Comparative Evolutions of Swarm Communication

Amy Jin and Murphy Austin

May 9, 2014

Abstract

In this paper, we attempt to replicate an experiment by Marocco and Nolfi in which a grounded communication system was evolved in robot swarms. We hoped to develop bidirectional communication without any hard-coding by testing neural networks on a navigation task in which effective communication was essential to achieving the best score. In order to build a communication system, controllers would have to associate certain frequency values in an undifferentiated, continuous range with meaningful information. Building off of previous unsuccessful attempts to replicate this experiment, we evolved neural networks with fixed topologies in our own custom-built simulation. Overall, we achieved fitness scores comparable to the original experiment, although or recurrent networks performed significantly worse. A qualitative examination of our evolved populations, moreover, shows that we did not successfully evolve bidirectional communication. Frequency output was largely limited to binary values that aided with robot navigation. We conclude that our robots learned to effectively navigate the environment, but did not develop more complex communication behavior.

1 Introduction

Evolving communication in robots has proven to be a consistently difficult task [2] [4]. Evolving an effective system among swarms is especially useful, as it can drastically improve the performance of group tasks. Moreover, communicative intelligence is essential to understanding how biological organisms socially interact and develop useful symbolic reasoning.

1.1 Related Work

In their 2007 article, Davide Marocco and Stefano Nolfi describe a remarkably successful experiment in developing effective communication in a homogeneous robot swarm. The robots they evolved were able to link different frequencies to relevant motor actions, and use these meaningfully communicative frequencies in order to better solve a navigation task. This communication was bidirectional: each robot played both a speaking and hearing role. Each robot gathered frequency information from its environment in order to better direct motor actions, while concurrently producing a frequency to provide information to other robots. Perhaps the most impressive aspect of their experiment was the fact that robots evolved a meaningful communication system without hard-coding. Evolution began with an undifferentiated, continuous range of possible frequency outputs, and evolved individuals learned both to consistently produce meaningful frequencies and to use information from other robots to better solve the task.

In the spring of 2012, Kristen Allen and Peter Ballen attempted to recreate this same experiment for their final project in CS81 [1]. They used Pyrobot as their simulation platform and NeuralEvolution of Augmenting Topologies (NEAT) [5] to evolve their populations. The experiment failed to produce individuals with significantly high fitness scores or meaningful communication systems. The authors persuasively argue that NEAT is fundamentally unsuited to this particular navigation task because its dual advantages of complexification and speciation actually hinder the performance. Navigation in the environment is not complex, but requires a very fine-tuned set of node connections. NEAT fails to discover these correct weights due to its tendency to mutate morphology. In addition, NEAT's speciation maintains diversity but de-emphasizes fitness scores, while Marocco and Nolfi's evolutionary algorithm was based only on individual fitness, regardless of diversity. Both of these NEAT features prevented evolution from discovering the first, most essential components of the task - successful navigation and interpretation of inputs - without which a meaningful communication system could not be built. As Allen and Ballen write, "individual fitness and the implied ability to usefully categorize environmental inputs are prerequisites to communicating meaningful information to fellow robots" [1, p. 8].

We hoped to improve on Ballen and Allen's approach in several ways. First, we made our evolutionary algorithm more similar to the original paper by evolving fixed topologies and eliminating speciation. While we did use NEAT, we deactivated its ability to add or remove nodes so that it would only evolve new behaviors through modifying connection weights.

In addition, we built our own simulation platform using the Python graphics library. By hand-coding the environment, we could have direct control over all sensors, measurements, and calculations, and thus more closely reproduce the original experiment. For example, Allen and Ballen had to approximate sonar values based on the different sensor configuration of the Explorer robot, and these sensor values exhibited the noisy interference of the Pyrobot environment. More significantly, our simulation was much simpler than Pyrobot, and therefore took much less time to evaluate and evolve. This allowed us to reasonably evolve a much greater volume of material. We evolved population sizes of 100 with three trials of 1,000 timesteps per evaluation; Allen and Ballen used population sizes of either 20 or 100, evaluated with three trials of only 200 timesteps. The length of our trials and size of our population made it more likely that evolution could discover and reward desirable behavior.

2 Experimental Setup

Our robots navigated in a bounded, 270x270 cm environment with two circular target areas on opposite sides of the arena. Each of the four robots began in a random location outside the target areas, emitting a random frequency between [0.0, 1.0]. Robots were represented by simple circles with 11 cm diameters; they could detect frequencies up to 100 cm away and had a sonar range of 15 cm. Each robot is equipped with communication, sonar, and target sensors, implemented as follows:

1. **Frequency**: Four communication sensors detect frequencies emitted by other robots from in front of, behind, to the left of, and to the right of the robot. Each sensor has a range of 100 cm and can hear only one frequency at a time. If more than one robot is detected within range of a single sensor, the sensor returns the frequency of the closest robot (shown in Figure 1). If no robots are within range, the sensor returns 0.0. In addition to the four communication sensors, each robot can sense the frequency it is currently producing.

2. **Sonar**: Eight sonar sensors detect obstacles around the body of the robot. The sonar returns a value between [0.0, 1.0]. A value of zero indicates no obstacle closer than 15cm, while a value of one indicates a direct collision.

3. **Target**: The target sensor is a binary reading returning 0 if the robot is not on a target and 1 if it is. There is no differentiation between the two target areas.

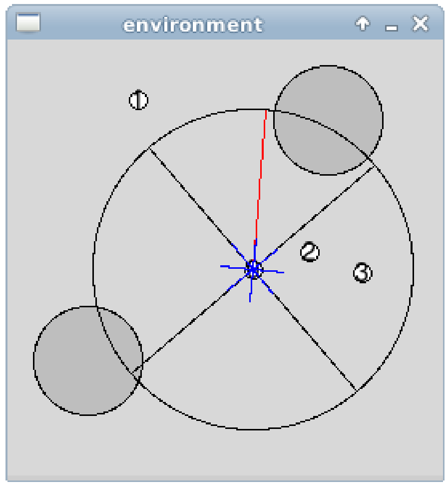

Figure 1: The simulated environment. The four smaller white circles are the robots, and the larger two grey circles are the target areas. For simplicity, sensor ranges have only been drawn for Robot 0. The red line is the robot's heading, while the eight blue lines are the robot's sonar sensors. The black circle is the range for the frequency sensors, which is divided up into a four quadrants: front, left, back, and right. In the right quadrant, the frequency sensor will return the frequency emitted by Robot 2, since Robot 2 is closer than Robot 3.

Robots produced three motor outputs in each timestep: two floating-point values between [-1.0, 1.0] controlling the motion of the left and right wheel, and one frequency output between 0 and 1.

We used NEAT to evolve our population, but disabled most of its features in order to better reproduce the original experiments genetic algorithm. Marocco and Nolfi implemented an evolutionary approach using fixed topology neural networks and no cross-breeding, achieving evolution solely via the mutation of weights. They hand-coded multi-variable activation functions for the networks nodes. New generations were produced simply by selecting the 20 best performing individuals, replicating each five times, and replacing "2% of their bits with a randomly selected value" [3, p. 58].

NEAT provided us with a robust, ready-made evolutionary algorithm, and we were able to configure its parameters to closely align with the original experiment. We disabled NEAT's ability to add or remove nodes so that our evolution would also progress solely through manipulating connections. Moreover, the probability of mutation connections was set to a very low value, reflecting the low levels of mutation between each generation in the original paper. These changes nullified NEAT's generally advantageous ability to complexify and speciate. As discussed in the introduction, we made these modifications due to past failures to evolve robot communication through NEAT for this task.

Each individual in the population was an ANN brain used for robot control. The same brain controlled all four robots, so that fitness was calculated for the entire homogeneous swarm. The robots were tested in the environment for three 100-second trials, which collectively constituted a single evolutionary evalution. We ranked the individuals using the following fitness function from Marocco and Nolfi's paper:

$$f(t) = \sum_{0}^{\max t} (0.25 \cdot n_t - 1.00 m_t)$$

where n is the number of robots located in a target area at time t, m is the the number of extra robots (i.e., a third or fourth robot) in one target at time t, and $\max t$ is the total number of time steps. In order to collapse fitness scores to a value between [0.0, 1.0], we divided by $\max t$ before rescaling and shifting the score. The fitness function provides the most reward when two robots are positioned in each target area (scaled fitness = 1.0), provides some reward for single robots finding the target area (if one robot is in each target area, scaled fitness = 0.8), and penalizes for too many robots resting in the same target area (for four robots in one area, scaled fitness = 0.0).

Note that since fitness is accumulated over time, a fitness value of 1.0 is impossible to achieve because it would require that the robots spend their entire lifetimes divided evenly between target areas. As mentioned above, we ensured that no robot began on a target area when implementing our environment.

In keeping with Marocco and Nolfi, we evolved two different network topologies: one strictly feed-forward with no hidden nodes, and one with recurrent connections enabled and two hidden nodes (Figure 2). The authors found that the topology with hidden nodes, which they called "non-reactive," achieved greater fitness, likely due to its recursive memory.

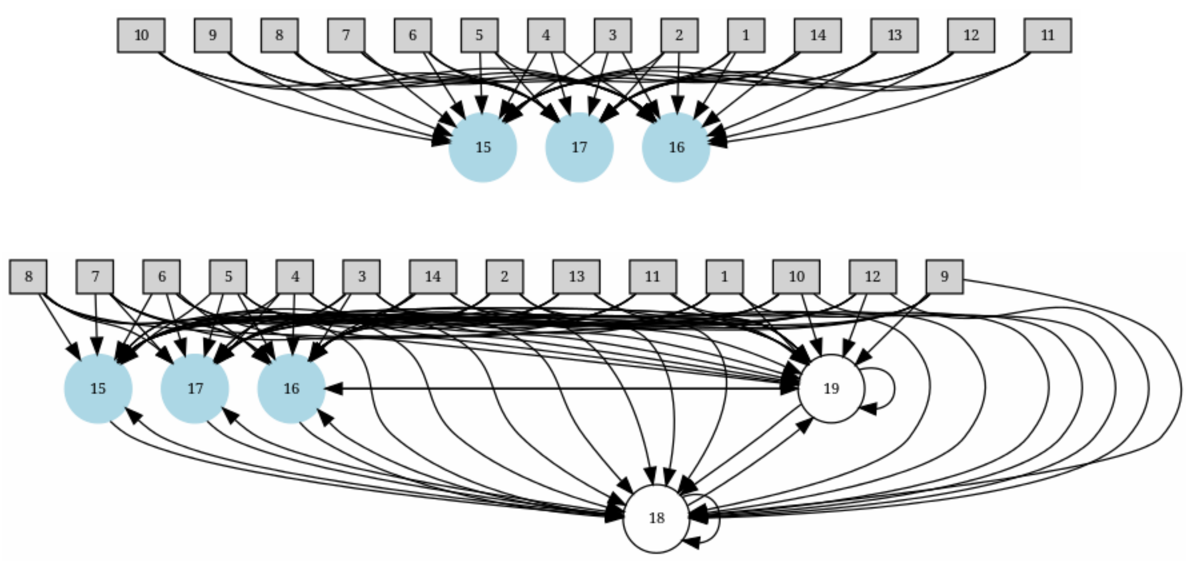

Figure 2: The feed-forward network with no hidden nodes (top) and the recurrent network with two hidden nodes (bottom). Inputs 1-8 are the sonar sensor values, inputs 9-12 are the four frequency sensors, input 13 is the target sensor, and input 14 is the robot's own frequency emission.

3 Results

For a thorough evaluation of our results, we compare them to those of Marocco and Nolfi. We do this in two ways: first, by evaluating relative fitness scores between varying experimental conditions, and then second, by a direct comparison of fitness values. Since the underlying goal of the experiment is to examine whether or not evolved communication methods can improve a collective navigation task, we compare results between three conditions:

1. **Normal**: Robots are evolved *with* communication, and then tested *with* communication
2. **Deprived**: Robots are evolved *with* communication, and then tested *without* communication, and
3. **Silent**: Robot are evolved *without* communication, and then tested *without* communication.

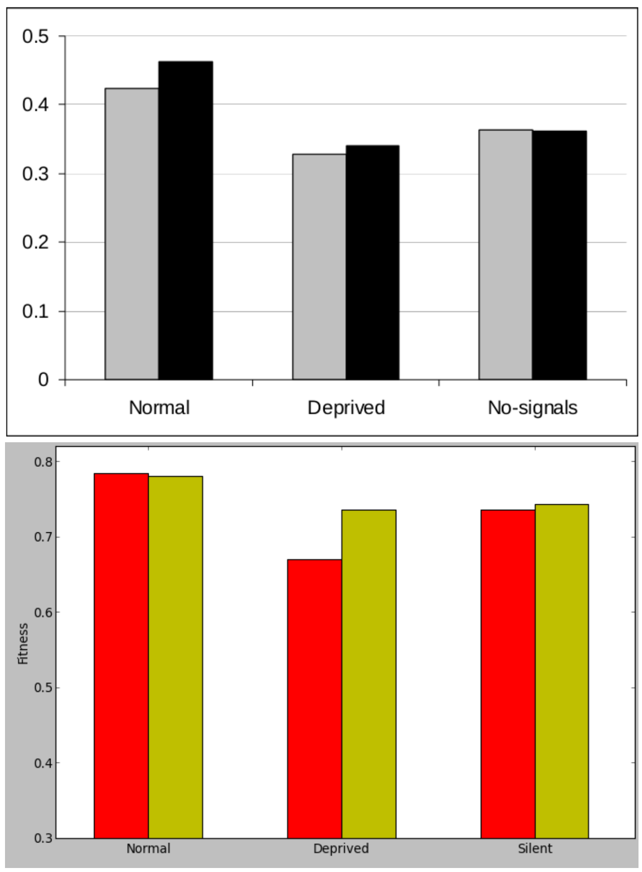

Figure 3: Average fitness of most fit individuals from last 20 evolutions of generation. Marocco and Nolfi's experimental results are on top, while our results are below. In Marocco and Nolfi's graph, grey and black histograms represent average fitness under feed-forward and recurrent networks, respectively. In our results, red and yellow histograms represent average fitness under feed-forward and recurrent networks, respectively. Fitness values are both normalized in the range [0.0, 1.0], and tests are run in both experiments for a total of 1000 trials.

First, let us look at the results obtained by running evolution on the feed-forward network with no hidden nodes. Fitness values achieved under the "Normal" experimental condition are higher than those achieved in the other two conditions, a strong indication that communication improves the robots' abilities to complete the navigation task. The fitness values achieved in the "Deprived" condition (where robots are then tested without the ability to communicate) confirm this, as the average fitness is significantly worse ($p < .0001$) [1].

Perhaps a more interesting experimental condition is the "Silent" condition. Since the fitness values achieved here are also significantly worse than in the "Normal" condition ($p = .0002$), this shows that evolution of communication is essential in improving the robots' abilities to better complete the task. In other words, evolution without communication noticeably limits the robots abilities to find a more optimal problem-solving strategy. However, these robots still perform better than the robots tested in the "Deprived" condition, which goes to show that the robots evolved in the "Normal" condition rely heavily on their evolved communication in order to achieve their goal.

Results of the recurrent networks follow the same trend, but with an interesting difference. Whereas the results between experimental conditions of the feed-forward networks statistically differ from one another, they do not for recurrent networks. This differs from Marocco and Nolfi's results, as they found statistically significant differences within their recurrent networks. There are a couple more distinctions to make between the data of our experiment and those of Marocco and Nolfi's. First, a direct comparison of fitness values shows that our fitness values are consistently about 3.5 higher. Since relative fitness trends are similar between the two experiments, we hypothesize that this is due to an inconsistency in how the authors normalized their fitness values. Also, the fitness values plotted in Marocco and Nolfi's graph are the averages of the best fit individuals from the last generation of evolution, where evolution was replicated 10 times. Since we only ran one evolution, our above values are calculated from the best fit individual from each of the last 20 generations of evolution, each tested 50 times. Both tests total to 1000 trials.

3.1 Communicative Behavior

Upon closer inspection of the robots' evolved communication systems, we observed the following frequency behavior:

- A frequency emitted of around 1.0, when a robot is outside of a target area and all no object or wall is within sensor range
- A frequency emitted of around 0.0, when a robot is inside of a target area or is traveling along a wall

The communication system evolved by the robots in our experiment is much less complex than in Nolfi and Marocco's. In the paper, four distinct frequency ranges are observed, while in our replicated experiment, robots mostly emitted one of the two binary values above. Even so, we were able to observe the effect of these emitted frequencies on the behavior of other robots. If a robot heard a high frequency of around 1.0, it would immediately change its path of exploration so that it would move opposite from the direction in which the emitting robot was exploring. Doing so increases the efficiency of the exploration of the environment, so that the robots are not

[1] p-values are two-tailed, calculated using an unpaired two-sample t-test.

redundantly following the same trajectory. Furthermore, it is more likely that the robots will end up dividing themselves amongst the two targets, since the targets are placed in opposite areas of the environment. Hearing a frequency of 0.0 was equivalent to not hearing any frequency at all, and would render the emitting robot invisible. This behavior is appropriate for when a robot is located on a target, since it does not want to repel an exploring robot.

Despite the fact that fewer distinct frequency ranges were defined, this simple system was still able to aid the robots in more effectively exploring the environment. We consider a simplified environment, in order to more clearly illustrate the effects of these communication strategies.

3.2 Communication in a Simplified Environment

Using our best evolved individuals from above, we now test them in a reduced environment, in which there lies only one target area and includes only two robots. Since there are only two interacting agents, there is no ambiguity as to which robot is being affected at any given time (Figure 4).

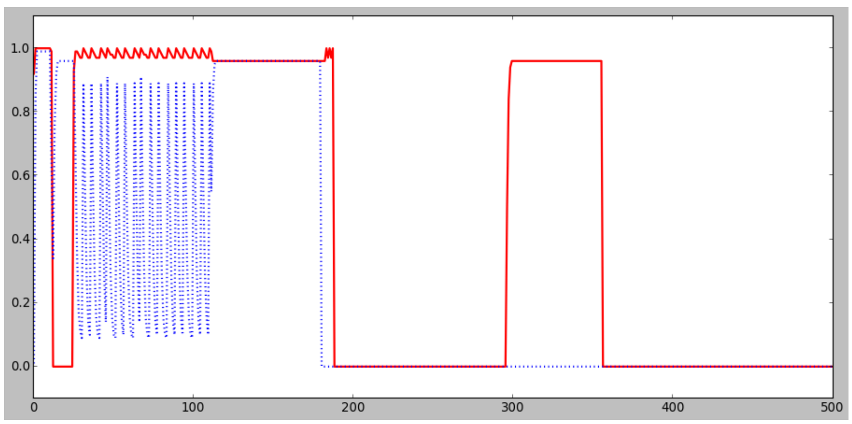

Figure 4: At the beginning of the trial, Robot 1 (dotted blue line) and Robot 2 (red line) are outside the target area and emitting a non-zero frequency, which allows them to explore opposite directions (Figure 5, left). Just before time step 200, Robot 1 reaches the target area and changes its frequency to 0.0. Shortly after, Robot 2 finds a wall, and begins following it while producing no frequency (Figure 5, center). Robot 2 leaves the wall, which causes its frequency to jump to 1.0, and approaches the target area. Because Robot 1 is producing a frequency of 0.0, Robot 2 cannot detect it and is not repelled. Although Robot 1 can hear Robot 2, it has learned to stay on the target area regardless of other robots. Both robots remain on the target area, producing a frequency of 0.0, for the remainder of the trial (Figure 5, right).

Figure 6 shows a different communication method found in an early generation of one our evolutions. In this case, the robot in the the target area periodically produces a distinctive frequency when it detects another robot. This is promising behavior - it suggests the possibility for other robots to “zero in” on this frequency and head towards the target area. However, the neural controller in this case has not yet discovered how to use this new frequency to increase performance. Significantly, this communicative behavior is not reproduced in future generations.

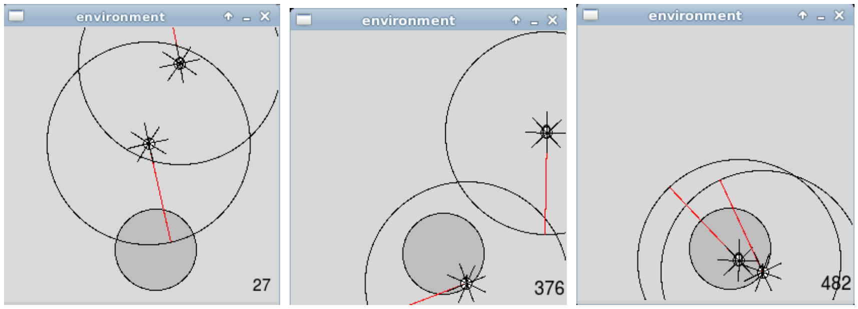

Figure 5: Our most successfully evolved neural controller, operating in our simplified environment. Three different situations arise based on what frequencies are heard and emitted (from left to right): avoidance, wall-following, and target cohabitation. The black circles indicate frequency range.

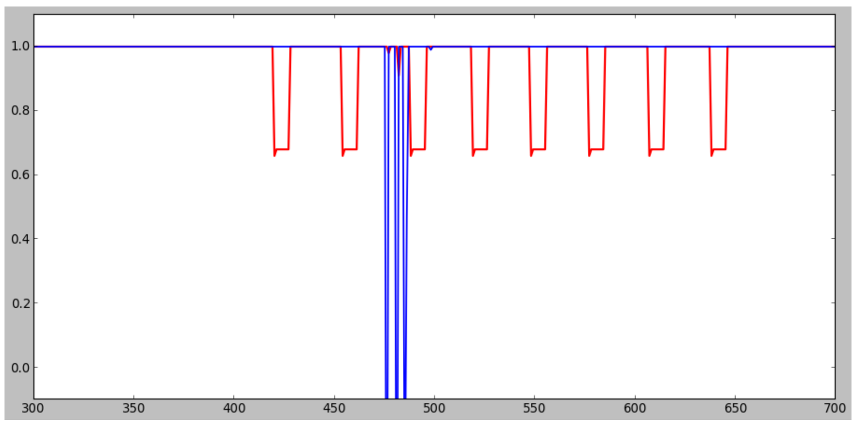

Figure 6: Robot 1 and Robot 2 explore the environment, both producing a frequency of 1.0. Robot 2 reaches the target and begins spinning, but does not change its frequency. Shortly after timestep 400, Robot 1 enters Robot 2s hearing range. While it is in range, Robot 2 emits a frequency around 0.68 whenever its front quandrant is facing Robot 2 (figure 7). This frequency does not seem to impact Robot 1s movement. Robot 1 does not find the target area continues past frequency range.

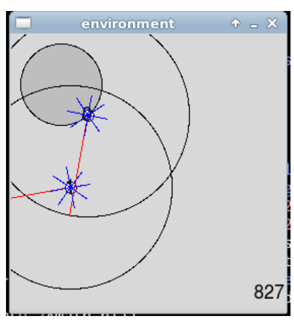

Figure 7: A robot locates a target area, and then begins to spin. When its front facing frequency sensor detects the other robot, the spinning robot emits a new frequency of 0.68.

4 Conclusion

Our results suggest that successful navigation behavior is significantly different from communication behavior. The best-performing individuals in our experiments used only binary communication:

a high frequency repelled neighboring robots, while a frequency of 0.0 did not. This strategy did significantly improve fitness because it allowed robots to spread out, find a target, and then become "invisible" to other robots so as not to repel them from the target area. The robots, however, had no way of knowing whether a target area was too full, and thus their fitness was limited. As our example of communication from an early evolutionary generation demonstrates, 4 useful information could be communicated, but robots did not have the navigational ability to utilize it, and the push for greater fitness eliminated this communication before it could improve results.

At this point in our results, our population seems to have discovered a deceptive solution. The robots utilize frequencies to navigate better but have not developed meaningful communication. A more effective strategy utilizing bidirectional communication, would have to change the frequency behavior of the robots, and would likely cause a temporary decrease in navigational fitness. A novelty search, or even NEAT with speciation, could be useful at this point in finding new communication behaviors.

Indeed, a closer examination of Marocco and Nolfi's experiment reveals that their evolutionary approach may not be the most practical. It seems strange that their experiment evolves a population of 100 individuals for 100 generations, using 20 trials per evolutionary evaluation, simply to change the connection weights of a neural network with no hidden nodes. Previous work by CS81 students has shown the NEAT's maintenance of diversity and complexifying topology prevented neural controllers from developing simple navigational capabilities in this environment. Once these capabilities have been found, however, NEAT or Novelty Search may be more effective in building meaningful communication.

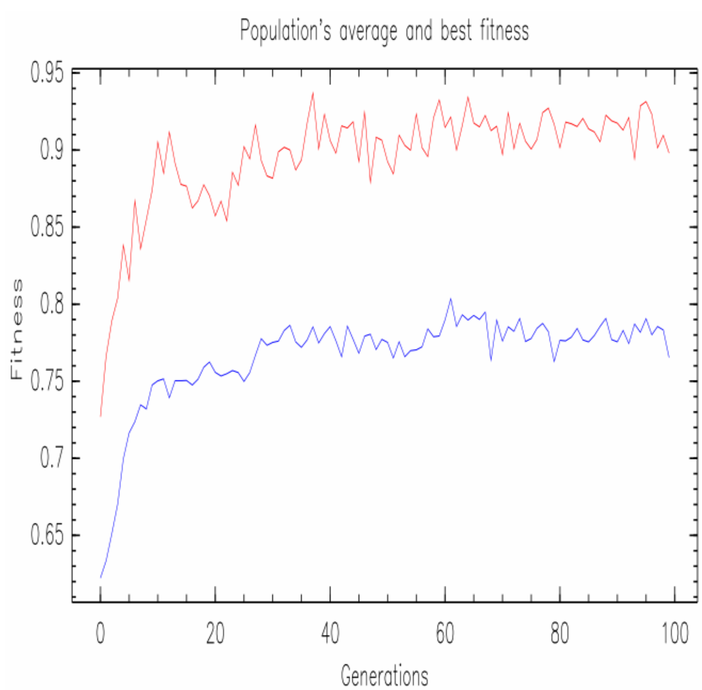

Figure 8: Average and best individual fitness for the feed-forward network with no hidden nodes.

One factor that may have detracted from our results is the number of trials each individual was tested on during an evolutionary evaluation. Marocco and Nolfi calculated an individuals fitness score based on the average fitness over 20 trials of 1,000 timesteps; for computational efficiency, we only used 3 trials of 1,000 timesteps. Because robots were positioned randomly at the beginning of each trial, it is very likely that fitness evaluations were skewed by more beneficial initial configurations, especially since we used far fewer trials per evolution than the original authors. Figure 8 supports this conclusion - the fitness of the best individual varies dramatically over the last 50 generations. These random variations could stunt evolution by preventing better solutions from being consistently rewarded.

The distinction between navigational and bidirectional communication may explain why our recurrent networks did not outperform the feed-forward networks, as expected. Recurrent connections allow a network to remember information gathered in previous timesteps. This capacity is likely more of a hindrance, however, when it is used for our populations simple navigation rather than complex communication.

References

[1] Kristen Allen and Peter Ballen. *Evolving Grounded Communication in a Navigation Task using NEAT*. Accessed: May 1, 2014, from `http://web.cs.swarthmore.edu/~meeden/cs81/s12/papers/KristenPeterPaper.pdf`. (2012).

[2] Angelo Cangelosi and Domenico Parisi. "The Emergence of a Language in an Evolving Population of Neural Networks". In: *Connection Science* 10.2 (1998), pp. 83–97.

[3] Davide Marocco and Stefano Nolfi. "Emergence of Communication in Embodied Agents Evolved for the Ability to Solve a Colletive Navigation Problem". In: *Connection Science* 19.1 (2007), pp. 53–74.

[4] Davide Marocco and Stefano Nolfi. "The Emergence of Communication in Evolutionary Robots". In: *Philosophical Transactions of the Royal Society London* A.361 (2003), pp. 2397–2421.

[5] Kenneth Stanley. "Competitive coevolution through evolutionary complexification". In: *Journal of Artificial Intelligence Research* 21 (2004).