Nurturing Promotes the Evolution of Generalized Supervised Learning

Bryan Hoke
Robotics, Evolution, Adaptation, and Learning Laboratory
School of Computer Science
Gallogly College of Engineering
University of Oklahoma
Norman, OK 73019–1101
Email: bryan.hoke@ou.edu

Dean F. Hougen
Robotics, Evolution, Adaptation, and Learning Laboratory
School of Computer Science
Gallogly College of Engineering
University of Oklahoma
Norman, OK 73019–1101
Email: hougen@ou.edu

***Abstract*—The ability to learn makes intelligent systems more adaptive. 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 also interesting from a theoretical perspective of understanding how the sophisticated learning seen in nature 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. Here we expand upon a foundational work in the evolution of learning to also enable the evolution of instincts and then examine 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 generalized supervised learning in these environments.**

I. Introduction

Numerous researchers have demonstrated that evolutionary processes can produce artificial systems that learn [1], [2]. Learning may be either specific to certain tasks or generalized for any task in the problem space. It is also possible that learning might not emerge if none of the behavior adjustment rules evolved perform better than adjusting behavior randomly.

In many works (e.g., [3]–[10]), 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, no one strategy would always be optimal for every possible kind of environment [11].

A disadvantage of learning is that individuals tend to have low fitness early in life before they have had the opportunity to learn correct behavior. This is in contrast to individuals that exhibit instinctive behavior, which are likely to 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*—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 [12]. Nurturing, often from a parent to a child, may increase the competitive viability of learning by protecting the learner from otherwise costly errors made before proper behavior has been acquired [12].

Moreover, learning may enable greater nurturing, as a learning individual may be able to obtain greater resources for use in nurturing or may be able to directly learn how to satisfy the needs of a particular individual it is nurturing. In this way, nurturing and evolution may form a virtuous cycle of self-reinforcing positive outcomes, with the evolution of nurturing as its entry point, as shown in Figure 1 [12].

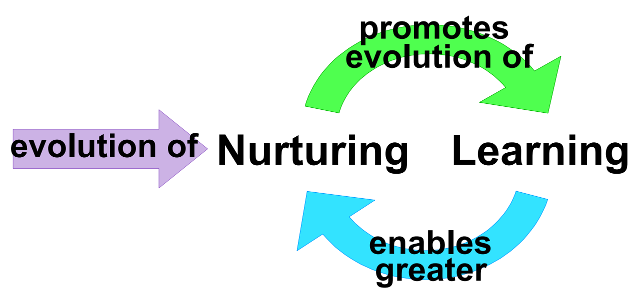


Fig. 1: Hypothesized virtuous cycle of nurturing and learning.

In order to understand a part of this putative cycle, this work examines how nurturing in the form of *safe exploration*—that is, exploration that does not count against an individual's fitness—influences the evolution of supervised learning and/or instincts. To explore these questions, this work uses a combination of genetic algorithms and artificial neurons applied to a set of supervised learning tasks.

II. Related Work

Any feature of an artificial neural network (ANN) can be subjected to an evolutionary process [13]. This work will focus on the evolution of connection weights and learning rules.

The connection weights of a neuron can be evolved as fixed values or as initial values that will be modified by the neuron's learning rule. This latter 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 this situation the values of the initial weights do not matter if the evolved learning rules will converge on solutions that are 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 [1], [2]. This work will examine the outcomes of the simultaneous evolution of initial connection weights and learning rules under different conditions.

A. The Evolution of Learning

Given the wide variety of artificial neural networks developed, 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 remain [2].

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 [14]–[17]. Evolved learning behavior that is scalable and adaptive to unknown or changing situations can be seen in countless species in nature [18].

Substantial and varied work on the evolution of learning has taken place since the early 1990s (e.g., [3]–[6], [19], [20]) and continues to this day (e.g., [2], [10], [21]), with work ranging from evolving learning rules for single neurons by encoding the parameters of a template formula [3] to evolving the structure of the rules themselves [6] to evolving the topology, components, and hyperparameters of deep ANNs [21]; covering supervised [3] and reinforcement learning [10], [22]; and showing that known rules could be rediscovered [3] as well as that novel rules could be discovered [19].

B. Nurturing

Nurturing is widespread in the biological world [18] and can be an important contributing factor to the evolution of learning [2], [12], [23]. Nurturing itself can be either instinctive or learned. Instinctive nurturing has been evolved in robots [24]. Nurturing can cover the initial costs of a learner and the resulting benefits of learning can be paid forward to the next generation, thereby promoting the evolution of learning [12].

It has been shown that neuroevolution can benefit from parent networks teaching their offspring using backpropagation [22]. Other research involving food patch estimation in uncertain environments demonstrated that nurturing as both social learning and safe exploration can promote the evolution of learning [23]. It has also been demonstrated that nurturing as task simplification can promote the evolution of reinforcement learning in changing environments [10], [17].

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

III. Hypotheses

Our hypotheses consider the relationship between nurturing and the quality of evolved instincts or generalized learning. It is expected that higher-quality generalized learning will evolve with nurturing than without because of the costs associated with the learning process without nurturing; individuals tend to have low fitness early in their life before they have had the opportunity to learn correct behavior, but nurturing means that individuals do not incur fitness penalties for mistakes made early in life. Likewise, it is expected that higher-quality instincts will evolve without nurturing than with 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.

$\mathbf{H_1}$: *When both instincts and learning are evolved, higher-quality generalized learning will evolve with nurturing than without.*

$\mathbf{H_2}$: *When both instincts and learning are evolved, higher-quality instincts will evolve without nurturing than with.*

H_1 and H_2 can be seen as counterparts to each other because they predict that the quality of evolved generalized learning rules and evolved instincts will exhibit opposite effects with respect to the presence or absence of nurturing. However, it is important to note that H_1 and H_2 have no dependencies on each other because they make predictions about generalized learning and instincts 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_1 would be supported but H_2 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* [25], [26], in which case H_2 would be supported but H_1 would not.

These hypotheses 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 [27].

IV. Approach

While previous authors had shown that nurturing can promote the evolution of reinforcement learning [10], [17], [22], [23], we have chosen to investigate whether that is true for supervised learning as well. Moreover, the only previous work to look at nurturing as safe exploration considered only whether evolution would turn learning on or off, not whether it would suitably evolve a generalized learning rule when task-specific options were available [23], and we wished to provide a broader range of possibilities from which evolution could select. For these reasons, the approach described here builds on the foundational paper of Chalmers that showed the evolution of supervised learning rules [3]. On that foundation, we have added the possibility for evolving instincts and contrasted nurturing and non-nurturing conditions.

A. Operational Definitions

An *individual* is a candidate solution to the genetic algorithm's fitness function whose characteristics are represented by a chromosome [28]. 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. The tasks used here were those used by Chalmers [3]. 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, *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, *weight-generalized 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; *task-generalized learning ability* is measured by evaluating the 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.

B. 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 a_j the activation of input unit j, o_i the activation of 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 [3], 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 k.

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

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

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

The portion of the genome that 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

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

C. 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, each in its own region of each chromosome.

Initial weights are encoded using 3 bits each. The first bit is the sign 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.

input : $chromosome$,
$task$,
$isNurturingCondition$,
$areInitialWeightsEvolved$
output: The fitness of the chromosome on the task

```
fitness = 0.0
denominator = 1
if areInitialWeightsEvolved == true then
    initialWeights = DecodeWeights(chromosome, task)
else
    initialWeights = InitializeRandomWeights(task)
end
learningRule = DecodeLearningRule(chromosome)
neuron = InitializeNeuron(initialWeights,
  learningRule)
for i = 0 to 9 do
    if isNurturingCondition == false then
        fitness += MeasureTaskAccuracy(neuron,
          task)
        denominator += 1
    end
    TrainNeuron(neuron, task)
end
fitness += MeasureTaskAccuracy(neuron, task)
return fitness / denominator
```

Algorithm 1: The procedure for evaluating the fitness of a chromosome on a given task.

D. Evaluation of Fitness

The procedure for evaluating the fitness of a chromosome for a particular task is shown in Algorithm 1. 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 after learning, whereas in the non-nurturing case the individual's fitness is the average of the evaluations during all training epochs and the final evaluation after learning.

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.

E. Parameters and Process of the Genetic Algorithm

A fixed population size of 40 is used in all experiments. Following the fitness evaluation process, 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 that are inserted into the new population;

(2b) the remaining 7 selected individuals are reproduced by cloning to create 7 new individuals that 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 change.

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

F. 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 of the nurturing condition and the non-nurturing condition.

After each evolutionary run, the individual with the highest fitness score in the history of the populations in that run is identified. This individual is evaluated on the evolutionary tasks and test 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*.

To measure weight-generalized learning ability, the individual's learning rule genes are isolated and evaluated on the evolutionary tasks (using both the intra-learning and post-learning tests) by initializing a set of neurons with random initial weights and the individual's 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.

To measure task-generalized learning ability, the individual's learning rule genes are isolated and evaluated on the test tasks (using both the intra-learning and post-learning tests) by initializing a set of neurons with random initial weights and the individual's learning rule, with the intention of determining the capacity for learning that is generalizable to new tasks.

To measure instinct quality, the individual's initial-weight genes are isolated and evaluated on the evolutionary tasks by initializing a set of neurons with the encoded weights (and no learning rule), with the intention of determining the contribution of just the portion of the chromosome that encodes the initial weights to the individual's total fitness.

Each evolutionary run is repeated 30 times to collect statistically meaningful results.

V. Results and Discussion

This section considers first the evolution of adaptions, then looks at the results of tests examining the hypotheses.

A. Evolutionary Runs

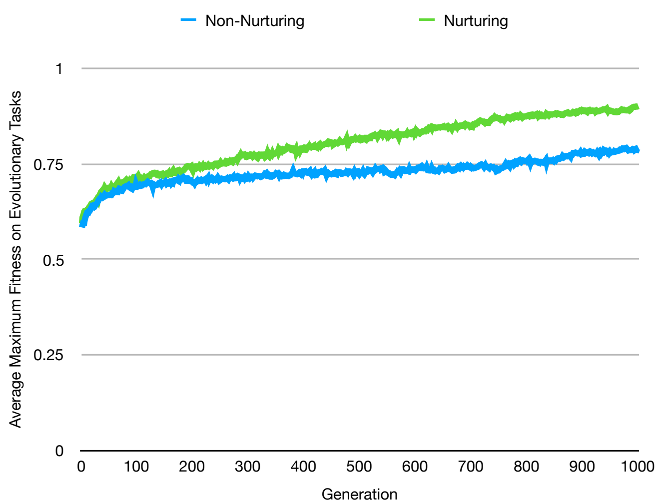


Fig. 2: The maximum fitness for populations evolved using 20 evolutionary tasks.

Figure 2 shows how the average maximum fitness in each population changes over the generations. Here, the populations are evolved using 20 tasks under either the nurturing or non-nurturing condition and the fitness shown is the average over 30 repetitions of the most fit individual in each population each generation, considering the individual's fitness when evaluated on the evolutionary tasks used during evolution.

Figure 2 shows the increase of maximum fitness with the progression of generations, which demonstrates the increasing adaptation of the individuals to their environment. This indicates that the populations are becoming adapted to their environments over the generations.

B. 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 3. 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.

As shown in Figure 3, the improvement in fitness due to learning increases as the number of evolutionary tasks increases under both the nurturing and non-nurturing conditions. This suggests that learning becomes an increasingly important evolutionary strategy as the environment becomes more varied.

The differences between the results for the nurturing and non-nurturing conditions are statistically significant (randomized ANOVA, algorithm $p < 0.001$, interaction $p < 0.001$). This suggests that learning makes a greater contribution toward fitness under the nurturing condition than under the non-nurturing condition.

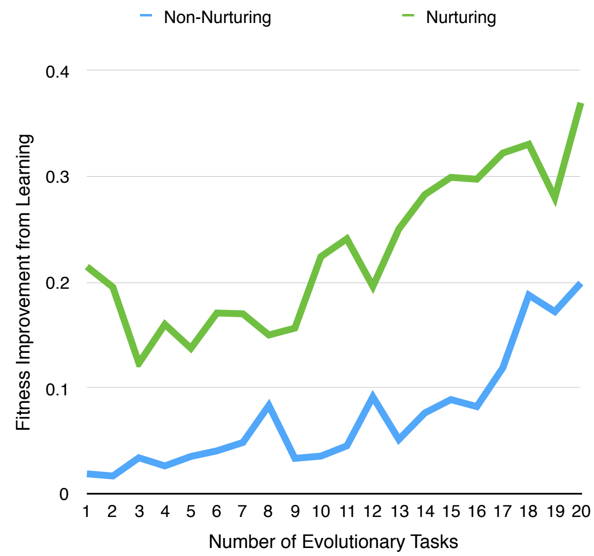


Fig. 3: The difference between post-learning fitness and pre-learning fitness.

C. Weight-Generalized Learning Test

Figures 4a and 4b show how the average weight-generalized fitness of the most fit individual in an evolutionary run changes with the number of evolutionary tasks on which that individual was evolved. Here, the populations are evolved using 1 to 20 tasks under either the nurturing or non-nurturing condition and the fitness shown is the average over 30 repetitions of the weight-generalized evolutionary fitness of the most fit individual in each evolutionary run as evaluated by the intra-learning test (Figure 4a) and the post-learning test (Figure 4b).

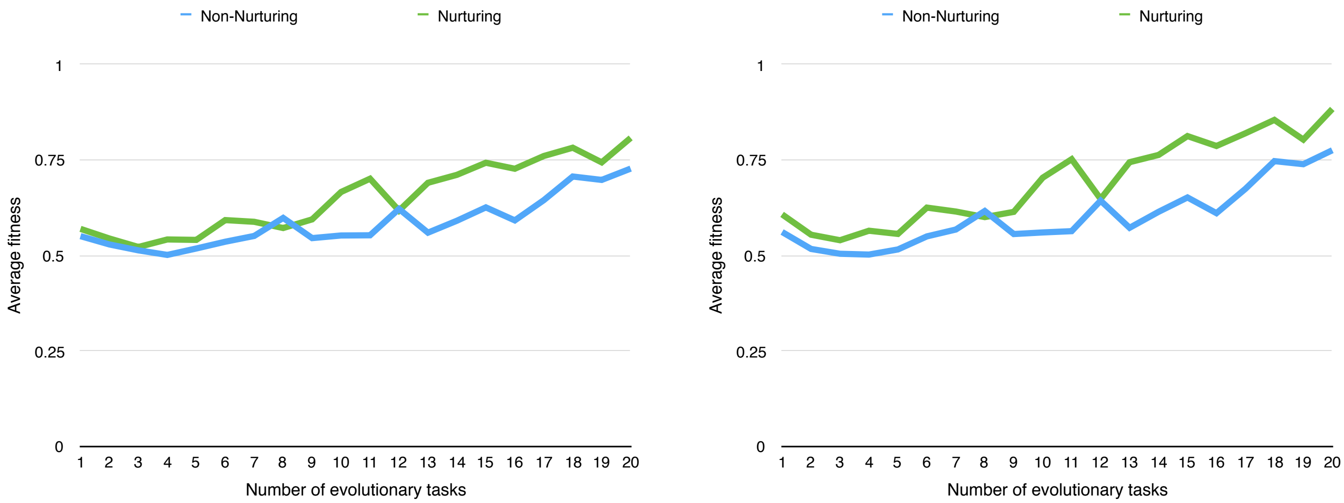
Figures 4a and 4b show that the learning rules evolved under the nurturing condition perform better than the learning rules evolved under the non-nurturing condition when using both the intra-learning and post-learning tests.

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

The statistically significant improvement of the weight-generalized fitness evolved under the nurturing condition compared to that evolved under the non-nurturing condition suggests that the learning rule makes a larger contribution to fitness under the nurturing condition than it does under the non-nurturing condition.

D. Task-Generalized Learning Test

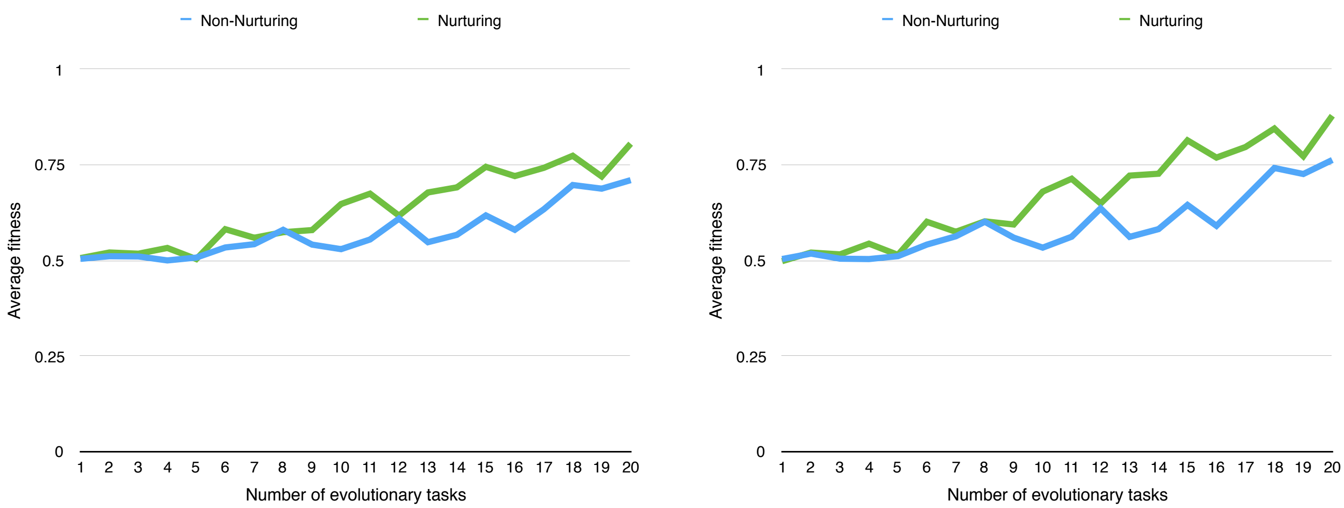
Figures 5a and 5b show how the average task-generalized fitness of the most fit individual in an evolutionary run changes with the number of evolutionary tasks on which that individual was evolved. Here, the populations are evolved using 1 to 20 tasks under either the nurturing or non-nurturing condition and the fitness shown is the average over 30 repetitions of the task-generalized fitness of the most fit individual in each



(a) Intra-learning.

(b) Post-learning.

Fig. 4: Weight-generalized learning test (on evolutionary tasks).



(a) Intra-learning.

(b) Post-learning.

Fig. 5: Task-generalized learning test (on non-evolutionary tasks).

evolutionary run as evaluated by the intra-learning test (5a) and the post-learning test (5b).

This data is directly relevant to hypothesis H_1 because it is interpreted as a measure of the generalized learning capabilities evolved in each evolutionary run.

Figures 5a and 5b show the increasing quality of generalized learning evolved as the number of evolutionary tasks increases, demonstrating that more complex environments provide more incentive for the evolution of generalized learning.

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

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, showing support for hypothesis H_1 which predicts that better generalized learning is expected to evolve under the nurturing condition and is the key hypothesis that this research set out to investigate. These results strongly support the idea that nurturing promotes the evolution of learning under these conditions.

E. Instinct Quality Test

Figure 6 shows how the pre-learning fitness of the most fit individual in each evolutionary run changes with the number of evolutionary tasks on which that individual was evolved. Here, the populations are evolved using 1 to 20 tasks under either the nurturing or non-nurturing condition and the fitness shown is the average over 30 repetitions of the pre-learning fitness of the most fit individual in each evolutionary run.

This data is relevant to hypothesis H_2 because it is interpreted as a measure of the quality of the instincts evolved.

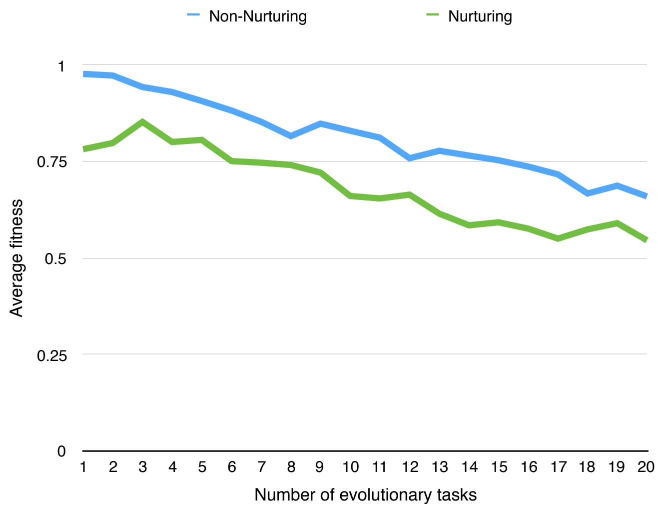


Fig. 6: Instinct quality test (on evolutionary tasks).

Figure 6 shows the decreasing quality of evolved instincts as the number of evolutionary tasks increases, demonstrating that it becomes more difficult to evolve suitable instincts as the number of tasks in the evolutionary environment increases.

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 statistically significantly higher performance of the pre-learning fitness evolved under the non-nurturing condition than that evolved under the nurturing condition shows support for hypothesis H_2.

Individuals evolved using the non-nurturing condition performed better on this evaluation than individuals evolved using the nurturing condition by a statistically significant margin, suggesting that instincts (initial weights) make a larger contribution to fitness in the non-nurturing case than in the nurturing case.

VI. Conclusions

This work shows strong support for both of our specific hypotheses:

$\mathbf{H_1}$: *When both instincts and learning are evolved, higher-quality generalized learning will evolve with nurturing than without.*

$\mathbf{H_2}$: *When both instincts and learning are evolved, higher-quality instincts will evolve without nurturing than with.*

It also supports our overall concept that nurturing promotes the evolution of learning.

This research explicitly considers the interaction of nurturing, learning, and instincts as does Shah [17] but is distinguished from Shah by the type of learning evolved (here, supervised learning; in Shah, reinforcement learning). This is the first research to consider the impact of nurturing on the evolution of supervised learning and is the first to measure the quality of the instincts and learning evolved as a function of task complexity.

It was found that nurturing promoted the evolution of generalized supervised learning compared to the absence of nurturing, as was expected. This provides concrete evidence that nurturing promotes the evolution of learning. This is in support of previous findings by Eskridge and Hougen [23], who demonstrated that nurturing promotes the evolution of learning in uncertain environments; Shah and Hougen [10], who demonstrated that nurturing promotes the evolution of learning in changing environments; and McQuesten and Miikkulainen [22], who first demonstrated that nurturing could promote the evolution of learning in an artificial system (although they did not use the unifying term “nurturing” to describe their work). Moreover, this supports the overall research agenda laid out by Woehrer et al. [12] in their call to the research community that first postulated that nurturing promotes the evolution of learning. This same idea has also recently been echoed by Soltoggio, Stanley, and Risi [2], apparently independently.

The quality of generalized learning evolved was found to be proportional to the number of tasks in the environment in both the nurturing and non-nurturing conditions, reconfirming the original findings of Chalmers [3] and demonstrating that they are valid even under the non-nurturing condition and with the possibility of evolving instincts.

Instincts evolved under the non-nurturing condition were found to be of higher quality than those evolved under the nurturing condition, and the quality of instincts evolved was found to be inversely proportional to the number of tasks in the environment. These appear to be original results because, while Niv et al. [29] and Shah [17] both allow for the evolution of instincts in their research, Niv et al. do not consider nurturing versus non-nurturing conditions and neither attempt to measure the quality of the evolved instincts. While Niv et al. and Shah both showed that learning could evolve even when the evolution of instincts was possible, neither considered their interaction with task numbers or nurturing, so this work provides clarification of the role that instincts may play in the evolutionary process.

These results are both practical and theoretical contributions. They help to fill in the picture regarding how and when learning and instincts evolve and the role that nurturing can play in this process. In addition, they should help researchers understand how to design their evolutionary environments in order to maximize the quality of generalized learning and/or instinctive behavior that is likely to be evolved.

VII. Future Work

There are numerous important theoretical concepts related to the evolution of learning [1], [2], [14]–[16] that should be considered in light of the proposed virtuous cycle [12].

For example, Baldwin argued that learning accelerates evolution because suboptimal individuals increase their baseline fitness by acquiring more adaptive characteristics during life [25], [26]. 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 what were once learned behaviors may gradually become assimilated into the genotype, reducing the need for learning, and ultimately resulting in a loss of learning capability [7], [20], [30]; this "second step" of the Baldwin effect [30] is not a beneficial effect if it is desirable for the evolved individuals to retain the ability to learn over evolutionary time.

Nurturing can reduce the evolutionary costs of learning and thus may be able to inhibit the undesirable second step of the Baldwin effect. Inhibiting this aspect of the Baldwin effect may result in evolved individuals that retain greater learning capabilities throughout the evolutionary process.

In addition, to ensure the generality of the conclusions found here, there are a number of ideas for future work that should be pursued using straightforward variations on this work.

This work investigated the role of nurturing as safe exploration in the evolution of learning, as did Eskridge and Hougen [23]. However, additional investigations should consider other forms of nurturing such as social learning, perhaps by having parent individuals use their learned weights to teach their offspring as in McQuesten and Miiklulainen [22]; task simplification as in Shah and Hougen [10]; etc.

This work investigated the role of nurturing in the evolution of supervised learning and previous work has considered the role of nurturing in the evolution of reinforcement learning [10]; additional investigations should consider the role of nurturing in the evolution of other forms of learning, such as unsupervised learning [28].

This work investigated the role of nurturing in the evolution of single neurons tasked with learning simple, linearly-separable boolean functions; future work should investigate the role of nurturing in the evolution of multi-layer neural networks tasked with optimizing more complex tasks, such as those investigated by Miikkulainen et al. [21].

References

[1] G. Weiss, "Towards the synthesis of neural and evolutionary learning," in *Progress in Neural Networks*. Norwood, NJ, USA: Ablex Publishing Corporation, 1997, vol. 5: Architecture, pp. 145–176.

[2] A. Soltoggio, K. O. Stanley, and S. Risi, "Born to learn: the inspiration, progress, and future of evolved plastic artificial neural networks," 2017, arXiv preprint arXiv:1703.10371.

[3] D. J. Chalmers, "The evolution of learning: An experiment in genetic connectionism," in *Proceedings of the 1990 Connectionist Models Summer School*, M. Kaufmann, Ed. Morgan Kaufmann, 1990, pp. 81–90.

[4] J. F. Fontanari and R. Meir, "Evolving a learning algorithm for the binary perceptron," *Network: Computation in Neural Systems*, vol. 2, no. 4, pp. 353–359, 1991.

[5] S. Bengio, Y. Bengio, J. Cloutier, and J. Gecsei, "On the optimization of a synaptic learning rule," in *Optimality in Artificial and Biological Neural Networks*. Univ. of Texas, 1992, pp. 6–8.

[6] S. Bengio, Y. Bengio, and J. Cloutier, "Use of genetic programming for the search of a new learning rule for neural networks," in *IEEE World Congress on Computational Intelligence, Proceedings of the First IEEE Conference on Evolutionary Computation*. IEEE, 1994, pp. 324–327.

[7] D. Floreano and J. Urzelai, "Evolutionary robots with on-line self-organization and behavioral fitness," *Neural Networks*, vol. 13, no. 4, pp. 431–443, Jun. 2000.

[8] S. Risi, C. E. Hughes, and K. O. Stanley, "Evolving plastic neural networks with novelty search," *Adaptive Behavior*, vol. 18, no. 6, pp. 470–491, Dec. 2010.

[9] O. Gigliotta, G. Pezzulo, and S. Nolfi, "Evolution of a predictive internal model in an embodied and situated agent," *Theory in Biosciences*, vol. 130, no. 4, pp. 259–276, Dec. 2011. [Online]. Available: https://link.springer.com/article/10.1007/s12064-011-0128-x

[10] S. N. H. Shah and D. F. Hougen, "Nurturing promotes the evolution of reinforcement learning in changing environments," in *2017 IEEE Symposium on Computational Intelligence in Control and Automation (CICA)*. IEEE, 2017.

[11] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, 1997. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/585893/

[12] M. Woehrer, D. F. Hougen, I. Schlupp, and B. E. Eskridge, "Robot-to-robot nurturing: Definition, importance, and evolution," in *IEEE International Conference on Development and Learning / Epigenetic Robotics*, 2012.

[13] X. Yao, "Evolving artificial neural networks," in *Proceedings of the IEEE*, vol. 87, 1999, pp. 1423–1447.

[14] R. Dukas, "Evolutionary ecology of learning," in *Cognitive Ecology: The Evolutionary Ecology of Information Processing and Decision Making*, R. Dukas, Ed. Chicago: University of Chicago Press, 1998, pp. 129–174.

[15] B. Kerr and M. W. Feldman, "Carving the cognitive niche: Optimal learning strategies in homogeneous and heterogeneous environments," *Journal of Theoretical Biology*, vol. 220, no. 2, pp. 169–188, Jan. 2003.

[16] K. Ellefsen, "Balancing the costs and benefits of learning ability," in *12th European Conference on Artificial Life*. MIT Press, Sep. 2013, pp. 292–299.

[17] S. N. H. Shah, "Nurturing promotes the evolution of learning in changing environments," Ph.D. dissertation, University of Oklahoma, 2015.

[18] J. B. Reece and N. Campbell, A., *Biology*, 9th ed. Boston, Mass.; London: Pearson, 2011.

[19] J. Baxter, "The evolution of learning algorithms for artificial neural networks," *Complex Systems*, pp. 313–326, 1993.

[20] F. Cecconi, F. Menczer, and R. K. Belew, "Maturation and the evolution of imitative learning in artificial organisms," *Adaptive Behavior*, vol. 4, no. 1, pp. 29–50, 1995.

[21] R. Miikkulainen, J. Liang, E. Meyerson, A. Rawal, D. Fink, O. Francon, B. Raju, H. Shahrzad, A. Navruzyan, N. Duffy, and B. Hodjat, "Evolving deep neural networks," *arXiv:1703.00548 [cs]*, Mar. 2017, arXiv: 1703.00548.

[22] P. McQuesten and R. Miikkulainen, "Culling and teaching in neuro-evolution," in *International Conference on Genetic Algorithms*, 1997, pp. 760–767.

[23] B. E. Eskridge and D. F. Hougen, "Nurturing promotes the evolution of learning in uncertain environments," in *IEEE International Conference on Development and Learning / Epigenetic Robotics*, 2012.

[24] A. Leonce, B. Hoke, and D. F. Hougen, "Evolution of robot-to-robot nurturing and nurturability," in *IEEE International Conference on Development and Learning / Epigenetic Robotics*, 2012.

[25] J. M. Baldwin, "A new factor in evolution," *The American Naturalist*, vol. 30, no. 354, pp. 441–451, 1896. [Online]. Available: http://www.jstor.org/stable/2453130

[26] ——, "A new factor in evolution (continued)," *The American Naturalist*, vol. 30, no. 355, pp. 536–553, 1896. [Online]. Available: http://www.jstor.org/stable/2453231

[27] J. H. Piater, P. R. Cohen, X. Zhang, and M. Atighetchi, "A randomized anova procedure for comparing performance curves," in *Machine Learning: Proceedings of the Fifteenth International Conference*, 1998, pp. 430–438.

[28] A. P. Engelbrecht, *Computational Intelligence: An Introduction*. Wiley, 2007.

[29] Y. Niv, D. Joel, I. Meilijson, and E. Ruppin, "Evolution of reinforcement learning in uncertain environments: A simple explanation for complex foraging behaviors," *Adaptive Behavior*, vol. 10, no. 1, pp. 5–24, 2002.

[30] R. Suzuki and T. Arita, "Interactions between learning and evolution: The outstanding strategy generated by the Baldwin effect," *Biosystems*, vol. 77, no. 1-3, pp. 57–71, Nov. 2004. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0303264704000565