes an object to appear to slide left from one particular visual location to another (e.g., by turning off one visual item and turning on the visual item to its left), then this would be true at *all* points in the retina. Such generalization would require a retinotopic map (e.g., not the unordered “bag of bits” currently employed) so that concepts such as “to the left of” could be inferred without exhaustively acquiring data about every adjacent pair of visual items. Without such a map; there is no way of even determining adjacency without such exhaustive experimentation. Such a retinotopic map *is* assumed in particular goals defined here (e.g., we assume that we know a priori, due to hardwiring, which coarse visual items actually correspond to the fovea, in certain goals), but this is not a general mechanism and cannot really be used by the schema system per se in order to increase its representational power.

2. For example, to change their hardwired mapping from the goal to the allowed set of percepts and actions.

6.2.3 Experimental strategy

In addition to the above ideas about the action selection system, a more intelligent experimentation strategy would be welcome. In particular, within the currently-allowed set of actions specified by the active goals, actions are chosen randomly. This leads to exploration of the state space in a manner which is probably quite inefficient. Consider the figure below, from [Thrun 94]. In this figure, we have a robot which must simply navigate from

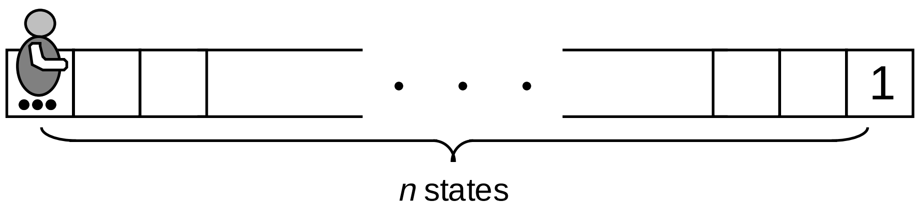

Figure 19: A task which is *O(1)* with counters, and $O(2^n)$ without

one end to the other, *without* learning. If the robot takes steps randomly either to the left or to the right, its expected time to reach the goal is exponentially bad, e.g., $O(2^n)$. If, on the other hand, the robot is allowed a *counter-based* approach in which it simply drops a counter on each visited square, and picks a square without such a counter when it can, its performance improves to *O(1)*. Indeed, it has been proven that, subject to some simple and common assumptions[3] any learning technique based on random walk is inefficient in time [Whitehead 91a] [Whitehead 91b]. On the other hand, even a very simple strategy such as "go to the least visited neighboring state" can reduce this inefficiency from exponential, e.g,. $O(2^n)$ time to polynomial, e.g., $O(n^2)$ time, regardless of whether or not one has a model that can predict the next state from the current one [Thrun 92] [Thrun 94].[4]

3. These are: a state space which is finite, deterministic, and ergodic (e.g., no states from which, once entered, the agent cannot escape), in which the agent receives a reward only in the goal state; there is no information available about the domain a priori; random actions change the distance to the goal state by only +1, 0, or -1, and can be expected to increase the distance to the goal on the average; and, finally, the size of the state space is polynomial in the largest possible distance to the goal state, e.g., the *depth* of the state space (this holds for most state spaces studied in literature, e.g., grids of arbitrary dimensionality).

The system explored in this research does not quite meet Whitehead's conditions. In particular, because of the stochastic characteristics of the domain microworlds studied, in which other entities may move and hence change parts of the state space, and because of the effects of sensor aliasing [Whitehead and Ballard 90], the system is nondeterministic. Similar results do not exist for nondeterministic domains, and for some *malicious* domains, it can be shown that *any* exploration technique will take exponential time to find a goal state [Thrun 92].

Nonetheless, given domains which are not malicious, and which are not *too* stochastic, it is quite possible that some sort of counter-based approach could increase the rate at which relatively-unexplored parts of state space are encountered, hence decreasing the amount of work per generated schema. One possible (but untried) approach would therefore be to use something akin to prioritized sweeping [Moore and Atkeson 93]. An even simpler approach could be to always pick that action A for which the average reliability of all currently-applicable[5] schemas containing action A is minimized.

6.2.4 Goals

The goal system implemented here is unsophisticated. It is unlikely to scale well to large numbers of goals, in part because of its rather nonhierarchical space of goals, and in part because goals and their relations to each other must currently be hardwired. It also offers little support for multiple concurrent strategic (as opposed to tactical) goals, or for sharing work between goals, nor do goals reason about their performance at a metalevel in order to better guide the learning, as is done in some current multistrategy learners [Hunter 94] [Ram and Leake 94]. A system which learned useful mappings from goals to the correct strategy for focus of attention, rather than having such a strategy hardwired in

4. Such a predictive model does help, but the problem is still $O(n^2)$.
5. E.g., context satisfied, meaning that their context agrees with the currently perceived state of the world.

for each goal, would also be quite useful. It would extend the current work from the realm of very simple animals further up the phylogenetic tree, and may help some of the probable scaling issues in the current design.

6.2.5 Occasional defocusing

The focus system's selectively is a bit sharp; it is essentially an all-or-nothing sort of focus. One that occasionally defocused might lead to more opportunistic exploration of the space without undue cost; integrating this into the system in an intelligent way touches upon many of the explore/exploit problems mentioned in [Maes 94].