Multi-Agent Reinforcement Learning for Swarm Retrieval with Evolving Neural Network

Neil Vaughan[1,2]

[1] University of Chester, Chster, UK
[2] Royal Academy of Engineering, London, UK
n.vaughan@chester.ac.uk

Abstract. This research investigates methods for evolving swarm communication in a simulated colony of ants using pheromone when foriaging for food. This research implemented neuroevolution and obtained the capability to learn pheromone communication autonomously. Building on previous literature on pheromone communication, this research applies evolution to adjust the topology and weights of an artificial neural network (ANN) which controls the ant behaviour. Comparison of performance is made between a hard-coded benchmark algorithm (BM1), a fixed topology ANN and neuroevolution of the ANN topology and weights. The resulting neuroevolution produced a neural network which was successfully evolved to achieve the task objective, to collect food and return it to a location.

Keywords: Artificial Ants, Neuroevolution, Swarm Communication.

1 Introduction

This research has developed a model of ant colony swarm intelligence behaviour. The novel aspect is that behaviour of pheromone navigation was not hard coded, as in most implementations, but has evolved using ANNs and an implementation of neu-rovevolution. Compared to previous research which failed to evolve standard and fixed topology ANNs for ant behaviour (Collins & Jefferson, 1990a), this research produces successful evolution and applies a more comprehensive neuroevolution methodology including complexification and augmentation of ANN topology and weights, as described by NEAT (Stanley, 2004).

Inspired by biological ants, this research aims to provide insights to advance understanding of how pheromone communication evolved in biological organisms. Application of neuroevolutionary computational modelling provides a useful analogy to how brains may have evolved to produce biological organism behaviours. Nature inspired algorithms largely enabled recent advances in multi-agent reinforcement learning.

1.1 Background of Swarm Communication Systems

The core interest of this work is how ant pheromone communication can be evolved in a computational model. There have been some interesting works attempting to evolve

ant pheromone communication, and evolving swarm communication in general. Literature on pheromone communication is described by various key words: ant evolution, pheromone simulation, central-place foraging algorithm (CPFA), pheromone recruitment (Letendre and Moses, 2013).

A milestone early attempt to use a computer simulation to evolve ant foraging strategies using pheromones which resemble behaviours of biological ants was AntFarm (Collins & Jefferson, 1990a). AntFarm implemented an early form of neuroevolution, which was used to evolve the ANNs which learn behaviour for effective ant pheromone communication (Collins & Jefferson, 1990b). Neuroevolution methods in AntFarm evolved both the ANN connectivity pattern (topology) and weights of the ANN which were under genetic control in a genotype. Limitations were that: (1) AntFarm did not successful evolve any cooperative foraging which was the main objective. (2) A basic, conventional ANN was used, when compared to the wider range of operators, sigmoids and activation functions with complexification as used in more recent neuroevolution models such as NEAT (Stanley, 2004). (3) The number of neurons and connections were not under genetic control. That feature is possible in this new research.

The first research to evolve Ant pheromone foraging was by Panait and Luke (2004). Previous research showed that 3D physical simulated robots can evolve longer legs and better eyesight by natural selection when in competition (Vaughan, 2015).

2 Implementation of Swarm Reinforcement Learning

Amongst foraging behaviours, signals can 'evaporate over time' simulated by decrementing the artificial pheromone – a global rather than local update. Ants have 13 input sensors: (1, 2) the location within the 9 adjacent cells (Moore neighbourhood) of the highest pheromone, (3, 4) the location within Moore neighbourhood which is closest to the nest, given by a 'compass sensor'. (5, 6) location within Moore neighbourhood of food. (7, 8) the direction of the ant's previous move, (9, 10) a direction picked at random, (11) a Boolean indicating whether the ant is currently carrying food, (12) a random number, (13) a fixed value of 1 (Bias). These are referred to as the pre-computed inputs and they remain the same even when the controller is changed (BM1, ANNs, NEAT).

The controller is a 'black box' brain which decides the animal behaviour at timestep t, based on the pre-computed inputs from the ant's sensors. The experiments were repeated using different controllers: a hard-coded benchmark (BM1), a fixed topology neural network and neuroevolution by adjusting the topology and weights of an ANN.

The resulting output of the controller determines the direction in which the ant moves. After each ant has moved, a number of post-move local updates are automatically applied. (1) If the ant is now standing on food and isn't carrying any, it automatically picks food up. (2) If the ant is carrying food, pheromone is deposited with strength inversely proportional to the time since collection. (3) If an ant is already carrying food and is now standing on a nest, it automatically drops the food. This representation realistically assumes that biological ants already could pick up and drop food before they evolved pheromone communication. These tasks are regarded as automatic responses

which we assume have been learnt previously. After a full iteration, when all ants have finished making a move, a global update is triggered in which all pheromone is evaporated (decremented). A number of different evaporation rates including decrementing and various percentage reductions were tested to identify how evaporation rate affects the ability to evolve navigation controllers.
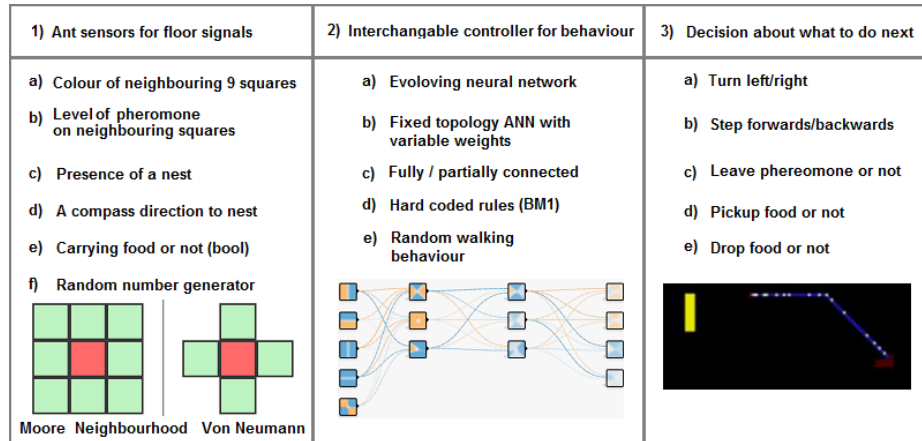

Fig. 1. The three main components of the developed evolving ant pheromone system.

3   Custom Designed Benchmark Algorithm

A hard-coded benchmark foriaging algorithm was developed and tested with 10 random levels of food and obstacles generated in a 500 x 300 grid. The benchmark was run for 5000 timesteps and the number of food collected to the nest was recorded. Each run took approximately 19 seconds. Some of the 10 levels resulted in food being returned very rapidly, due to random positioning close to the nest (Fig. 1)

Every food square has 30 foods so it requires 30 collections either by the same or different ants. Food squares are in clusters or around 10-15 on each cluster.
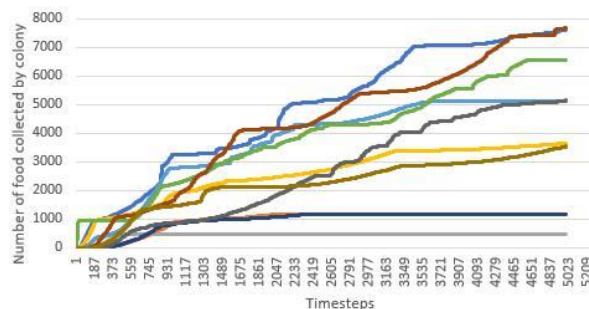

Fig. 2. Food collected over time during ten runs applying BM1 for pheromone foriaging.

A scenario was tested in which there was no compass, to identify whether a benchmark locate the nest direction, using only pheromone. The algorithm did not locate the nest, ants moved randomly, leaving pheromone everywhere, attracting other ants in the wrong directions (Fig. 3). Also obstacles (yellow) cause blockages (Fig. 4b)

Food distance from nest has various effects. With closer food, the pheromone trail will be stronger and it takes less time to get back to the nest. But longer trails have greater chance of other ants walking into them by accident, so further food may attract more ants that way (Fig. 4a). Two foods were discovered: a small food in the upper right is favoured compared to a larger food in the bottom left, because it is closer, the pheromone is stronger and all ants abandon the larger food until the pheromone evaporates and knowledge of its location is lost to the swarm.
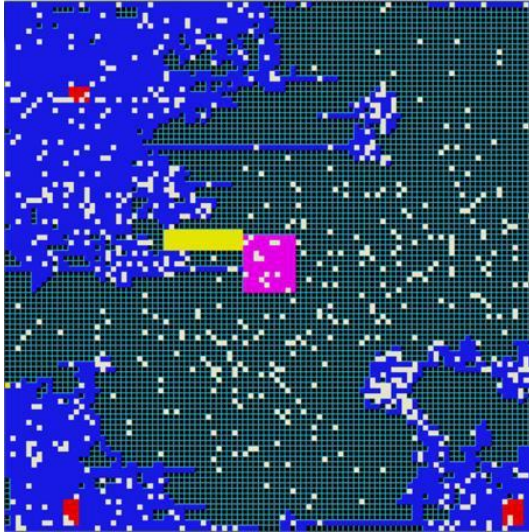

Fig. 3. The BM1 algorithm running without a compass sensor – ants have no way of finding the nest once food is discovered and pheromone is scattered randomly.
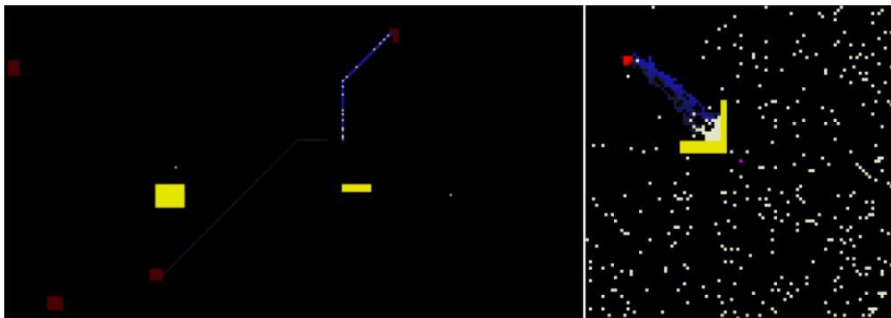

Fig. 4. (a) A large food supply (lower left) is abandoned in favour of a small food (upper right), closer to the nest with stronger pheromone. (b) a V shaped obstacle.

Each decision that an ant times is subject to probability so that it is always possible for an ant to do something unpredictable at any time. The effect of introducing a probability of random decisions.

4 Conclusions

This paper has investigated the benefits of neuroevolution compared to fixed topology ANN by testing how pheromone behaviour can evolve in both, in relation to a hard coded designed benchmark (BM1).

This paper has demonstrated neuroevolution applied to evolve pheromonome communication in simulated ant colonies. The core intelligence required to perform pheromone communication was summarised in form of the hard coded benchmark BM1, comprised of an IF block with 5 conditions and this behaviour was demonstrated to be effectively mimicked by the evolving ANN.

Acknowledgment

References

1. Collins, R. J., & Jefferson, D. (1990a). Antfarm: Towards simulated evolution. Computer Science Department, University of California.
2. Vaughan, N., (2015). Simulated Robotic Autonomous Agents with Motion Evolution. European Conference on Artificial Life ECAL. MIT Press, p27, http://www.cs.york.ac.uk/nature/ecal2015/late-breaking/ecal2015-late-breaking-proceedings.pdf
3. Stanley, K. O. (2004). Efficient evolution of neural networks through complexification (Doctoral dissertation).
4. Letendre K and Moses ME. Synergy in ant foraging strategies: Memory and communication alone and in combination. In Proceedings of the 15th Annual Conference Companion on
5. Genetic and Evolutionary Computation (GECCO '13 Companion), pages 41–48, New York, NY, 2013. ACM.
6. Panait, L., & Luke, S. (2004). Learning ant foraging behaviors. In proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems (ALIFE9) (pp. 575-580).