Evolving Multi-Agent Behaviors Using a Tunable Problem Landscape

Jeffrey K. Bassett
George Mason University
Computer Science Department
Fairfax, VA 22030
jbassett@cs.gmu.edu

R. Paul Wiegand
George Mason University
Computer Science Department
Fairfax, VA 22030
paul@tesseract.org

Kenneth A. De Jong
George Mason University
Computer Science Department
Fairfax, VA 22030
kdejong@gmu.edu

One of the most challenging aspects of building intelligent systems is the design and implementation of control programs for intelligent autonomous agents. Manually designing and implementing control programs that are sufficiently robust to handle dynamically changing environments and uncertainty has proved to be extremely difficult. As a consequence, there has been considerable interest in the use of machine learning techniques to help automate this process.

A good deal of progress has been made in this area in the past few years using a variety of representations (rules, neural nets, fuzzy logic, etc.) and a variety of learning techniques (symbolic, reinforcement, evolutionary, etc.). A natural extension of this work is to consider solving difficult tasks with teams of cooperating agents, for example see Stone and Veloso (1998); Bull (2001). In this research we address the issue of designing fitness functions that reward individual agents in ways that produce desired team behaviors. As a first step we have developed a method of *tuning* the way in which fitness is computed and individuals are evaluated such that adjustment of parameters of this fitness landscape will result in agent behaviors with different kinds of emergent properties.

The first question one asks when evolving behaviors for these kinds of agents is what kind of control system will be evolved. For simplicity, we elected to evolve weights in a statically structured two-layer feed-forward neural network. The weights are represented directly as real-valued parameters, so individuals consist of a single string of these weights. A simple (μ, λ) Evolution Strategy (ES) is used to evolve the weights of the neural network.

Our research focuses on applying teams of automated micro air vehicles (MAVs) to the task of surveillance (Bassett and De Jong, 2000) in a simulated environment. The problem is for teams of MAVs to learn to fly above ground targets, attempting to keep interesting targets within the field of vision of their on-board camera. The MAVs should learn to discriminate between different levels of interest in order to maximize coverage, while avoiding one another. Such a task requires that they first *find* an interesting target (explore), and then keep it under surveillance as best as they are able (exploit). Better behaviors are those where the agents cooperate to attain these goals.

In order to allow us to investigate this balance, we developed a fitness landscape generator for our architecture that allows us to meter the availability of a renewable resource. In our case the resource is land to survey. Fitness is calculated by accumulating the area that all the MAVs see on each timestep, with more credit given for interesting areas. We can limit the availability of this resource by not counting an area for some number of time steps after an MAV has viewed it. This is referred to this as the "regrowth rate". It is our hypothesis that as the resource is diminished, more cooperative team behaviors will emerge in order to achieve higher fitness.

When we experimented with a variety of regrowth rates we found that controllers evolved in environments with high regrowth rates (abundant resources) tended to be very greedy and didn't explore the space much. On the other hand, those evolved in low regrowth rate environments tended to explore the space much more, and even appeared to use each other's positions as state information in order to build formations to cover the space more effectively. There was a tradeoff though. These individuals were not very good at exploiting the more interesting areas once they discovered them.

References

J. K. Bassett and K. A. De Jong. Evolving behaviors for cooperating agents. In *Proc. of the Twelfth ISMIS*, pages 157–165, Charlotte, NC., 2000.

L. Bull. On learning classifier systems in multi-agent environments. *Machine Learning*, 2001.

P. Stone and M. Veloso. A layered approch to learning client behaviors in the robocup soccer server. *Applied Artificial Intelligence*, 12, 1998.