UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

NURTURING AS SAFE EXPLORATION PROMOTES THE EVOLUTION OF

GENERALIZED SUPERVISED LEARNING

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

BRYAN HOKE
Norman, Oklahoma
2017

NURTURING AS SAFE EXPLORATION PROMOTES THE EVOLUTION OF
GENERALIZED SUPERVISED LEARNING

A THESIS APPROVED FOR THE
SCHOOL OF COMPUTER SCIENCE

BY

_____
Dr. Dean Hougen, Chair

_____
Dr. Amy McGovern

_____
Dr. Ingo Schlupp

Contents

Abstract

The ability to learn is often a desirable property of intelligent systems which can make them more adaptive. However, it is difficult to develop sophisticated learning algorithms that are effective. One approach to the development of learning algorithms is to evolve them using evolutionary algorithms. The evolution of learning is interesting as a practical matter because harnessing it may allow us to develop better artificial intelligence; it is interesting also from a more theoretical perspective of understanding how the sophisticated learning seen in nature—including that of humans—could have arisen. A potential obstacle to the evolution of learning when alternative behavioral strategies (e.g., instincts) can evolve is that learning individuals tend to exhibit ineffective behavior before effective behavior is learned. *Nurturing*, defined as one individual investing in the development of another individual with which it has an ongoing relationship, is often seen in nature in species that exhibit sophisticated learning behavior. It is hypothesized that nurturing may be able to increase the competitiveness of learning in an evolutionary environment by ameliorating the consequences of incorrect initial behavior. The approach taken is to expand upon a foundational work in the evolution of learning to enable also the evolution of instincts and then examining the strategies evolved with and without a nurturing condition in which individuals are not penalized for mistakes made during a learning period. It is found that nurturing promotes the evolution of learning in these environments.

Chapter 1

Introduction

Chalmers (1990) demonstrated that evolutionary processes can produce artificial systems that learn. The learning that emerged could be either specific to certain tasks or generalized for any task in the problem space. It was also possible that learning might not emerge if none of the behavior adjustment rules evolved performed better than adjusting behavior randomly.

In Chalmers's work, learning was the only viable phenotypic strategy that could emerge from evolution. However, if other strategies, such as "instinctual" behavioral responses (i.e., responses that are innate), were able to emerge alongside the strategy of learning, it might be found that learning is not always the strategy that emerges from evolution. Presumably, the evolutionary environment (or objective function) would have the effect of influencing which strategies emerge; after all, it is reasonable to think that no one strategy would always be optimal for every possible kind of environment.

A disadvantage of learning is that individuals tend to have low fitness early in their life before they have had the opportunity to learn correct behavior. This is in contrast to individuals that exhibit instinctive behavior and have high fitness at the start of life. Such effects could make it less likely for learning to evolve in favor of instinctual behavior.

However, if the penalties incurred during the early stages of learning could be mitigated then learning may be more likely to evolve. In the biological world this can be observed in the form of nurturing, often from a parent to a child. In this

way, nurturing may increase the competitive viability of learning by protecting the learner from otherwise costly errors made before proper behavior has been acquired (Woehrer et al., 2012).

This work examines whether learning is less likely to evolve when both instincts and learning have the potential to emerge, and how the introduction into the environment of a "nurturing" condition—where mistakes made during learning are not penalized—influences the evolution of learning as opposed to an instinctual strategy where behavior is fixed during an individual's lifetime.

To explore these questions, this work uses a combination of genetic algorithms and artificial neurons applied to a set of supervised learning tasks (the same tasks, in fact, that were used by Chalmers). The rest of this chapter introduces the following topics, in turn: genetic algorithms, artificial neurons, artificial neural networks, neuroevolution, the evolution of learning, and nurturing.

1.1 Genetic Algorithms

Genetic algorithms were developed by Holland (1975) to explore using artificial evolution as an optimization process. Many others have built on and applied Holland's ideas for genetic algorithms over the intervening decades, and several other forms of evolutionary computation have since been developed (Engelbrecht, 2007). What follows in this section is a brief explanation of one typical version of a genetic algorithm that focuses on the mechanisms used in this thesis, but it should be noted that many other variations exist.

Genetic algorithms are a class of search algorithms inspired by the process of evolution by natural selection. Solutions to a problem are encoded in strings referred to as *chromosomes*. The distinct variables in the solution representation are encoded in substrings of the chromosome referred to as *genes*.

The procedure begins with a population of chromosomes which are typically randomly generated. The chromosomes are evaluated by an objective function—the problem to be solved—referred to as the *fitness function* to determine the fitness value of each one. The goal of the genetic algorithm is to find the chromosome which maximizes the fitness function and has the highest fitness; in this sense, a genetic algorithm can be used as a search algorithm that looks for the best solution for an objective function.

A process of selection is used to determine the chromosomes which will reproduce into the next generation based on their fitness values, where individuals with higher fitness are more likely to be selected. In *roulette-wheel selection*, individuals are selected with a probability that is linearly proportional to their fitness. In *elitist selection*, the individual with the highest fitness is automatically selected.

The selection chromosomes then reproduce to create the next-generation population. One operation used in reproduction is *crossover*, where segments of two chromosomes are combined to create a new chromosome. The offspring chromosomes may also undergo a process of *mutation*, where each gene has its value changed with some probability, in order to explore the problem space.

This process is repeated until some stopping condition is met, typically after a certain number of generations.

Figure 1.1: A high-level view of a genetic algorithm.

1.2 Artificial Neurons

The concept of an artificial neuron was introduced first by McCulloch and Pitts (1943) as a theoretical model of biological nervous activity. Many others have expanded upon this neural model and have developed a number of different neural models (Engelbrecht, 2007). What follows in this section is a brief explanation of one typical version of an artificial neuron that focuses on the mechanisms used in this thesis, but it should be noted that many other variations exist.

Artificial neurons are computational elements inspired by the biological neurons found in nature. An artificial neuron is comprised of a number of weighted input connections, an activation function which implements a response to input, and a number of output connections.
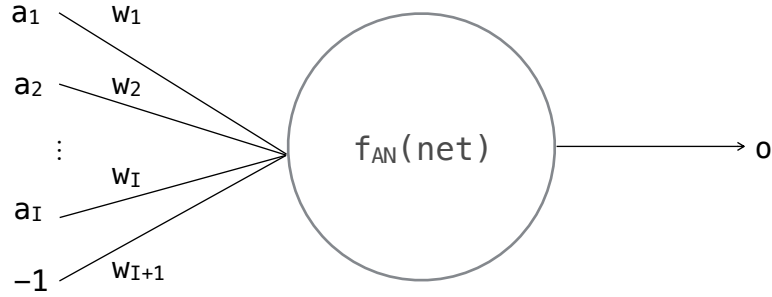
Figure 1.2: An artificial neuron.

Figure 1.2 depicts an artificial neuron with I inputs, one output, and an activation function f_{AN}. The neuron receives a vector of $I + 1$ input signals, $\mathbf{a} = (a_1, a_2, \ldots, a_I, a_{I+1})$, where a_{I+1} is known as a *bias unit* and always has a value of -1. The input vector is modulated by a weight vector, $\mathbf{w} = (w_1, w_2, \ldots, w_I, w_{I+1})$, where each weight w_i modulates the input signal a_i. The neuron computes the net

input as the sum of the weighted input signals, giving

$$net = \sum_{i=1}^{I+1} a_i w_i.$$

The output signal, o, is then computed by applying the activation function to the weighted input, so that $o = f_{AN}(net)$.

A common choice for f_{AN} is the sigmoid function:

$$f_{AN}(net) = \frac{1}{1 + e^{-\lambda net}}$$

where the parameter λ influences the steepness of the function, but usually $\lambda = 1$.

Artificial neurons can be used to compute linearly separable functions without error. This means that, for such a function, there exists at least one threshold value such that the neuron can separate the space of I-dimensional input vectors which produce an above-threshold output from those which produce a below-threshold output by an I-dimensional hyperplane. This threshold is determined by the bias weight w_{I+1}, meaning that the neuron can be used to separate the input vectors for which $net > 0$ from the input vectors for which $net \leq 0$.

Learning is a technique by which the weights of an artificial neuron are updated to realize functions given by data. One type of learning is known as *supervised learning*, where the neuron is provided with a data set, known as the *training set*, consisting of training patterns of input vectors with associated target outputs. The weights of the neuron are then adjusted until the error between the actual outputs of the neuron and the target outputs in the patterns is minimized.

An example of a supervised learning rule is the so-called *delta rule* (Widrow and Hoff, 1960). The delta rule requires the definition of an error function, $\mathcal{E}$, to measure the neuron's error in approximating training targets. The sum of squared errors is usually used, given by

$$\mathcal{E} = \sum_{p=1}^{P_T} (t_p - o_p)^2$$

where t_p and o_p are the target and actual output for the pth pattern, and P_T is the total number of patterns in the training set.

Given a single training pattern, weights are updated using

$$w_i(t) = w_i(t-1) + \Delta w_i(t)$$

with

$$\Delta w_i(t) = \eta(-\frac{\partial \mathcal{E}}{\partial w_i})$$

where

$$\frac{\partial \mathcal{E}}{\partial w_i} = -2(t_p - o_p)\frac{\partial f_{AN}}{\partial net_p}a_{i,p}$$

and η is a constant known as the learning rate, net_p is the net input for pattern p, and $a_{i,p}$ is the ith input signal in pattern p.

The delta rule requires that its activation function f_{AN} is continuous and differentiable; this is so the gradient of the error can be followed toward lower values, giving the direction in which the weights should be updated. If we use the sigmoid function for f_{AN} then

$$\frac{\partial f_{AN}}{\partial net_p} = o_p(1 - o_p)$$

giving

$$\frac{\partial \mathcal{E}}{\partial w_i} = -2(t_p - o_p)o_p(1 - o_p)a_{i,p}$$

Artificial neurons are generally connected together into artificial neural networks to perform more sophisticated computation; however, this work is based closely on the work of Chalmers (1990) which considered single neurons in isolation.

1.3 Artificial Neural Networks

Artificial neural networks were first described by McCulloch and Pitts (1943) in conjunction with their artificial neuron model, and the field has been developed by many others in the intervening decades (Engelbrecht, 2007). Rosenblatt (1958) developed the perceptron, an algorithm for pattern recognition and one of the first artificial neural networks to be produced. Werbos (1975) and Rumelhart et al. (1986) developed the backpropagation algorithm for training multi-layer networks.

Today, deep learning network architectures are used for sophisticated pattern recognition and machine learning tasks (Schmidhuber, 2015). The *Neocognitron* (Fukushima, 1980) was both the first deep neural network and the first convolutional neural network. LeCun et al. (1998) later developed *LeNet-5*, a convolutional network that became widely used for recognizing hand-written numbers. The concept of deep learning was further popularized by Hinton et al. (2006), who referred to "learning" for "deep belief nets."

1.4 Neuroevolution

Any feature of an artificial neural network, such as topology, connection weights, activation functions, learning rules, and input features, can be subjected to an evolutionary process after being string-encoded (Yao, 1999). However, this work will focus on the evolution of connection weights and learning rules.

The connection weights of a non-learning neuron can be evolved as the fixed response of that neuron to input. This is analogous to the evolution of instincts or what we think of as "hard-wired" behavior.

Alternately, the initial connection weights of a learning neuron can be evolved as the initial behavior which will be modified by the neuron's learning rule. This can be

beneficial for evolving passably fit initial behavior that is then fine-tuned by learning but it can also be the case that the values of the initial weights are not adaptive and weights are evolved that are effectively random.

Learning rules can be evolved stand-alone by applying them to neurons initialized with random weights in the fitness function. In this situation the values of the initial weights do not matter because the idea is to evolve learning rules that will converge on a solution that is independent of the values of the starting weights.

The initial connection weights and the learning rules that modify them can be evolved simultaneously. In this situation, instincts, learning, or a combination of both can be evolved depending on the environment. This work will examine the outcomes of the simultaneous evolution of connection weights and learning rules under different conditions, whereas Chalmers examined only the possibility of evolving learning rules.

1.5   The Evolution of Learning

Given the wide variety of artificial neural networks developed over the decades, the necessity of having appropriate learning rules for each, and the difficulty in determining such rules, the evolution of learning in artificial neural networks is a topic of great interest to the research community, and while many researchers have investigated many different approaches to the evolution of learning, many open questions still remain (Soltoggio et al., 2017).

Evolution and learning are both forms of adaptation. Evolution takes place across generations whereas learning takes places within an individual's lifetime. Evolution can produce adaptive behaviors such as learning or non-adaptive behaviors such as instincts.

If an environment is sufficiently diverse or dynamic, instinctual behavior is not reasonably adequate for good task performance; this is because such an environment will

invariably present novel tasks to which the instinctual behavior is not well-adapted. Evolved learning behavior that is scalable and adaptive to unknown or changing situations can be seen in countless species in nature.

Bengio et al. (1990) first proposed the optimization of a learning rule via evolutionary search. Chalmers (1990) evolved learning rules by encoding the parameters of a template formula as strings, demonstrating that the delta rule for supervised learning can be evolved in certain situations for artificial neurons in static environments. Fontanari and Meir (1991) expanded upon Chalmers's approach to evolve a learning algorithm for single-layer networks with binary weights.

McQuesten and Miikkulainen (1997) evolved reinforcement learning by culling subpar individuals and having parent networks teach their offspring using backpropagation. Niv et al. (2002) demonstrated the evolution of reinforcement learning in uncertain environments for networks with a single binary output. Shah (2015) similarly demonstrated the evolution of reinforcement learning in changing environments for networks with large numbers of rational-valued outputs.

1.6   Nurturing

Nurturing is defined as "the contribution of time, energy, or other resources by one individual to the expected physical, mental, social, or other development of another individual with which it has an ongoing relationship" (Woehrer et al., 2012).

Nurturing is prevalent in the biological world and can be an important contributing factor to the evolution of learning (Woehrer et al., 2012; Eskridge and Hougen, 2012), an idea which has recently been echoed by others (Soltoggio et al., 2017).

Nurturing itself can be either instinctive or learned. Leonce et al. (2012) demonstrated the evolution of instinctive nurturing in robots.

Nurturing can cover the initial costs of a learner and the resulting benefits of

learning can be paid forward to the next generation. In this way, nurturing can promote the evolution of learning.

The work of McQuesten and Miikkulainen (1997), where it was shown that neuroevolution could benefit from parent networks teaching their offspring using backpropagation, is the first known example demonstrating that nurturing can promote the evolution of learning; although McQuesten and Miikkulainen did not refer to the teaching they used as nurturing, it conforms to the definition followed in this thesis. Eskridge and Hougen (2012) conducted experiments involving food patch estimation in uncertain environments, and their results demonstrated that nurturing as both social learning and safe exploration can promote the evolution of learning. Shah (2015) demonstrated that nurturing as task simplification can promote the evolution of reinforcement learning in changing environments.

This work explores nurturing as safe exploration in that individuals do not incur fitness penalties for mistakes made during the learning process.

1.7 Contents of the Thesis

The rest of this thesis is organized as follows: Hypotheses about how nurturing affects the evolution of learning and instincts are detailed in **Chapter 2**. Operational definitions for terms such as "nurturing" and "instincts" are introduced in **Chapter 3**. Procedures for the experiments examining the effect of nurturing on the evolution of learning and instincts are detailed in **Chapter 4**. The experimental results, which demonstrate that nurturing facilitates the evolution of learning, are outlined in **Chapter 5**. The conclusion that nurturing facilitates the evolution of learning is described in **Chapter 6**. Finally, ideas for exploring the interaction between nurturing and the Baldwin effect in future work are described in **Chapter 7**.

Chapter 2

Hypotheses

Three factors are considered with respect to their effects on the evolution of learning: (1) the number of tasks, (2) the presence or absence of nurturing, and (3) the possibility of evolving instincts. Six hypotheses are formulated to explore these factors, and we consider them in related pairs. The first pair, which makes predictions about the addition of a non-nurturing condition to Chalmers's experimental setup, is discussed in Section 2.1. The second pair, which makes predictions about the addition of instincts to the experimental setup used by the first pair, is discussed in Section 2.2. The third pair, which makes predictions about the effects of the nurturing and non-nurturing conditions on the evolution of instincts and generalized learning, is discussed in Section 2.3. Finally, appropriate hypothesis tests for these hypotheses are discussed in Section 2.4.

2.1   Hypotheses H_1 and H_2

The first two hypotheses consider the effects of adding a non-nurturing condition to Chalmers's original model.

Chalmers demonstrated that the quality of evolved generalized learning mechanisms is proportional to the number of tasks in the evolutionary environment. When there are few tasks in the evolutionary environment, high-quality generalized learning is not expected to evolve because there is no need for evolution to produce learning that is fit for more general contexts; instead, we would expect specialized learning rules to evolve which happen to perform well in the evolutionary environment.

However, as the number of tasks in the evolutionary environment increases, the environment becomes too complex for any specialized learning to perform well in, and so we then expect generalized learning to evolve out of necessity.

Given environments like those used by Chalmers, where instincts cannot be evolved, the addition of nurturing alone is expected to have little impact on the evolution of generalized learning because learning is still the only viable strategy in that setup. Therefore, it is expected that the quality of evolved generalized learning mechanisms will remain proportional to the number of tasks and the quality of the learning evolved will be no greater with nurturing than without.

H$_1$: When only learning is evolved, the quality of generalized learning evolved will be proportional to the number of tasks in the evolutionary environment, both with and without nurturing.

In other words, we expect to be able both to re-confirm Chalmers's conclusion about the evolution of generalized learning (since the condition under which Chalmers's experiments were performed is effectively the same as the nurturing condition here) and to demonstrate that Chalmers's conclusion remains valid even when evaluated under the new non-nurturing condition.

H$_2$: When only learning is evolved, the quality of generalized learning evolved will be equal with nurturing and without.

This is to say that we do not expect the non-nurturing condition to significantly influence the quality of evolved generalized learning with respect to the quality of generalized learning evolved under the nurturing condition (i.e., the conditions Chalmers used). We expect this to be the case because while the non-nurturing condition does make learning less competitive, there still are no viable alternative strategies to learning that can be evolved.

It is important to note that H_1 and H_2 are independent hypotheses, with H_1 making a prediction about the effect of varying the number of tasks in the evolutionary

environment and H_2 making a prediction about the presence and absence of nurturing. As such, it is possible for one hypothesis to be supported while the other is not. For example, it could be the case that more complex environments promote the evolution of generalized learning but that generalized learning is also less likely to evolve without nurturing because the fitness penalties incurred during learning cause the generalizability of the learning rules evolved to be somewhat compromised by more specialized features that reach higher fitness values more quickly during training (but which perhaps have a lower post-training fitness), and so the quality of generalized learning evolved is proportional to the number of tasks in the evolutionary environment but higher-quality generalized learning is evolved with nurturing than without, in which case H_1 would be supported but H_2 would not. Alternately, it could be the case that the absence of nurturing doesn't affect the quality of generalized learning evolved but does induce a gating effect on environmental complexity which makes it much more difficult to evolve generalized learning until some threshold complexity is reached, and so the quality of generalized learning evolved is equal with nurturing and without but the quality of generalized learning evolved is not linearly proportional to the number of tasks in the evolutionary environment without nurturing, in which case H_2 would be supported but H_1 would not.

2.2 Hypotheses H_3 and H_4

The introduction of the ability to evolve initial network weights alongside learning rules allows for the evolution of "instincts" as an alternate strategy to learning. The remaining hypotheses consider the effects of adding the ability to evolve initial network weights to the model developed in the first two hypotheses.

It is expected that higher-quality instincts will evolve for small numbers of evolutionary tasks, but higher-quality learning will evolve as the number of evolutionary

14

tasks increases. This will be because there will be a smaller number of instinct genes than learning genes to "tune" to a correct solution for small number of tasks, but because the number of genes required to represent instincts will increase as the number of tasks increases the number of instinct genes will become much larger than the number of learning genes as the number of tasks increases.

H$_3$: When both instincts and learning are evolved, the quality of instincts evolved will be inversely proportional to the number of tasks in the evolutionary environment, both with and without nurturing.

In other words, higher-quality instincts will evolve with low numbers of tasks in the evolutionary environment, and lower-quality instincts will evolve with higher numbers of tasks in the evolutionary environment. We expect this to be the case because the number of genes required to represent instinctive responses to the environment will be proportional to the number of tasks in the environment, and while it is easy to find good solutions when only small numbers of genes are evolved, it becomes increasingly difficult to find good solutions as the number of genes involved increases.

H$_4$: When both instincts and learning are evolved, the quality of generalized learning evolved will be proportional to the number of tasks in the evolutionary environment, both with and without nurturing.

In other words, higher-quality generalized learning will evolve with low numbers of tasks in the evolutionary environment, and lower-quality generalized learning will evolve with higher numbers of tasks in the evolutionary environment. We expect this to be the case for the same reasons we expect H_1 to be supported; indeed, H_4 can be seen as a prediction that the mechanisms behind H_1 will still hold despite the addition of the possibility of evolving instincts.

H_3 and H_4 can be seen as sort of counterparts to each other because they predict that the quality of evolved instincts and generalized learning rules will exhibit opposite trends with respect to the number of tasks in the evolutionary environment; that

instincts will be strong under the conditions where generalized learning is weak, and vice-versa. However, it is important to remain clear about the fact that the outcomes of H_3 and H_4 have no dependencies on each other because they make predictions about instincts and generalized learning with respect to the number of tasks in the evolutionary environment, not each other; it just so happens that we expect instincts and generalized learning to exhibit opposite trends but we do not necessarily expect a zero-sum relationship between the two. For example, it could be the case that learning and instincts are mutually beneficial, and so the quality of both evolved instincts and generalized learning would be proportional to the number of tasks in the evolutionary environment, in which case H_4 would be supported but H_3 would not. Alternately, it could be the case that the larger number of genes required for more tasks and the resulting increase in complexity of the search space means that the quality of both evolved instincts and generalized learning are inversely proportional to the number of tasks in the evolutionary environment, in which case H_3 would be supported but H_4 would not. There are, of course, other possible relationships between the factors involved such that one or the other or both of these hypotheses are false.

2.3   Hypotheses H_5 and H_6

The final two hypotheses consider the relationship between the nurturing condition (present or absent) and the quality of evolved instincts or generalized learning. Neither of these hypotheses makes a prediction about the relationship between evolved instincts and generalized learning and the number of tasks in the evolutionary environment.

It is expected that higher-quality generalized learning will evolve in the nurturing condition than in the non-nurturing condition because of the costs associated with the learning process in the non-nurturing condition; individuals tend to have low fit-

ness early in their life before they have had the opportunity to learn correct behavior, but in the nurturing condition individuals do not incur fitness penalties for mistakes made early in life. Likewise, it is expected that higher-quality instincts will evolve in the non-nurturing condition than in the nurturing condition because, without nurturing, the learning individuals will suffer fitness consequences from their unfit initial behavior, increasing the competitive value of instinctual behavior in the evolutionary environment.

H_5: When both instincts and learning are evolved, higher-quality instincts will evolve without nurturing than with.

H_6: When both instincts and learning are evolved, higher-quality generalized learning will evolve with nurturing than without.

Like H_3 and H_4, H_5 and H_6 can be seen as counterparts to each other because they predict that the quality of evolved instincts and generalized learning rules will exhibit opposite effects with respect to the presence or absence of nurturing. However, as with H_3 and H_4, it is important to remain clear about the fact that H_5 and H_6 have no dependencies on each other because they make predictions about instincts and generalized learning with respect to the presence or absence of nurturing, not each other. For example, it could be the case that under the nurturing condition the increased competitiveness of learning does not inhibit the evolution of instincts, and so higher-quality generalized learning would evolve with nurturing than without but the instincts evolved with nurturing would be equal to those evolved without, in which case H_6 would be supported but H_5 would not. Alternately, it could be the case that the penalties incurred during learning under the non-nurturing condition promote the evolution of faster, more effective generalized learning than that which tends to evolve under the nurturing condition, and so higher-quality generalized learning evolves without nurturing than with, which in turn promotes the evolution of better instincts in an example of what is known as the *Baldwin effect* (Floreano

and Mattiussi, 2008), in which case H_5 would be supported but H_6 would not.

2.4   Hypothesis Testing

Hypotheses H_1, H_3, and H_4 make predictions about proportionality, and so they will be tested using the Pearson correlation coefficient, which is a measure of the linear correlation between two variables. Hypotheses H_2, H_5, and H_6 make predictions about performance under one condition being better than performance under another condition, and so they will be tested using a randomized ANOVA procedure for comparing performance curves (Piater et al., 1998).

Chapter 3

Operational Definitions

An *individual* is a candidate solution to the genetic algorithm's fitness function whose characteristics are represented by a chromosome (Engelbrecht, 2007). In this work, individuals are manifest as artificial neurons.

A *task* is a set of input-output patterns, where for each pattern in the task an individual is evaluated on its ability to produce the pattern output when activated with the corresponding pattern inputs.

An *environment* is a set of tasks by which individuals are assessed.

An individual's *lifetime* is the period during which it is manifest as an artificial neuron and assessed by the fitness function. Specifically, an individual's lifetime consists of a supervised learning period followed by a final evaluation.

Instincts are to be represented by the genetically encoded initial weights of a neuron. The tasks are independent, so there will need to be a separate set of instinct weights for each task.

Learning is to be represented by the genetically encoded learning rule that is applied to a neuron throughout its lifetime.

For evolved neurons where both initial weights and learning rules are genetically encoded, instinct quality is measured by removing the learning rule and evaluating the fitness of the neuron in the environment in which it was evolved. For the same neurons, *learning ability* is measured by replacing the genetically encoded initial weights with random initial weights and evaluating the fitness of the neuron in the environment in which it was evolved; *generalized learning ability* is measured by evaluating the

19

fitness of this instinct-removed neuron in an environment different from the one in which it was evolved.

By default, an individual's lifetime fitness is the average of all fitness evaluations made during learning—when the individual is prone to making mistakes—and the final fitness evaluation that follows learning; this is the "non-nurturing" condition. In the "nurturing" condition, then, an individual's lifetime fitness is simply the result of the final fitness evaluation taken after the individual has had an opportunity to learn correct behavior.

Chapter 4

Procedure

4.1   Genetic Coding of Learning Mechanisms

The weight update rule for an artificial neuron needs to be based on the components of that neuron which are, as covered in Section 1.2, a_j the activation of the input unit j, o_i the activation of the output unit i, t_i the training signal on output unit i, and w_{ij} the current value of the connection strength from input j to output i.

Following Chalmers (1990), the genome encodes a function F such that $\Delta w_{ij} = F(a_j, o_i, t_i, w_{ij})$, where F is a linear function of its four parameters and their six pairwise products. Thus F is determined by specifying ten coefficients.

The genome encodes these ten coefficients, as well as an eleventh "scale" parameter. That is,

$$\Delta w_{ij} = k_0(k_1 w_{ij} + k_2 a_j + k_3 o_i + k_4 t_i + k_5 w_{ij} a_j + k_6 w_{ij} o_i + k_7 w_{ij} t_i + k_8 a_j o_i + k_9 a_j t_i + k_{10} o_i t_i),$$

where k_0 to k_{10} are the encoded coefficients.

The portion of the genome which encodes Δw_{ij} consists of 35 bits. The first five bits encode the scale parameter k_0 such that it can represent the values 0, $\pm 1/256$, $\pm 1/128$, ..., ± 32, ± 64, via exponential encoding. The first bit encodes the sign of k_0 (0: negative, 1: positive), and the next four bits encode the magnitude. If these four bits are interpreted as an integer j between 0 and 15, we have

$$|k_0| = \begin{cases} 0 & \text{if } j = 0 \\ 2^{j-9} & \text{if } j = 1, ..., 15. \end{cases}$$

The other 30 bits encode the other ten coefficients in groups of three. The first bit of each group expresses the sign, and the other two bits express a magnitude of 0, 1, 2, or 4 via a similar exponential encoding. If we interpret these two bits as an integer j between 0 and 3, then

$$|k_i| = \begin{cases} 0 & \text{if } j = 0 \\ 2^{j-1} & \text{if } j = 1, 2, 3. \end{cases}$$

4.2   Genetic Coding of Initial Weights

Initial weights may be evolved alongside learning rules. It would not be meaningful for each chromosome to encode a single set of weights to be used on all tasks, so instead each chromosome simultaneously encodes a distinct set of weights for each task in the evolutionary environment. The weights to be applied to each evolutionary task are encoded in distinct, consistent regions of each chromosome.

Initial weights are encoded using 3 bits each. The first bit is the sign of the weight, and the other two bits express a magnitude of 0, $\frac{1}{2}$, 1, or 2 via an exponential encoding. If we interpret the remaining two bits as an integer j between 0 and 3, then

$$|k_i| = \begin{cases} 0 & \text{if } j = 0 \\ 2^{j-2} & \text{if } j = 1, 2, 3. \end{cases}$$

For each task the number of weights encoded is equal to the number of inputs in that task plus one bias weight.

4.3   Evaluation of Fitness

There are two conditions of evaluation: the "nurturing" case and the "non-nurturing" case. During evaluation, each network is first trained for 10 epochs using its learning rule. Following this, the network is evaluated on the same tasks. In the "nurturing" case the individual's fitness is simply the result of the evaluation, whereas in the "non-nurturing" case the individual's fitness is the average of evaluation during all training epochs and the final evaluation.

Fitness evaluation for the nurturing case:

(1) Create a network with the appropriate number of input units for the task and a single output unit.

(2) Initialize the connection strengths of the network using the values encoded in the chromosome.

(3) For 10 epochs, cycle through the training exemplars for the task, where for each exemplar:

(3a) propagate input values through the system, yielding output values; then

(3b) adjust the weights of the system according to the formula specified by the weight update rule, on the basis of inputs, output, training signal, and current weights.

(4) At the end of this process, fitness on the task is measured by testing the network on all training exemplars, and dividing the total error by the number of exemplars, and subtracting from 1. This yields a fitness value between 0 and 1.

Fitness evaluation for the non-nurturing case:

(1) Create a network with the appropriate number of input units for the task and a single output unit.

(2) Initialize the connection strengths of the network using the values encoded in the chromosome.

(3) For 10 epochs, cycle through the training exemplars for the task, where for

each exemplar:

(3a) test the network on the exemplar and measure the error in the network output; then

(3b) propagate input values through the system, yielding output values, then adjust the weights of the system according to the formula specified by the weight update rule, on the basis of inputs, output, training signal, and current weights.

(4) At the end of this process, test the network on all training exemplars and divide the total error of this test and all tests from step 3a by the total number of tests that occurred, and subtracting from 1. This yields a fitness value between 0 and 1.

Fitness of the chromosome is obtained by evaluating its performance on each of the (typically 20) tasks, and taking the mean fitness over all tasks. In this way every chromosome is assigned a fitness between 0 and 1.

4.4 Parameters of the Genetic Algorithm

A fixed population size of 40 is used in all experiments. Following the fitness evaluation process described in Section 4.3, a new population of 40 individuals is created as follows:

(1) Elitist selection is applied, meaning that an exact copy of the individual with the highest fitness in the previous population is inserted into the new population.

(2) Roulette-wheel selection (where an individual's probability of being selected is linearly proportional to its fitness) is used to select 39 individuals from the previous population with replacement (so an individual can be selected more than once), then:

(2a) the first 32 selected individuals are reproduced by gene-wise two-point crossover in pairs to create 32 new individuals which are inserted into the new population;

(2b) the remaining 7 selected individuals are reproduced by cloning to create 7

new individuals which are inserted into the new population.

(3) Each individual in the new population is mutated such that each bit in each chromosome has a 1% chance of changing.

This cycle of fitness-evaluation and reproduction is repeated for 1000 generations.

4.5   Post-Evolutionary Evaluation

Before each evolutionary run, 10 tasks are selected at random from the task pool and designated as "test tasks," and up to 20 tasks are selected at random from the remaining 20 tasks and designated as "evolutionary tasks." New evolutionary tasks and test tasks are selected for each repetition of an evolutionary run and between evolutionary runs nurturing condition and non-nurturing condition . After each evolutionary run, the chromosome with the highest fitness score in the history of the populations in that run is identified. From this chromosome, the learning rule bits are isolated and evaluated on the test tasks (using both the nurturing and non-nurturing conditions) and the bits encoding the initial weights are isolated and evaluated on the evolutionary tasks. Each evolutionary run is repeated 30 times.

Chapter 5

Results

This chapter presents the results of all experiments performed.

5.1   Results of Initial Evolutionary Runs

This section presents the results of experiments where learning rules can be evolved but initial weights cannot, performed under both the nurturing and non-nurturing conditions. In these experiments, each evolutionary run lasted for 4000 generations and only 10 repetitions were performed under each condition.

The linear correlation between the generalized learning fitness and the number of evolutionary tasks is statistically significant for both the nurturing condition (Pearson correlation, $\rho = 0.64$, $p < 0.001$) and the non-nurturing condition (Pearson correlation, $\rho = 0.62$, $p < 0.001$).

The differences between the results for the nurturing and non-nurturing conditions are statistically significant (randomized ANOVA, algorithm $p < 0.001$, interaction $p = 0.082$) Piater et al. (1998).
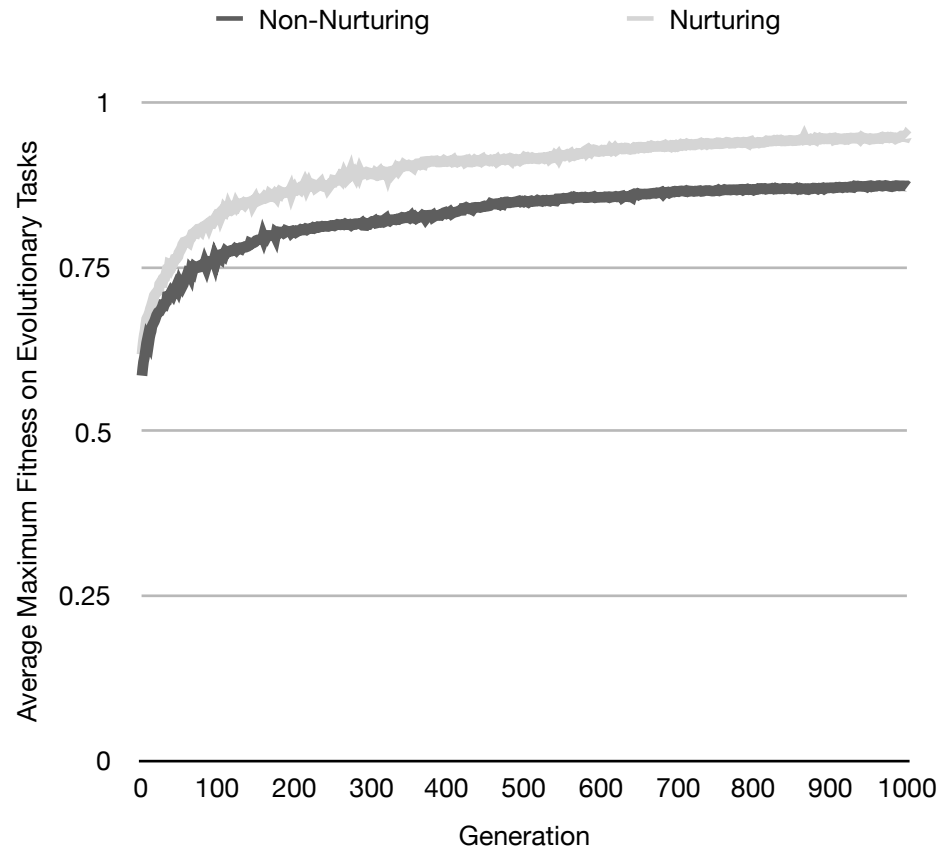
Figure 5.1: The maximum fitness for populations evolved using 20 tasks under the nurturing and non-nurturing conditions.
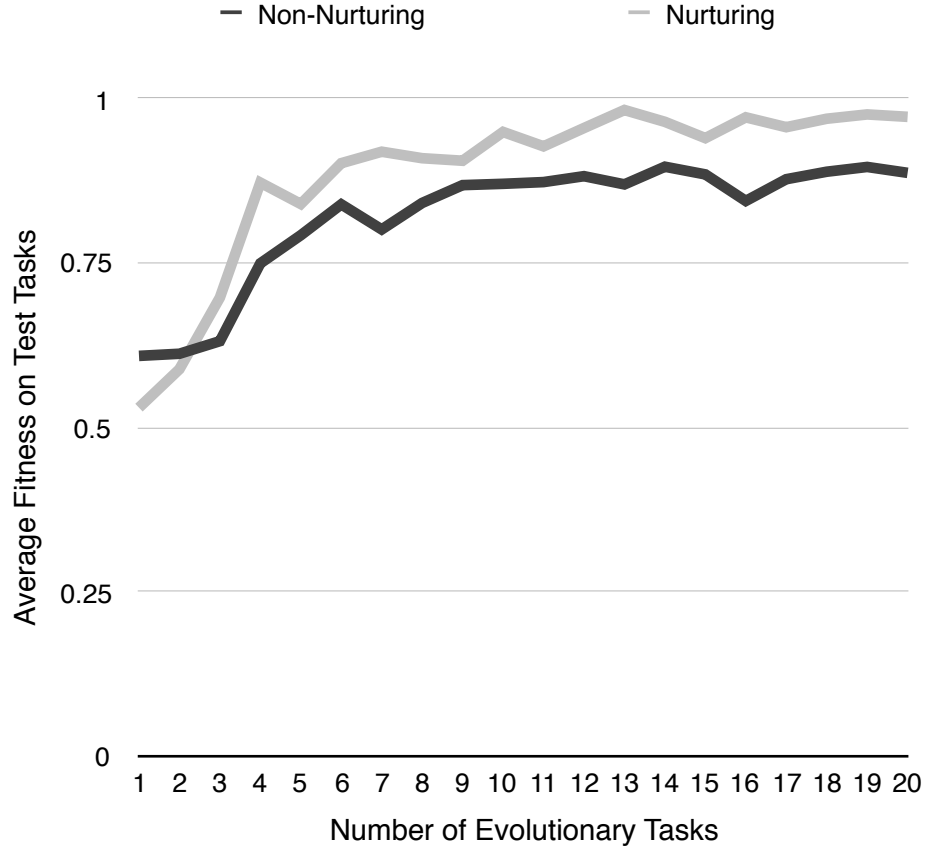
Figure 5.2: The average generalized learning capabilities for the best individuals evolved under the nurturing and non-nurturing conditions.

5.2 Intra-Learning and Post-Learning Fitness Test

This section and the following sections present the results of experiments where learning rules and initial weights can be evolved, performed under both the nurturing and non-nurturing conditions.

The individual with the highest fitness in each evolutionary run is re-evaluated on the evolutionary tasks from that run using both the nurturing and non-nurturing conditions (recall that it is evaluated using only one of these conditions during evolution).

The fitness test for the non-nurturing condition can be seen as testing how much

fitness an individual would collect on average during the learning process (and one final evaluation immediately afterward); as such, it will be referred to as the "intra-learning fitness test."

Similarly, the fitness test for the nurturing condition can be seen as testing how much fitness an individual would collect on average after learning as been completed; it will be referred to as the "post-learning fitness test."
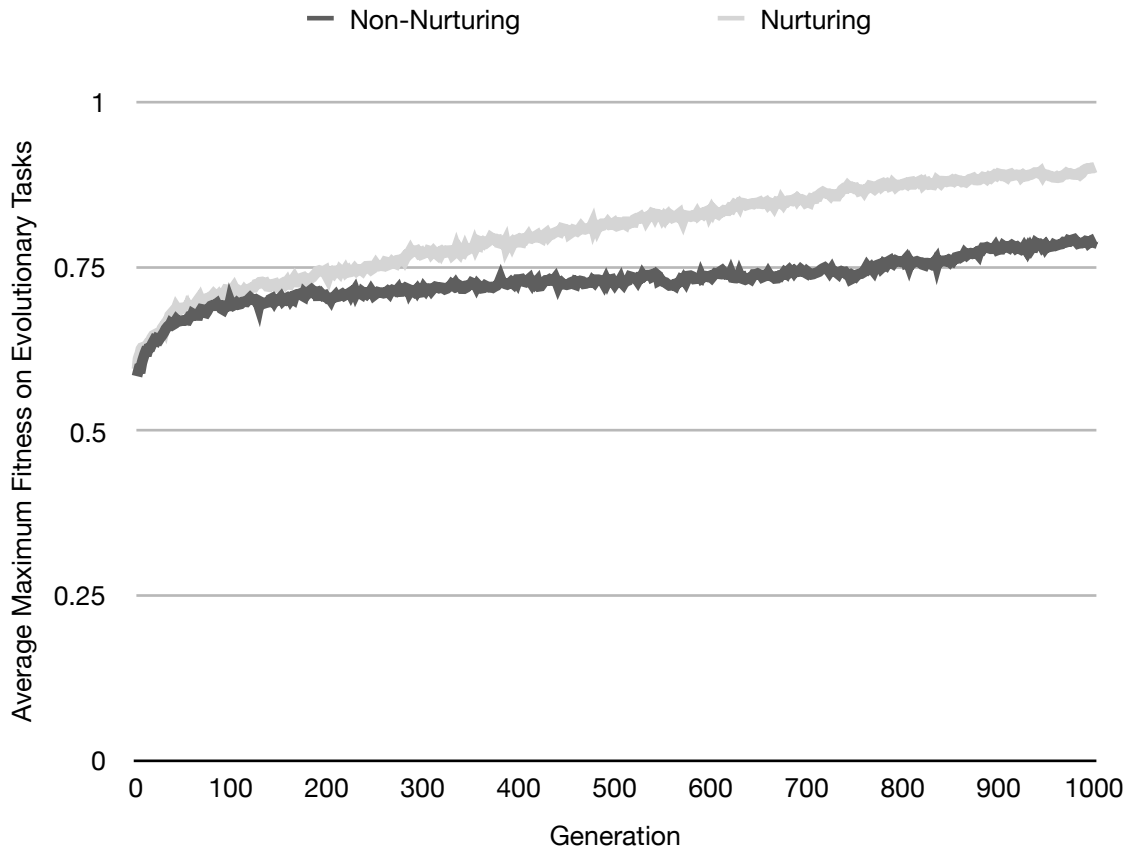

Figure 5.3: The maximum fitness for populations evolved using 20 evolutionary tasks under the nurturing and non-nurturing conditions.
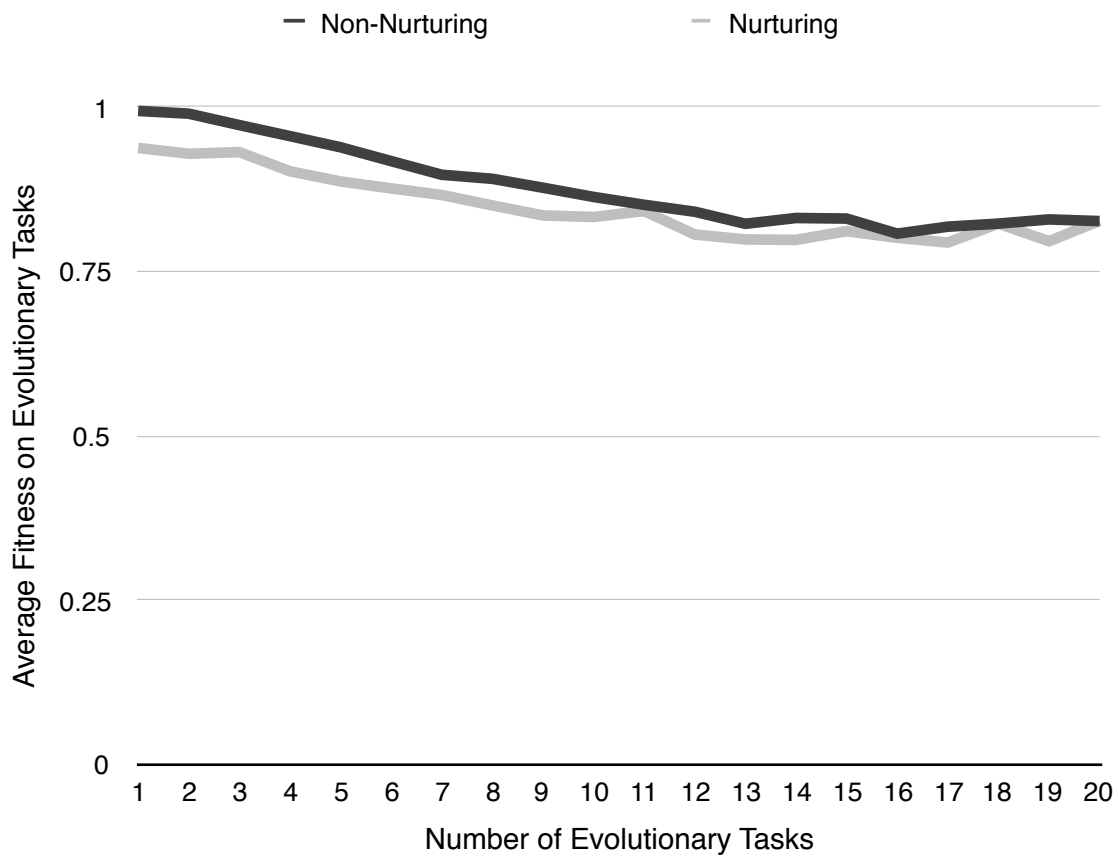
Figure 5.4: Intra-learning/non-nurturing fitness test on evolutionary tasks for individuals evolved under the nurturing and non-nurturing conditions.
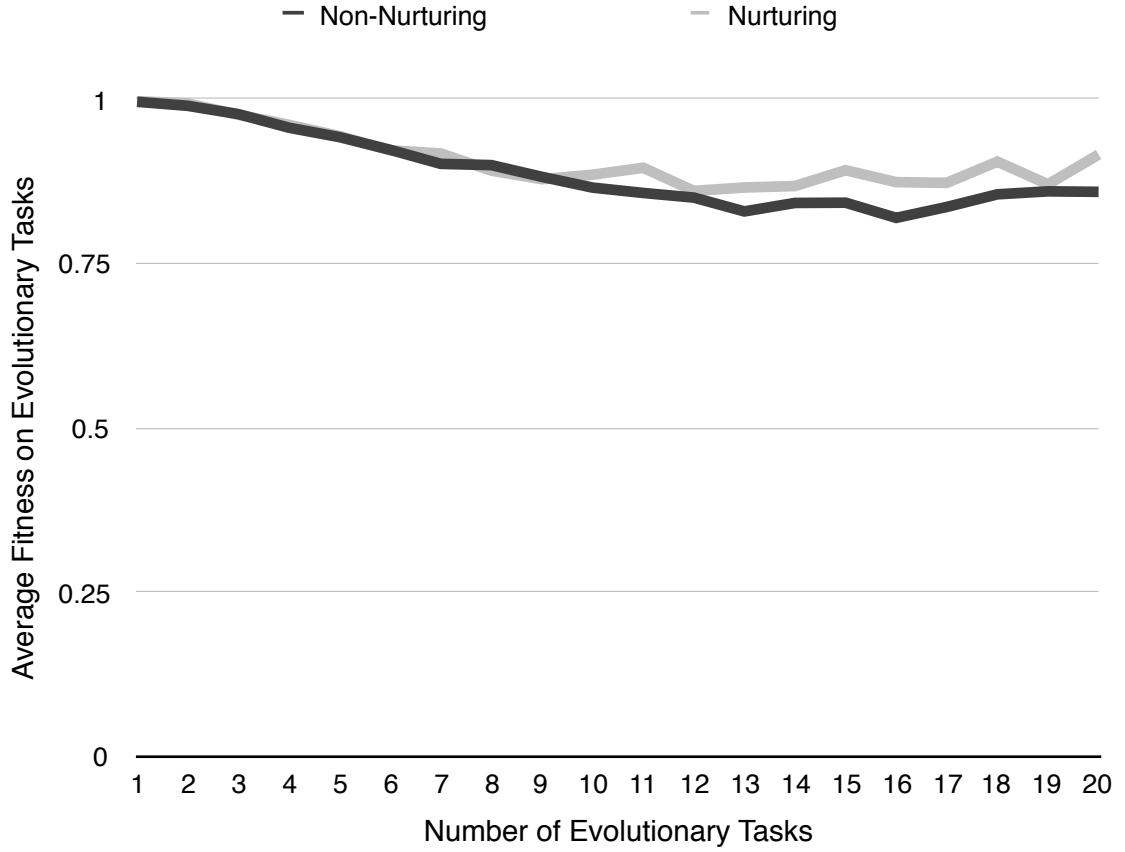
Figure 5.5: Post-learning/nurturing fitness test on evolutionary tasks for individuals evolved under the nurturing and non-nurturing conditions.

5.3 Pre-Learning Test

The portion of the chromosome that encodes the initial weights is isolated and evaluated on the evolutionary tasks by initializing a neuron with the encoded weights, with the intention of determining the contribution to the fitness of the whole chromosome of just the portion of the chromosome that encodes the initial weights. In other words, this is intended to test the instincts of the individual in the absence of learning.

The inverse-linear correlation between the pre-learning fitness and the number of evolutionary tasks is statistically significant for both the nurturing condition (Pear-

son correlation, $\rho = -0.63$, $p < 0.001$) and the non-nurturing condition (Pearson correlation, $\rho = -0.83$, $p < 0.001$).

The differences between the results for the nurturing and non-nurturing conditions are statistically significant (randomized ANOVA, algorithm $p < 0.001$, interaction $p < 0.001$).
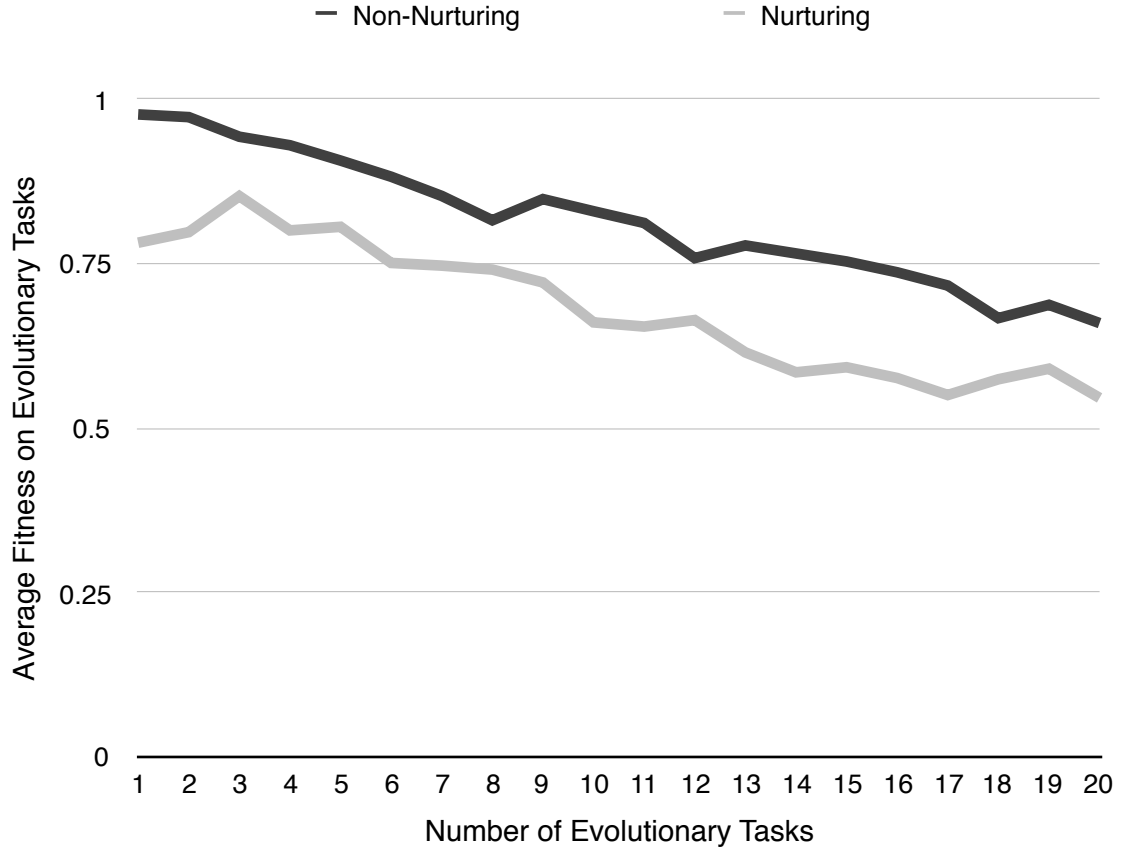

Figure 5.6: Pre-learning/instinct test on evolutionary tasks for individuals evolved under the nurturing and non-nurturing conditions.

5.4 Learning Improvement

To determine the average amount of useful learning that evolves for a given number of tasks we can examine the difference between the post-learning fitness and the pre-learning (instinctive) fitness, as shown in Figure 5.7. This is because the pre-learning

fitness is the individual's fitness before it learns, whereas the post-learning fitness is an individual's fitness after it learns.
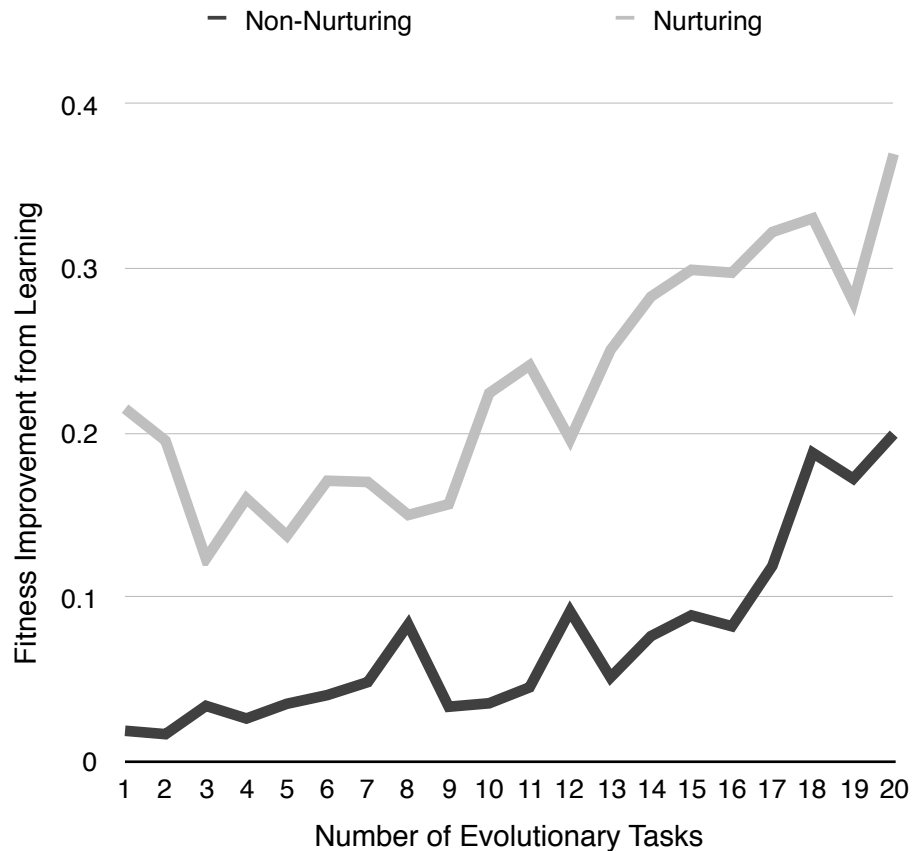

Figure 5.7: The difference between evolutionary fitness and network fitness for individuals evolved under the nurturing and non-nurturing conditions.

5.5 Weight-Generalization Learning Test

The portion of the chromosome that encodes the learning rule is isolated and evaluated on the evolutionary tasks by initializing a network with random weights and the encoded learning rule, with the intention of determining the contribution of the portion of the chromosome that encodes the learning rule to the fitness of the whole chromosome.
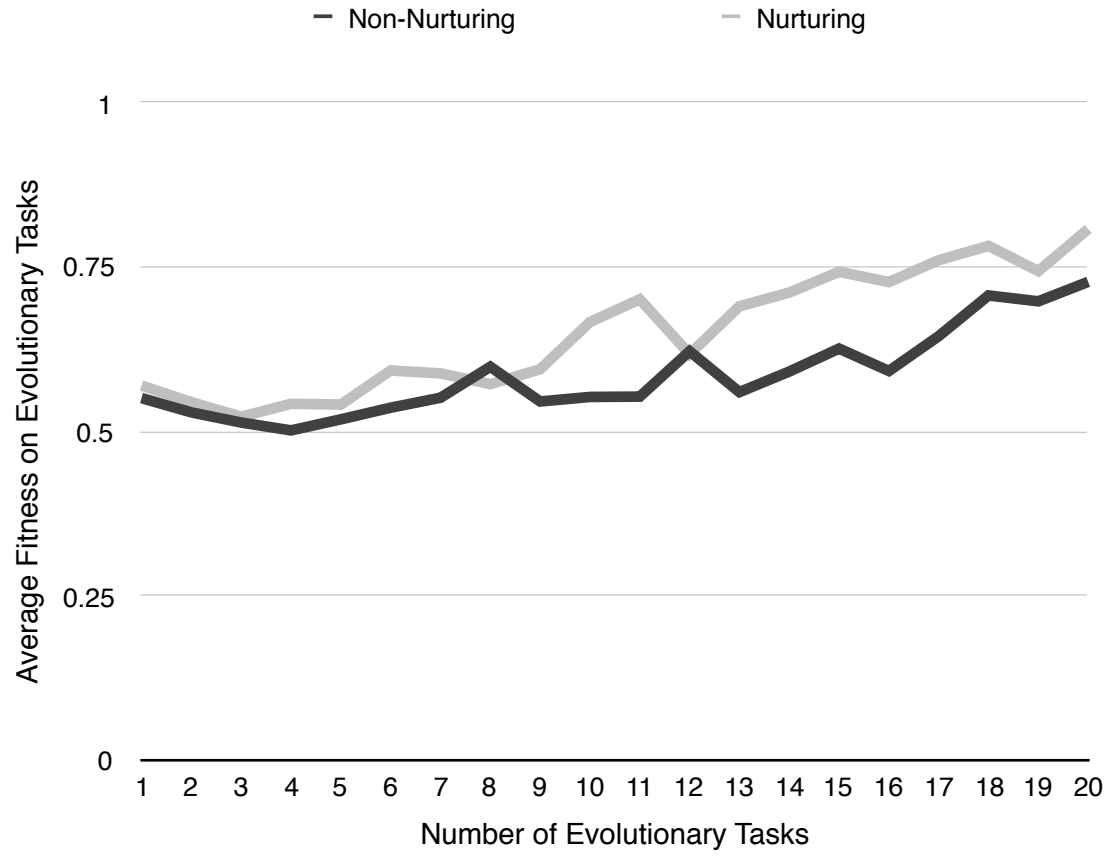
Figure 5.8: Random-start, intra-learning/non-nurturing learning test on evolutionary tasks for individuals evolved under the nurturing and non-nurturing conditions.
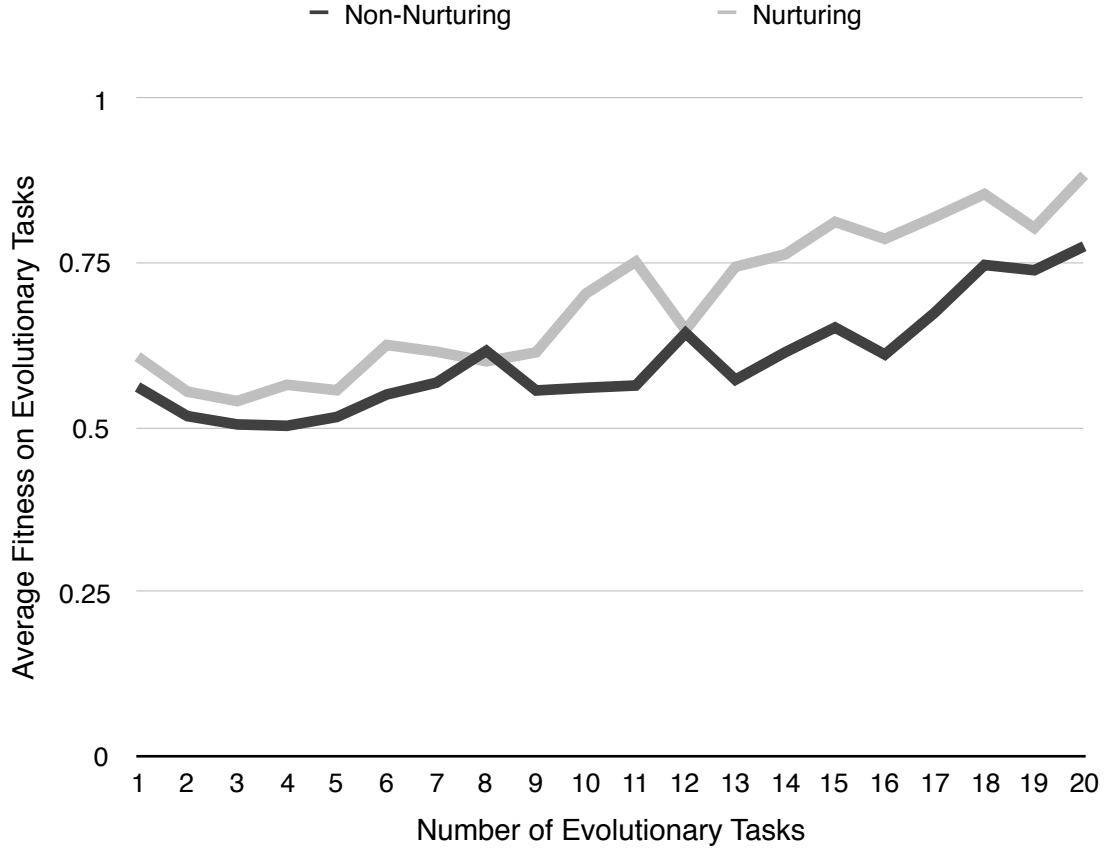
Figure 5.9: Random-start, post-learning/nurturing learning test on evolutionary tasks for individuals evolved under the nurturing and non-nurturing conditions.

5.6   Task Generalization Learning Test

The portion of the chromosome which encodes the learning rule is isolated and evaluated on the test tasks by initializing a network with random weights and the encoded learning rule, with the intention of determining the capacity for generalized learning that was evolved. The differences between the results for the nurturing and non-nurturing conditions are statistically significant (randomized ANOVA, algorithm $p < 0.001$, interaction $p < 0.001$).

The linear correlation between the generalized learning fitness and the number of evolutionary tasks is statistically significant for the intra-learning test under both the

nurturing condition (Pearson correlation, $\rho = 0.62$, $p < 0.001$) and the non-nurturing condition (Pearson correlation, $\rho = 0.43$, $p < 0.001$) and the post-learning test under both the nurturing condition (Pearson correlation, $\rho = 0.61$, $p < 0.001$) and the non-nurturing condition (Pearson correlation, $\rho = 0.42$, $p < 0.001$).
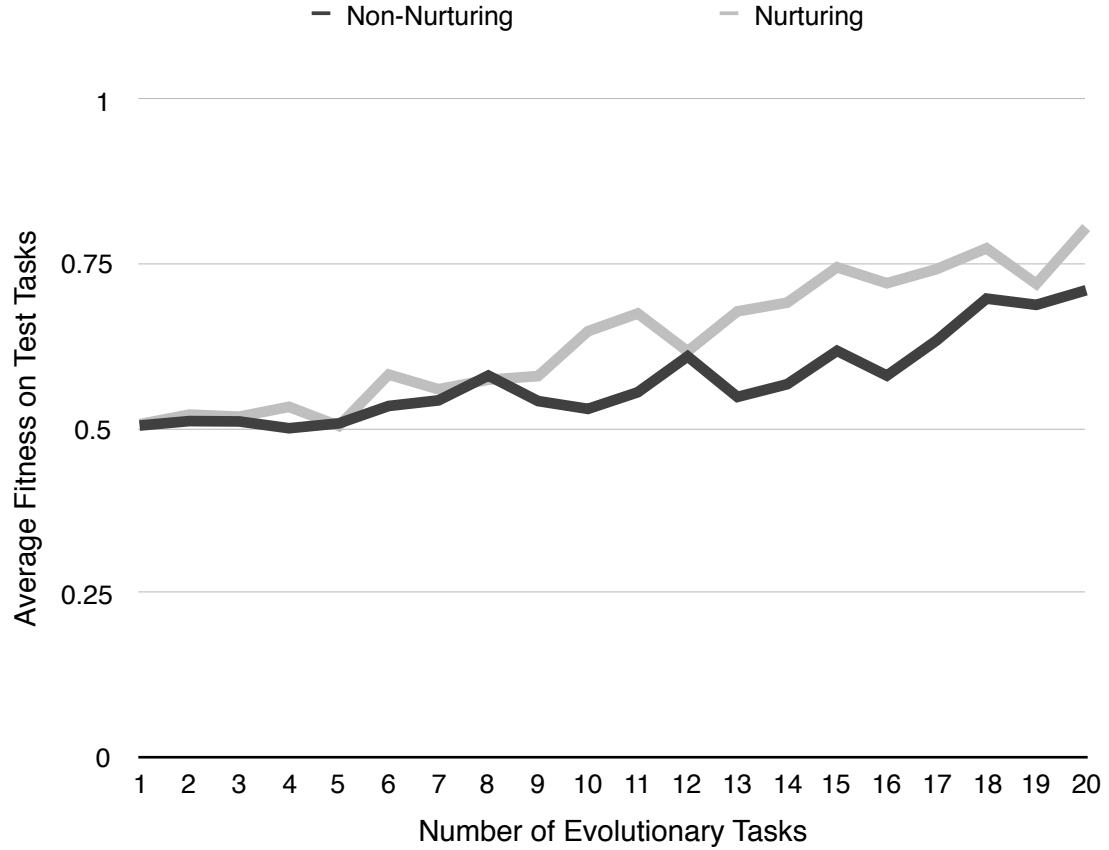

Figure 5.10: Intra-learning/non-nurturing generalization test on test tasks for individuals evolved under the nurturing and non-nurturing conditions.
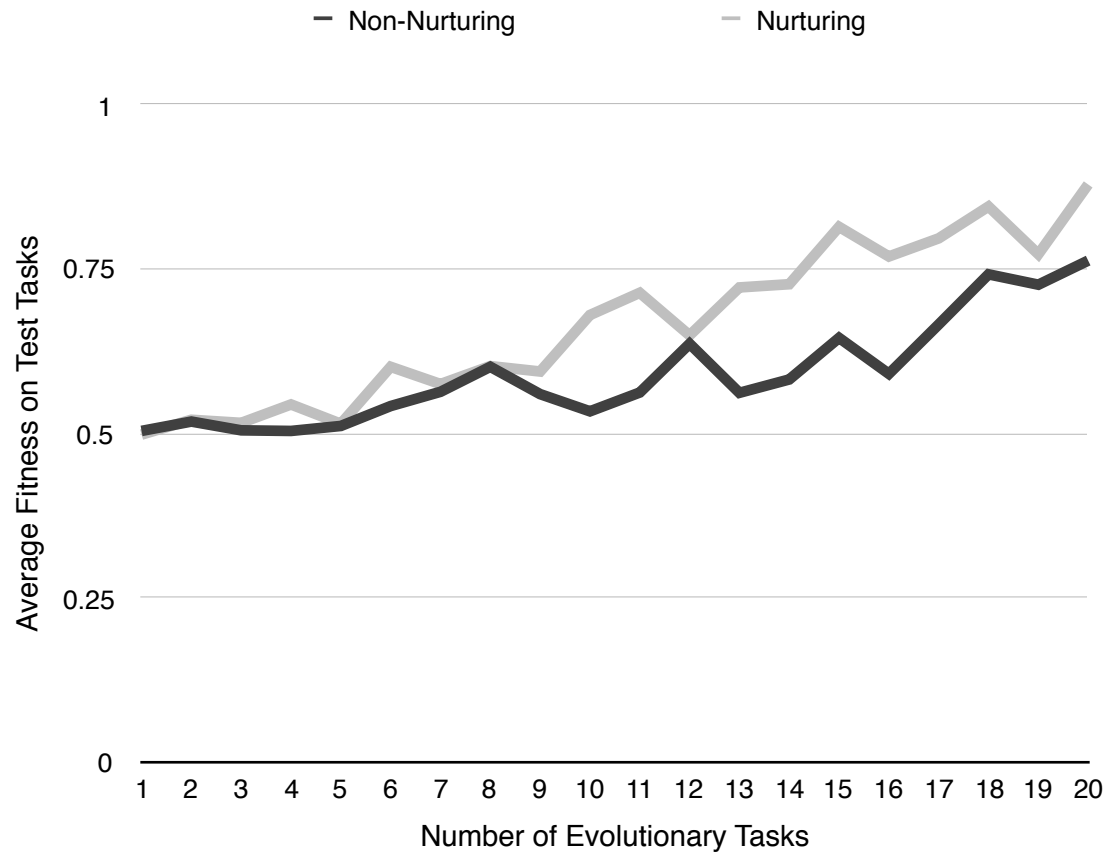
Figure 5.11: Post-learning/nurturing generalization test on test tasks for individuals evolved under the nurturing and non-nurturing conditions.

Chapter 6

Discussion

6.1 Initial Evolutionary Runs

Figure 5.1 shows the increase of maximum fitness with the progression of generations, which demonstrates the increasing adaptation of the individuals to their environment. This is the same trend seen in Figure 1 of Chalmers (1990), which depicted the evolution of maximum fitness in the populations evolved in that work. This similarity to previous findings helps to validate our setup as sufficient to observe the evolution of learning and indicates that we have sufficiently replicated Chalmers's setup to be confident in comparing our results to his.

Figure 5.2 demonstrates that the generalized learning capabilities of evolved individuals increases with the number of evolutionary tasks.

Under both evaluation conditions, the generalized learning fitness was statistically significantly proportional to the number of evolutionary tasks, suggesting that the quality of evolved generalized learning increases with the number of evolutionary tasks. This result serves as a re-confirmation of Chalmers's original findings (Chalmers, 1990).

However, in both evaluation conditions, the learning rule evolved under the nurturing condition performed better than the learning rule evolved under the non-nurturing condition by a statistically significant margin, suggesting that better generalized learning is more likely to evolve under the nurturing condition. This is contrary to our hypothesis that the presence or absence of nurturing would not affect

the quality of evolved generalized learning when learning rules could be evolved but initial weights could not. A possible explanation for this unexpected result is that the individuals evolved under the nurturing condition were evaluated using only the post-learning fitness test, while the individuals evolved under the non-nurturing condition were evaluated only under the intra-learning test. The intra-learning test is expected to yield lower fitness scores than the post-learning test, and this difference may account for the result found.

6.2   Intra-Learning and Post-Learning Fitness Test

Figure 5.3 demonstrates the increase of maximum fitness with the progression of generations, as well as the increasing adaptation of the individuals to their environment.

Under both conditions it is clear that the objective function becomes more difficult to satisfy as the number of evolutionary tasks increases.

Individuals evolved using the non-nurturing condition perform better when re-evaluated using the intra-learning fitness test/non-nurturing condition than when using the post-learning test/nurturing condition, and individuals evolved using the nurturing condition perform better when re-evaluated using the post-learning test/nurturing condition than when using the intra-learning test/non-nurturing condition.

6.3   Pre-Learning Test

Under both evaluation conditions, the pre-learning fitness was statistically significantly inversely proportional to the number of evolutionary tasks, suggesting that the quality of evolved instincts decreases with the number of evolutionary tasks when it is also possible to evolve learning.

Networks evolved using the non-nurturing condition performed better on this evaluation than networks evolved using the nurturing condition by a statistically signifi-

cant margin, suggesting that instincts (initial weights) make a larger contribution to fitness in the non-nurturing case than in the nurturing case.

6.4   Learning Improvement

As shown in Figure 5.7, the improvement in fitness due to learning increases as the number of evolutionary tasks increases. This suggests that learning becomes an increasingly important evolutionary strategy as the environment becomes more varied.

6.5   Weight-Generalization Learning Test

In both evaluation conditions, the learning rule evolved under the nurturing condition performed better than the learning rule evolved under the non-nurturing condition, suggesting that the learning rule makes a larger contribution to fitness under the nurturing condition than it does under the non-nurturing condition.

6.6   Task Generalization Learning Test

Under both evaluation conditions, the generalized learning fitness was statistically significantly proportional to the number of evolutionary tasks, suggesting that the quality of evolved generalized learning increases with the number of evolutionary tasks when it is also possible to evolve instincts.

In both evaluation conditions, the learning rule evolved under the nurturing condition performed better than the learning rule evolved under the non-nurturing condition by a statistically significant margin, suggesting that better generalized learning is expected to evolve under the nurturing condition.

Chapter 7

Conclusions

7.1 Specific Conclusions

H_1: When only learning rules are evolved, the quality of generalized learning evolved will be proportional to the number of tasks in the evolutionary environment, both with and without nurturing.

This hypothesis was supported by the results.

H_2: When only learning rules are evolved, the quality of generalized learning evolved will be equal with nurturing and without.

This hypothesis did not seem to be supported by the results. However, as per Section 6.1, this may be an issue with the data comparison methods used rather than a true contradiction of the hypothesis.

H_3: When both initial weights and learning rules are evolved, the quality of instincts evolved will be inversely proportional to the number of tasks in the evolutionary environment, both with and without nurturing.

This hypothesis was supported by the results.

H_4: When both initial weights and learning rules are evolved, the quality of generalized learning evolved will be proportional to the number of tasks in the evolutionary environment, both with and without nurturing.

This hypothesis was supported by the results.

H_5: When both initial weights and learning rules are evolved, higher-quality instincts will evolve without nurturing than with.

This hypothesis was supported by the results.

H$_6$: When both initial weights and learning rules are evolved, higher-quality generalized learning will evolve with nurturing than without.

This hypothesis was supported by the results.

7.2   Overall Conclusions

The impact of varying the nurturing condition alone was insignificant; fitnesses measured under the nurturing condition were generally higher than those measured under the non-nurturing condition, but this is to be expected because of the additional fitness penalties imposed by the non-nurturing condition.

When instincts and learning were allowed to evolve alongside each other, it was found that instinct-based solutions were more likely to evolve for small numbers of evolutionary tasks, but that learning became more likely to evolve as the number of evolutionary tasks increased.

By varying the nurturing condition when instincts were allowed to evolve alongside learning, it was found that higher-quality generalized learning was expected to evolve under the nurturing condition than under the non-nurturing condition.

7.3   Contributions

The contributions of this research are concrete evidence that nurturing promotes the evolution of learning, clarification of the role that instincts may play in the evolutionary process, and a reconfirmation of Chalmers's basic conclusions (Chalmers, 1990).

Chapter 8

Future Work

8.1 The Baldwin Effect

Baldwin (1896) argued that learning accelerates evolution because suboptimal individuals increase their baseline fitness by acquiring more adaptive characteristics during life. Lifetime learning often involves a cost because the individual may be at risk at an early stage of its life or it may modify its behavior in ways that are not functional for its survival. Baldwin suggested that evolution tends to select individuals that are born with some of the useful features that would otherwise be learned. However, a consequence of this effect is that learned features may gradually become assimilated into the genotype and reduce the role of learning; this is not necessarily a good effect if it is desirable for the individuals to retain the ability to learn over evolutionary time (Floreano and Mattiussi, 2008).

Nurturing can reduce the evolutionary costs of learning and thus may be able to inhibit the Baldwin effect. Inhibiting the Baldwin effect may result in evolved individuals that retain greater learning capabilities over evolutionarily significant periods of time.

Bibliography

Yoshua Bengio, Samy Bengio, and Jocelyn Cloutier. Learning a synaptic learning rule. Technical report, Université de Montréal, Département d'Informatique et de Recherche Opérationnelle, 1990.

David J. Chalmers. The evolution of learning: An experiment in genetic connectionism. In Morgan Kaufmann, editor, *Proceedings of the 1990 Connectionist Models Summer School*, pages 81–90. Morgan Kaufmann, 1990.

Andries P. Engelbrecht. *Computational Intelligence: An Introduction*. Wiley, 2007.

Brent E. Eskridge and Dean F. Hougen. Nurturing promotes the evolution of learning in uncertain environments. In *IEEE International Conference on Development and Learning / Epigenetic Robotics*, 2012.

Dario Floreano and Claudio Mattiussi. *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies*. The MIT Press, 2008.

J. F. Fontanari and R. Meir. Evolving a learning algorithm for the binary perceptron. *Network: Computation in Neural Systems*, 2(4):353–359, 1991.

Kunihiko Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36:193–202, 1980.

Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 2006.

John Henry Holland. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, 1975.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), November 1998.

Armand Leonce, Bryan Hoke, and Dean F. Hougen. Evolution of robot-to-robot nurturing and nurturability. In *IEEE International Conference on Development and Learning / Epigenetic Robotics*, 2012.

Warren S. McCulloch and Walter H. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 1943.

P. McQuesten and R. Miikkulainen. Culling and teaching in neuro-evolution. In *International Conference on Genetic Algorithms*, pages 760–767, 1997.

Yael Niv, Daphna Joel, Isaac Meilijson, and Eytan Ruppin. Evolution of reinforcement learning in uncertain environments: A simple explanation for complex foraging behaviors. *Adaptive Behavior*, 10(1):5–24, 2002.

Justus H. Piater, Paul R. Cohen, Xiaoqin Zhang, and Michael Atighetchi. A randomized anova procedure for comparing performance curves. In *Machine Learning: Proceedings of the Fifteenth International Conference*, pages 430–438, 1998.

Frank Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 1958.

David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, October 1986.

Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.

Syed Naveed Hussain Shah. *Nurturing Promotes the Evolution of Learning in Changing Environments*. PhD thesis, University of Oklahoma, 2015.

Andrea Soltoggio, Kenneth O. Stanley, and Sebastian Risi. Born to learn: the inspiration, progress, and future of evolved plastic artificial neural networks. arXiv preprint arXiv:1703.10371, 2017.

Paul John Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Harvard University, 1975.

B. Widrow and M. E. Hoff. Adaptive switching circuits. In *IRE WESCON Convention Record*, volume 4, pages 96–104, August 1960.

Mark Woehrer, Dean F. Hougen, Ingo Schlupp, and Brent E. Eskridge. Robot-to-robot nurturing: Definition, importance, and evolution. In *IEEE International Conference on Development and Learning / Epigenetic Robotics*, 2012.

Xin Yao. Evolving artificial neural networks. In *Processings of the IEEE*, volume 87, pages 1423–1447, 1999.