Cooperative Multi-step Behavior in an Evolved Robot Team

Ryan Walker
May 12, 2009
CS 81: Adaptive Robotics
rwalker1@swarthmore.edu

Abstract

Many tasks and situations may benefit from teams of robots working cooperatively to accomplish a task. Increasingly, robots are called upon to accomplish tasks that are too dangerous, too repetitive, or too costly for humans to complete. Developing a single generalized robot to complete such tasks may be expensive or inefficient. Teams of cooperative, autonomous robots hold the potential to efficiently complete challenges using a single evolved neural network.

Based upon previous success with teams of two robots, these experiments show that more complex group behavior can develop in a group of robots sharing an identical evolved 'brain'.

This set of experiments uses a team of three simulated robots in a biologically inspired scenario with opportunities for actions involving multiple steps and strategies. An underlying health system supplements the fitness function as a powerful shaper of behavior.

The gatherers shared the environment with three hand-coded predators, capable of quickly eliminating lone gatherers via health reduction. This reduction was mitigated or limited by the proximity of allies however. This mechanism ultimately led to a variety or grouping and herding behaviors that shaped final strategies.

Additionally, a home base that provided greater rewards for returned food served to create multiple step challenges that ultimately encouraged and produced novel state-based behavior.

1. Introduction

The results of this paper are an extension of a previous, simplified version of this experiment. The earlier version had fewer robots, a simplified health and fitness system, and little room or reward for complex behavior. While simplified, this earlier experiment showed that NEAT evolved robots could form rudimentary cooperative behavior and patterns to combat predators and find food.

The addition of more complex rules and objects in the environment is designed to test a neural network's capability to adapt and utilize more complex strategies across a larger team of robots. The simulation now allows robots to take multiple actions to acquire food, health and fitness, with more complex solutions rewarded through a multiple-component fitness function.

Evolutionary algorithms such as NEAT have the potential to be a powerful tool. They allow robots to find novel solutions to problems without direct human programming. Using a pre-defined base neural network of interconnected input, output, and hidden nodes, NEAT randomly adjusts the weights and connections between nodes, testing each solution in an environment with a defined fitness metric.

More successful weights are saved to pass on to future generations, building upon each other to maximize fitness. Through random changes, different strategies can be found and saved, ideally leading to a satisfactory solution.

In addition to randomly changing the weights associated with neural networks, NEAT randomly adds nodes and connections, slowly building a more complex network with increased computing capability. For these experiments, these probabilities were set at .07 for a new connection, and .05 for a new node.

The NEAT algorithm itself also allows for speciation by tracking the performance of related networks. This ensures that solutions with novel ideas have a chance to thrive despite temporary sub-optimal performance. This measure helps to avoid the local maxima problem intrinsic to randomized fitness-based systems.

2. Related Work

As in my earlier experiments, the concept and design of the experiment is influenced by the robot duel experiment in Stanley et al. This paper utilized NEAT in a competitive environment to encourage higher fitness and more novel solutions.

The paper also lays down the basic experimental framework of this sort of experiment. Notions such as the gradual increase in complexity over time were used successfully in the Stanley paper and carry over to this situation as well.

The experiments differ however, in complexity and motivation. Stanley pits two networks in an arms race of sorts in which both robots seek to become better than the other in a somewhat simple task. This experiment instead seeks to reward cooperation and versatility in a single evolved network. The scenario and variables are also more complex, with multiple environmental interactions and a fitness function derived from multiple components.

Potter et al. also inspired the group mechanics in the experiment. While Potter et al. relied on multiple specialized brains to accomplish a task, the paper was valuable in its discussion on encouraging cooperation between robots. The paper was highly focused on specialization as a solution to multi-robot problems, whereas this experiment examines a task where generality and adaptability are favored. The relationship between difficulty and cooperation discussed in the paper was helpful as my results also show that difficulty alone is not a primary motivator towards cooperation.

The Parker paper did not serve as a specific guide for this experiment, but it provided motivation and ideas when designing the simulation. The paper uses an un-evolved algorithm to determine motivation and action. The resulting behaviors and conditions however, helped me to create the challenges in the environment. The paper also discusses the importance of sensor input and communication in cooperative teams, concepts that helped to inspire the neural net and design of the simulation.

Additionally, as an example of

hard-coded teamwork, it serves as an interesting reference and comparison to behavior achieved in this experiment.

3. Experimental Rules and Environment

Each experiment involved the same essential setup with experimental parameters. All feature a team of three evolved gatherers in competition with a set of environmental rules. The evolved experimental robots must gather food while avoiding hand-coded predatory robots that can cause harm. A homogenous evolved neural network controls the gatherers. Each robot utilizes a copy of the most recent neural network to govern their actions. The alternative would be a heterogeneous system in which each gatherer robot evolves their controlling network independently of the others. The one-network system demands a generalized brain that can be applied successfully to three robots working cooperatively.

The neural network that controlled all three gatherers was evolved using the NEAT algorithm. The evolution was initialized with a hidden layer of size two or three depending on the experiment. Evolutions ran for 60 to 120 generations, with each generation consisting of ten chromosomes.

Each variation underwent two trials of 1000 steps to evaluate fitness. All robots were reset to their initial health state and pose for each trial, and the lights were randomly scattered about the environment. Initial experimentation used three trials with different focuses: food gathering, predator contact, and mixed. Results however, were less successful, and robot behavior was more erratic, prompting a return to uniform trials.

The simulated environment consisted of a square room filled with sixteen randomly scattered light sources that acted as 'food' for the gathering robots. In the upper-right corner of the room was a cyan box that acts as a home base for the robots. The three gatherers started each trial arrayed about the base in static locations towards the middle of the environment.

Three predator robots started opposite the gatherers. The predators were hand-coded to wander the map without eating the light-food. When predators detected a gatherer, they would make a general effort to pursue, but due to the small room size and variety of light sources, they were not overly aggressive. They were equipped with short range light sensors, which were easily confused by the initial light sources. As food was eaten however, the predators became more efficient at finding prey. The predators served an important role as a primary, semi-random challenge to the gatherers, motivating group behavior and health recovery.

The experimental robots were all simulated Pioneer robots, each equipped with front light sensors and an identifying colored light source. The gatherers also had basic front sonar sensors, averaged into front-left and front-right values. Light sources thus act as the primary form of sensory interaction in the environment, with colored lights according to role. Gatherer robots featured blue lights on their backs. The predators were equipped with red lights. Food sources were identified as green lights that disappeared when eaten. Home base was a cyan box illuminated slightly.

3b. Health System

The rules of the simulation were enforced via a health system that affected both predator and gatherer robots. Internal counters tracked health values of all robots, with different interactions increasing or decreasing the health of individual robots. Gatherers started each trial with 200 health, while predators had 120. The most dramatic health change occurred when predator and gatherer robots met. If one gatherer came within a small distance (< 1 robot length) of one or more predators, then the gatherer suffered dramatically, losing 30 health per time step for every predator in maintained proximity. In order to foster cooperative behavior however, if two or more gatherers were in proximity to the same predator then the penalties were far less. Gatherers lost only 10 health per time step while predators suffered 25. Figure 1 displays the gatherers in tight formation, safe from the otherwise deadly predator. This became a common tactic to avoid health loss.

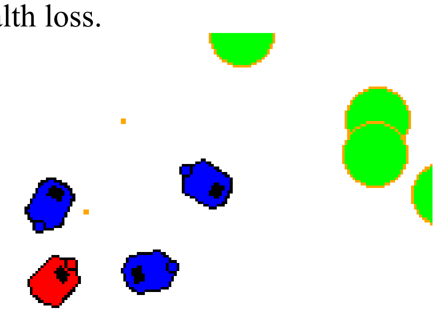


Figure 1

If, at the end of a time step, a robot's health was less than or equal to zero, it was deactivated for the remainder of the trial. The robot remained in the environment, but motor functions were halted and the identifying lights were turned off.

This system was in place in previous experiments and successfully allowed the evolved robots a chance to develop basic cooperative behavior.

In addition to this interaction, gatherers now had two other opportunities to affect their health through interactions with food and the home. Instead of simply consuming food, the gatherers in the simulation were forced to make a choice when they picked up food.

Depending on a neural net output, the robots had a choice to either 'hold' or 'eat' food with various consequences. Immediately eating food benefited the individual by restoring 20 health. The robots however, could 'choose' to hold food without eating it for no immediate health gain. If they held the food and happened to come close to the home base then they deposited their food and were instantly restored to full health if there was food in the base. This allowed the robots to maintain high individual and group health if they accomplished a more complex multi-step task.

3c. Fitness function

The health system governed each individual trial, but does not directly affect the fitness function greatly, the measure of fitness for each chromosome. In previous version of this experiment, fitness of a species was determined directly by food consumed. This simplistic fitness function rewarded repetitive and basic food-finding behavior, but failed to directly reward complex behavior.

In this series of experiments, fitness was determined by a series of measures that progressively rewarded more complex behavior. Exact values varied depending on the experiment, but in general the fitness function was defined as:

Eat food immediately:
health gain, +1 fitness

Food held*: +3-15 fitness*

Food returned to home*: health gain, +10-50 fitness*

The gatherers also gained fitness if the mean health of all three was above 190, rewarding generally high, sustained health.

This combined health and fitness system allows for gains in direct and indirect ways. Eating food immediately is an easy, low effort task with a small direct reward. The health gained from eating food however, may allow a gatherer to stay alive longer to gather more food, for instance.

Multi-step behavior with longer-term benefits, such as returning to base with food, proved difficult to explicitly reward, so each step of the process grants higher rewards. This ensured that robots exhibiting more complex behavior were more likely to pass these traits on.

3d. Neural Network

In all iterations and experiments, the robots were governed by an evolved neural-network. In all cases, the inputs and outputs of the network were identical. As seen in figure 2 (a fully evolved network), the network started with twelve inputs and three outputs, with a hidden layer of two nodes. Input nodes included sensory information from sonar and RGB values on both left and right sides. Additional sensors provide specific information to the robot relevant to the tasks.

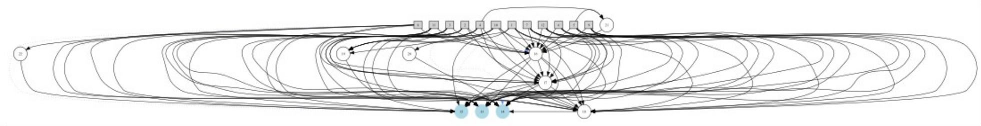


Figure 2

One node corresponds to the robots current health, enabling the robot to decide how to treat food and predators. Two additional nodes provide distance and dTheta to home base. While this knowledge may predispose the robot to interact with its home, it is valuable to provide directions. Early prototypes of the experiment used more lights to signify the home base, but the abundance of different light sources made it even more difficult for the robot to discern objects in its environment. Providing this information, which would not be unrealistic in many practical situations, allowed the robots to make distinctive moves towards home. This ultimately became key to more complex behaviors and patterns.

The three output nodes controlled the robots' interactions with the environment. Two nodes were linked to left and right motor controls. A third neural net output was used as a binary decision indicator as described in the health and fitness rules, queried whenever a robot ate food.

The neural net used a tanh activation function, with all inputs and outputs scaled from -1 to 1. This allowed for outputs that were easily translated into left and right motor commands, scaled for forwards and backwards movement.

Ultimately, the networks evolved many more connections and nodes to handle the multiple states and calculations required to succeed in the experiment. Figure 2 displays the final neural network from a moderately successful trial that utilized a pattern/group proximity based solution to finding food.

5. Experiments

The final experiments were divided into two general groups based upon fitness function and evolutionary parameters. Within each group, multiple evolutions using the same general parameters were tested.

Group 1 had the same environmental and health features as group 2, but the fitness function varied slightly between the two groups. Group 1 rewarded eaten food, held food, and food returned to base with +1, +3, and +10 respectively. Group 2 placed a much higher fitness emphasis on returning to base and holding with corresponding fitness rewards of +1, +10, and +50. This higher reward was designed to heavily promote more complex behavior so as to test whether the network would alter its behavior or priorities. Within each of these groups, three finalized evolutions were run from 80 – 120 evolutions.

Both groups also featured several preliminary evolutions of fewer generations in order to tune and refine NEAT parameters. Fitness graphs are an unhelpful indicator of success, as fitness functions varied greatly between the groups. Additionally, the true gauge of success was the complexity of the behavior, a metric difficult to judge without description.

Observations

Three representative evolutions are described for each group as described above. Other evolutions were either: very similar to one described, preliminary tests, or displayed little complex behavior of interest.

Trial 1A
Trial A in group 1 evolved an individualistic strategy, with each individual robot capable of autonomous navigation and avoidance of predators. This is in opposition to some later solutions in this group that relied primarily upon pattern-based motion. The network quickly evolved basic navigation skills including forward and backward movement, and wall avoidance. As the species advanced however, it struggled to integrate more advanced behavior seamlessly into its operation. The simulated robots often became stuck in a rotating ‘waiting state’ in between more complex actions.

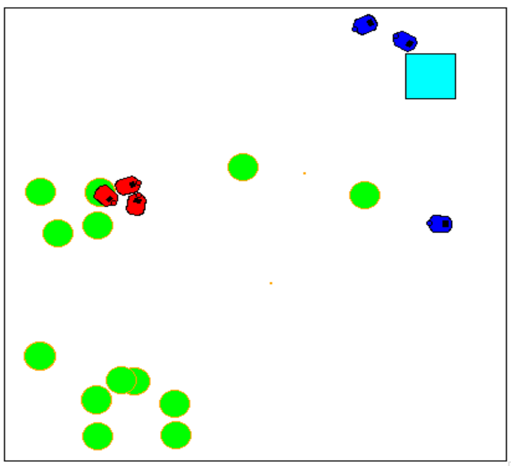
The robots ultimately learned to hold food, but upon collecting food tended to spin, occasionally recognizing a stimulus and acting quickly. Robots that held food occasionally moved towards base quickly after a period of waiting. Overall, this group was generally successful but exhibited little group interaction.

Trial 1B
The second evolution evolved a radically different strategy for dealing with food and predators. Instead of developing tuned individual reactions, the neural network became pattern based, with robots moving about each other, always attempting to maintain sight of another friendly robot.

To facilitate this, even highly evolved generations used backwards motion so that their forward-located light sensors were able to maintain contact with teammates. Initially, the gatherers moved around each other in a rotating triangle. This behavior is likely motivated by the health system that greatly rewards gatherer proximity to avoid predation.

After the robots generally cleared the middle area of food, they broke pattern and each attempted to move to a seemingly random corner. This seems to be a direct strategy to return food to the home base, as the chance of picking the corner with the home base is 25-75% depending on many gatherers died en route or moved to duplicate corners. Figure 3 shows a particularly successful trial in which all three robots moved in the correct general direction after eating a few units of food. Only the upper two are close enough to potentially drop off food.

Figure 3



Trial 1C

Trial C also utilized backwards-moving robots to maintain visual group contact. The third evolution evolved a much more intelligent form of the second's behavior however. Like the previous evolution, the group of gatherers moved inwards and together to quickly gather initial food, which can make up the majority depending on the random distribution. After gathering food however, the gatherers do not blindly run to the corners to become stuck.

Instead, the robots utilize individual reactions to move outwards, circling more towards the edge. When a robot comes within close visual distance of the home base, it rapidly moves into contact, effectively returning food to the base. Additionally, gatherers that encountered, but survived predator encounters often seemed to return to the middle seeking food.

These behaviors represent a more complex combination of the previous evolutions, in which the neural network has evolved state-based reactions presumably linked to the state determining inputs. While the behavior was unrefined at times, the robots were capable gatherers, fulfilling the task through a combination of efficient group-influenced patterns and individual reactions. Figure 4 displays a trial of 1C in which the robots, having consumed central food, have moved to the periphery with varying success.

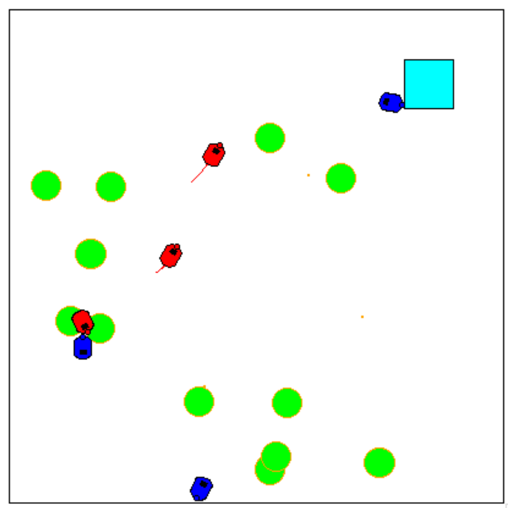


Figure 4

Trial 2A

This evolutionary trial represents the first of group 2, in which the fitness function more heavily rewarded holding and returning food. As in previous trials, the gatherers adopted a backwards motion, moving about in a tight cluster when in visual range of each other. When out of view of the other robots however, lone robots strayed farther towards the edge, usually returning to center in reaction to a nearby gatherer or predator. This strategy ensures that robots are generally protected, yet utilizes greater exploration.

Trial 2B

The second main evolution of the second group took an alternative strategy, utilizing forward motion and good general reactions. Instead of relying on a pattern to return home, this group relies on random group movements combined with sharp individual actions to consistently maintain health.

While this group rarely returned to the base with food, when it did it usually did so with multiple robots due to group following.

Trial 2C

The last notable trial of group 2 was one of the most successful in both score and behavior. In this evolution, the evolved brain uses more complex patterns and state-determined actions.

The gatherers employed a loose group-based behavior, wandering about after food yet seeking general proximity to allies. After collecting food, robots individually looped definitively towards base in an arc. While the robots occasionally became stuck on the home itself, those that did not continued to wander about the edge and 'eat' visible food.

6. Analysis and Conclusions

The trials exhibit several key strategies to maximize fitness. Ranging from basic patterns to advanced decisions, the behavior patterns exhibited in the different trials demonstrate that a single brain applied to a team can produce complex behavior.

One of the most striking and common behaviors was the tendency to group and move together. In a direct response to the threat of predation, the gatherers often moved in response to one another. This behavior was so important in the simulation that many of the evolutions moved backwards so as to stay in constant visual contact with their teammates. Oftentimes, by simply attempting to stay in visual range of another gatherer, the robots were able to move across the screen in a flock-like behavior, collecting food and moving towards base as a unit. Figure 5 displays the gatherers in formation, having gathered food and moving towards the home as a unit.

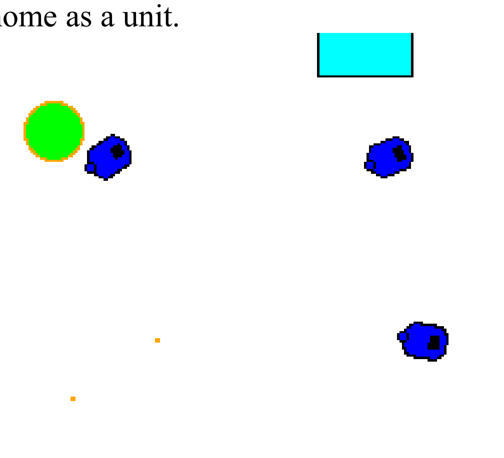


Figure 5

This mirrors biological scenarios in which weak individuals group and move together for food or protection. This experiment demonstrates that such behavior is relatively simply to attain if it is in every individual's best interest to maintain communication and proximity with their close partners.

Alternatively, evolutions 1A and 2B focused heavily on individual reactions to the environment, with group movement a lesser concern. This approach led to individuals that dodged predators and chased food much more specifically and accurately. This may have been a necessity due to the loss of group protection.

The best examples of complex behavior however, were demonstrated in the 1C and 2C. In these evolutions it is apparent that the gatherers are not reacting purely to direct stimulus, but status as well. In these species, the robots acted as a group to secure initial food but, once holding food, the robots were able to change to a new pattern, either following the edge in 1C, or looping directly towards the home in 2C. Both are examples of multi-step behavior and basic decision-making.

This complexity demonstrates the flexibility and strength of evolved neural networks. While the behavior was inconsistent in many of the trials, the successes in 2C show the potential strength of a homogenous control system for groups of robots. Despite using duplicate copies of the same network, the gatherers were able to simultaneously group for defense, explore, avoid enemies, and complete state-based navigation calculations to find the home base.

Achieving this behavior was difficult however, for many evolutions tended towards more simplistic, less responsive patterns, simply relying on probability and consistency to score high fitness. The implementation of NEAT used also does not support recurrent connections between nodes, limiting the computational ability, and therefore capabilities of the evolved networks.

Despite these issues, the experiment demonstrates that a single evolved brain can be used across multiple robots to achieve behavior that benefits both the individual and the group.

References

[Parker, 1999] L. E. Parker. Adaptive heterogeneous multi-robot teams. *Neurocomputing,* 28, 1999.

[Potter, 2001] M. Potter, L. Meeden, and A. Schultz. Heterogeneity in the coevolved behaviors of mobile robots: The emergence of specialists. *Seventh International Conference on Artificial Intelligence*, pages 1337–1343, Seattle, Washington, 2001.

[Stanley, 2004] Stanley, K. O., and Miikkulainen, R. Competitive coevolution through evolutionary complexification. *Journal of Artificial Intelligence Research*, 21:63–100, 2004.