Nurturing Promotes the Evolution of Learning in Uncertain Environments

Brent E. Eskridge
CSNE Department
Southern Nazarene University
Bethany, OK 73008
Email: beskridge@snu.edu

Dean F. Hougen
School of Computer Science
University of Oklahoma
Norman, OK 73019-1101
Email: hougen@ou.edu

***Abstract*—Adapting to a changing and uncertain environment is vital for the long-term success of individuals, whether they are biological or artificial. While learning can be powerful in the adaptation process, a lack of understanding exists in the factors that promote or inhibit its evolution. Nurturing is widely thought to be a contributing factor, if not a requirement, for high levels of learning. In the study described here, we investigated how nurturing contributes to the evolution of learning in uncertain environments using a simple, biologically inspired evolutionary simulation. The simulation predicts that the two models of nurturing used here, nurturing as social learning and nurturing as a means of safe exploration, both promote the evolution of learning in uncertain environments in which learning would otherwise not be a viable strategy. The results also indicate that nurturing is also a factor in improving fitness in environments in which learning is already viable.**

I. Introduction

The evolution of biological learning is poorly understood [1]. Several theoretical models exist (e.g., [1]–[4]) and recent empirical studies (e.g., [1], [5], [6]) have begun to influence theory in this domain, but many questions remain. In the artificial domain there have been several empirical investigations into the evolution of machine learning, particularly in the evolution of various learning rules for artificial neural networks (e.g., [7]–[11]). Besides artificial neural networks, other encodings for complex representations are possible and learning rules have been evolved for them as well [12]. Although there have been a wide range of learning techniques developed in recent years, the evolution of learning is of particular interest since it allows for a better exploration of the different types of learning and can serve as a significant step towards removing humans from the need to participate in the learning process. However, a great deal remains unknown about the factors that promote or inhibit the evolution of learning in both the natural and artificial domains. A better understanding of the internal and environmental conditions that affect the evolution of learning would be of immense benefit in promoting learning in artificial systems.

In the biological domain, it is widely believed that high parental investment per offspring (substantial biological nurturing) and long juvenile phases (extended nurturable periods) are necessary prerequisites for the evolution of high levels of intelligence and learning [13]. In the artificial domain, extensive robot learning requires oversight during the learning process and this oversight could be provided through robot-to-robot nurturing [14]. Nonetheless, the role of nurturing in the evolution of learning is underexplored, with a few notable exceptions (e.g., [4]).

We define *nurturing* 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. Nurturing is prevalent in the biological world and, accordingly, various aspects of nurturing are extensively studied in biology, cognitive science, psychology, child development, sociology, and education, among other disciplines.[1] Prototypically in biology nurturing would involve parental investment in offspring, while in education it would involve instructors teaching students. Here we investigate how two distinct models of nurturing could promote the evolution of learning in uncertain environments.

II. Hypotheses & Approach

Although there are a variety of potential nurturing models, two models motivated from observations of biological systems (e.g., meerkat societies [15]) were chosen for evaluation. The following models were chosen because of their simplicity and the ease with which they could be combined with the abstract simulations used in this study: *a*) nurturing as a method of social learning in which offspring learn to mimic the parent, and *b*) nurturing as a means of giving offspring the opportunity to experience the environment through exploration without risk. While both models provide an opportunity for an individual to learn before being exposed to the risks associated with experiencing the real world, they differ in what can be learned. In nurturing as social learning, offspring are restricted to learning only from a parent. While this can be straightforward, especially in artificial domains, it is limited to the information, and its accuracy, that the parent used to construct its model of the environment. On the other hand, nurturing as safe exploration offers offspring the opportunity to directly experience the environment without risk. Although this is the preferable option of the two, its implementation

[1] While various aspects of nurturing are studied in various disciplines under their own domain-specific terms, there is no cross-disciplinary term that captures this concept—hence our need to rigorously define this term.

is more complicated, especially with the understanding that, ultimately, nothing is without risk and there is always a cost.

Using these nurturing models, we explore the following hypotheses in this study:

- $\mathbf{H_1}$ Nurturing as social learning promotes the evolution of learning in uncertain environments.
- $\mathbf{H_2}$ Nurturing as safe exploration promotes the evolution of learning in uncertain environments.

To investigate these hypotheses, we performed evolutionary simulations designed using an experiment from the biological domain as inspiration. Dunlap and Stephens [1] evaluated the value of a learning strategy in uncertain, changing environments using *Drosophila*. They identified the fixity of the best action and the reliability of experience as the primary methods of change in an environment that affect the value of learning. As such, this study also uses these factors to determine the uncertainty of an environment.

Fixity of the best action, denoted F, can also be described as the consistency of the environment. A high level of fixity indicates that the environment is unlikely to change throughout an individual's lifetime, whereas a low fixity indicates that the environment is likely to change frequently throughout an individual's lifetime. In contrast, the reliability of experience, denoted R, can be described as the confidence an individual has that the experience it gains in the environment is indicative of the experience it can expect in the future. A high level of reliability indicates that experience is likely to be a true indicator of the environment, whereas a low reliability indicates that experience is unlikely to be a true indicator. Although experience and the environment are related, they are distinct in that experience is only relevant in learning when it is an accurate indicator of the environment. For example, a forest fire would be related to fixity as it would change the environment for many years to come, while a drought would be related to reliability as experience in the environment during a drought would not be indicative of experience in the environment in the normal, non-drought years. Experimental results have shown that low fixity environments promote learning as a means of adaptation, while high fixity environments promote non-learning since adaptation can be accomplished through evolution [1]. Experimental results have also shown that high reliability environments promote learning, while low reliability environments promote non-learning since experience gained through the exploration involved in learning is unreliable. To confirm our hypotheses, nurturing must compensate for the environmental bias due to one or both of these avenues of environmental change.

III. Experimental Setup

To test our hypotheses with different combinations of fixity and reliability, a simulation was used in which an individual estimated a single food patch's yield over its lifetime. The patch's yield varied over the course of a year. This variation was consistent from year to year over the course of the individual's lifetime. Depending on the fixity of the environment, the patch's yield for a particular time period could permanently change, necessitating adaptation through either the genotype or phenotype (i.e., learning). Also, depending on the reliability of the environment, the patch's yield for a particular time period could temporarily change for a single year. Individual's were evaluated on the accuracy of their estimates over their lifetime. Within this simulation, individuals could achieve high accuracy either by instinctually knowing the correct estimate using the genotype, or by learning the correct estimate. If the environment promoted learning, learning individuals should have higher fitness and evolution would select for learning. However, if the environment promoted adaptation of the genome over learning, evolution would select against learning.

TABLE I: Experimental parameters for the evolutionary simulations are shown.

Parameter	Value
Time periods per year	20
Years per generation	10
Years spent nurturing	2
Population	100
Generations	1000
Genome size	161
Codon size	8
Crossover	80%
Elite cloning	20%
Mutation rate	1%
Tournament size	6
Learning rate	0.80

To determine the effect of nurturing on the evolution of learning, the following three treatments were evaluated: *a*) no nurturing, *b*) nurturing as social learning, and *c*) nurturing safe exploration. The treatment without nurturing acted as a control against which the nurturing treatments could be compared. Individuals that were nurtured using social learning were given the opportunity to learn from their parent for a few years without any consequences to their fitness. This method of nurturing consisted of the offspring learning to mimic the parent using the parent's estimate of the food patch's yield for a given time period to update its own estimates. In contrast, individuals that were nurtured using safe exploration were given the opportunity to experience the environment for a few years without any consequences to their fitness. This method of nurturing consisted of the offspring gaining direct experience from the food patch yields just as in the non-nurturing case, but without any consequences to their fitness.

Individuals capable of learning used the following update rule for each time period's patch yield:

$$e_i' = (1 - \lambda)e_i + \lambda o_i \quad (1)$$

where λ was the learning rate, e_i was the individual's previous yield estimate at time period i, o_i was the observed yield at time period i, and e_i' was the individual's updated yield estimate. This update rule is commonly used in machine learning and in describing natural systems (e.g., [16], [17]). All learning individuals used a fixed learning rate of 0.8, which emphasized the most recent experience over past experiences.

A complete listing of the experimental parameters can be found in Table I.

A genetic algorithm (GA) [18] was used to evolve the initial estimates for the patch's yield at each time period in a year and determine if the individual used learning to update these estimates. Patch yield estimates were encoded in binary with eight bits per codon, which corresponded to a single time period. Estimates were encoded using a gray code and were normalized to the range $[0, 1]$ when decoded. These estimates represented the initial estimates used by learning individuals and the fixed estimates used by non-learning individuals. A single bit in the genome denoted whether or not the individual used learning. Individuals within a generation experienced 20 time periods per year for 10 years. Fitness was calculated as the negative of the root mean squared error over all the time periods in which the individual was evaluated. This ensured that although nurtured individuals were evaluated in fewer time periods, the fitness between nurtured and non-nurtured individuals could be compared without bias.

At the start of each generation, a random number for each time period was drawn from the uniform distribution in the range $[0, 1]$ and was compared to the fixity level F of the environment. If the random number exceeded F, the patch's yield for that time period was permanently changed to a random value. Furthermore, for every year, a second random number was drawn from the uniform distribution in the range $[0, 1]$ and was compared to the reliability level R of the environment. If the random number exceeded R, the patch's yield for that time period was temporarily changed to a random value. Thus, if an environment exhibited low fixity, but high reliability, a learning individual had an advantage since it had the ability to adapt to a changing environment using experience as a reliable indicator of the change in the environment. However, if an environment exhibited high fixity, but low reliability, a learning individual was at a disadvantage since it would attempt to adapt to perceived changes in the environment, when the best course of action was to ignore the unreliable experience and remain consistent. As such, the learning gene acted as a means to evolutionarily select for or against learning. Although individuals persisted beyond their own generation to "nurture" offspring in the nurturing treatments, their fitness was not evaluated further. This was because the individual had already had its reproductive opportunity and further fitness evaluations would not provide any further benefits. Individuals were selected for reproduction through tournament selection. In tournament selection, a specific number of individuals, in this case six, are chosen from the population and the individual with the highest fitness is selected for reproduction. Note that this selection method is not dependent on the magnitude of the differences in fitness, only that there are differences.

Data was collected every five generations over a total of 1,000 generations for each evolutionary run and each treatment consisted of 30 evolutionary runs, each performed with a different random seed. All individuals in a population for a given evolutionary run experienced the same patch yield values. Given the stochastic nature of the environment and the possibility of yield value changes in the final generation affecting the results, results were drawn from the final 50 generations of each evolutionary run to ensure that any such changes did not skew the results.

IV. Results

Figure 1 shows the mean percentage of individuals in the final 50 generations' populations that used learning in each of the three treatments. In both the nurturing treatments, learning proved to be a viable strategy in all but the most low reliability and high fixity environments (see Figures 1c and 1e). Furthermore, the variation in learning percentage was primarily restricted to the environments that were previously heavily biased against learning (see Figures 1d and 1f).

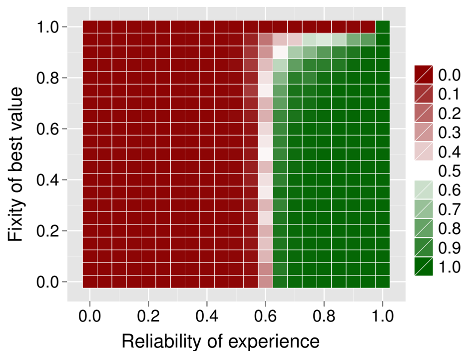
Figure 2 shows the p-values calculated by the bootstrapped Kolmogorov-Smirnov equality test. These indicate that the learning percentages differences visible in Figure 1 were statistically significant with $p < 0.01$ for environments that are theoretically classified as non-learning [1]. Also, the statistically significant differences between the nurturing models were sporadic and largely confined to environments that are thought to strongly promote non-learning.

Figure 3 shows the mean fitness and standard deviation for the best of generation individuals in the final 50 generations for each of the three treatments. While not shown in the figure, the mean fitness for a random, non-learning strategy was approximately -0.40. As one would expect, individuals in rapidly changing environments had far lower fitness levels than those in relatively stable environments, regardless of their ability to learn. Although the fitness for individuals in both the nurturing as social learning and nurturing as safe exploration models were comparable to individuals not nurtured in environments with low reliability, there was a marked improvement in environments with high reliability and low fixity (see fitness values for reliabilities of 0.95 and 1.0 in Figures 3c and 3e). Furthermore, the variation in fitness for these environments was high for individuals that used nurturing as social learning, while it was quite low for individuals that used nurturing as safe exploration.

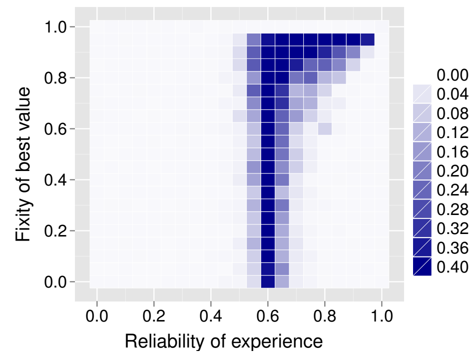
As with the learning percentage results, the bootstrapped Kolmogorov-Smirnov equality test was performed between each of the different types of nurturing for each fixity and reliability combination. Each nurturing type had a statistically significant difference in fitness with $p < 0.01$. While there were environments in which this difference was clear from Figure 3, such as those previously mentioned, this was not readily apparent in many of the environments in which individuals consistently had low fitness, regardless of the treatment. However, upon inspection, the relative differences in fitness values for these environments were small and the statistical significance was primarily due to the low standard deviations.

V. Discussion

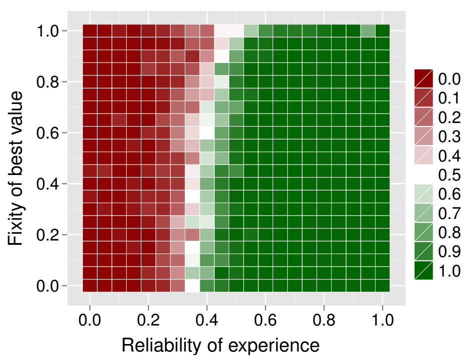
We draw three primary conclusions from these predictions. First, both hypotheses are confirmed as both nurturing as social learning and nurturing as safe exploration are predicted to



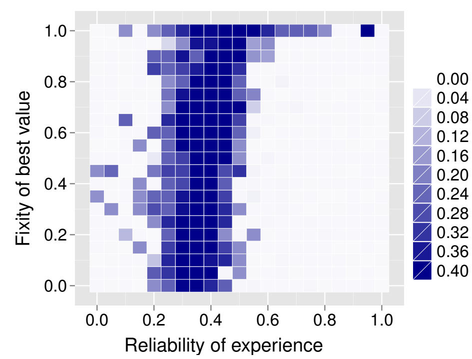
(a) Learning Percentage: No Nurturing



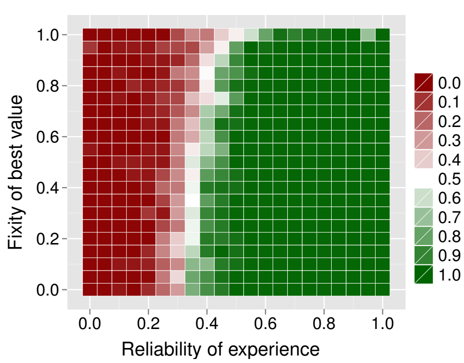
(b) Std Deviation: No Nurturing



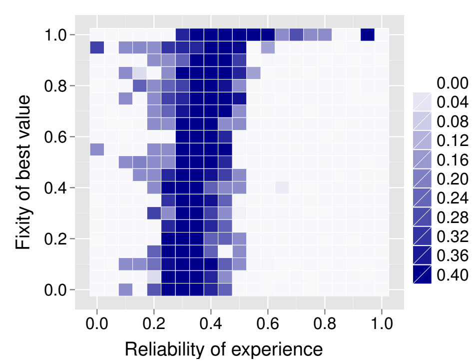
(c) Learning Percentage: Social Learning



(d) Std Deviation: Social Learning



(e) Learning Percentage: Safe Exploration



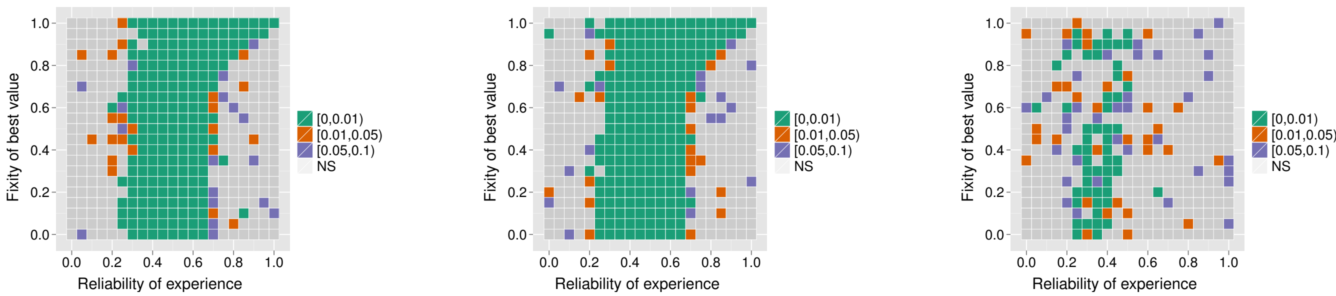
(f) Std Deviation: Safe Exploration

Fig. 1: The mean percentage of individuals in the final few generations' populations that use learning in each of the three treatments are shown. The standard deviation for each value is also shown to illustrate the variation of the learning percentage between evolutionary runs.

promote learning in environments in which it would otherwise not be a practical strategy. Not only did nurturing promote learning in low reliability environments, its use also resulted in higher fitness in high reliability environments. This indicates that nurturing is an integral component not only in the evolution of learning in uncertain environments, but, it is also an integral component to learning overall.

Although the fitness differences between learning and non-learning individuals were minor in the more uncertain environments, there were relatively few environments that were selectively neutral. While this may be indicative of the environment used or the learning rate used, our analysis leads us to conclude that the use of tournament selection contributed significantly to this result as it selects the most fit individual, regardless of the relative difference in fitness. Alternative selection mechanisms would most likely produce results that would be found in biological domains and expand the range of environments which could be considered selectively neutral.

Second, while both nurturing models resulted in statistically significant improvements, the use of nurturing as safe exploration resulted in more fit offspring. However, this was not surprising as it is always preferable to explore an environment without risk when compared to the opportunity to learn from an individual that uses potentially incorrect information. In



(a) None vs. Social Learning (b) None vs. Safe Exploration (c) Social Learning vs. Safe Exploration

Fig. 2: The p-values calculated by the bootstrapped Kolmogorov-Smirnov equality test in comparing the differences between the mean learning percentages for the treatments are shown. While p-values in the range $[0.05, 1)$ are generally not considered statistically significant, they are included here purely for informational reasons.

spite of this fact, nurturing as social learning still provided statistically significant improvements over the non-nurturing case. Furthermore, the fitness differences between individuals for each of these treatments were minor. Our analysis leads us to conclude that the improved fitness and the increased viability of learning as a strategy has more to do with how the environment is changed through nurturing than the means by which nurturing occurs. This indicates that merely ensuring a safe environment for learning provides significant benefits and may justify the investment required for nurturing.

Lastly, these results indicate that the reliability of experience in an environment has a larger effect both on the viability of learning and the fitness of an individual than the fixity of the best value or action. While this might be an artifact of the highly abstract nature of the environment used here, further investigation both in biological domains and in artificial domains that use higher-fidelity environments is warranted.

VI. Conclusions and Future Work

Learning is a vital component in adapting to a changing and uncertain environment in both the biological and artificial domains. Although the factors that promote or inhibit its evolution are poorly understood, previous work has shown that the fixity of the best action and the reliability of experience in an environment have a significant impact of the viability of learning [1]. In this study, we investigated how nurturing influences the evolution of learning in a simple food patch estimation simulation in uncertain environments using two models of nurturing: *a*) nurturing as a method of social learning in which offspring learn to mimic the parent, and *b*) nurturing as a means of giving offspring the opportunity to experience the environment through exploration without risk. Evolutionary simulations predict that both models of nurturing promote the evolution of learning in uncertain environments in which learning would otherwise not be a viable strategy at statistically significant levels. Furthermore, these predictions indicate that nurturing plays an integral role in the evolution of learning by improving the fitness of individuals in environments in which learning is already the preferred strategy.

The work described here is just the foundation for a much larger research effort investigating the role of nurturing in the evolution of learning in both biological and artificial domains. As such, there are a myriad of opportunities for future work. However, there are two directions of future work that are of immediate concern. First, the work described here assumed the existence of nurturing. To realize the full potential of nurturing, the situations that lead to its emergence need to be understood, along with the ways in which nurturing is realized, such as the social learning or safe exploration approaches used in this work. A variety of nurturing methods are known to exist, each with its own benefits and drawbacks. Second, the environments used in this work were highly abstract, which made the implementation of nurturing relatively easy. However, for nurturing to be truly useful, it must be effective in the complex and dynamic environments for which traditional learning methods are not effective.

References

[1] A. S. Dunlap and D. W. Stephens, “Components of change in the evolution of learning and unlearned preference,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 276, no. 1670, pp. 3201–3208, 2009.

[2] D. W. Stephens, “Change, regularity, and value in the evolution of animal learning,” *Behavioral Ecology*, vol. 2, no. 1, pp. 77–89, 1991.

[3] B. Kerr and M. W. Feldman, “Carving the cognitive niche: Optimal learning strategies in homogeneous and heterogeneous environments,” *Journal of Theoretical Biology*, vol. 220, pp. 169–188, 2003.

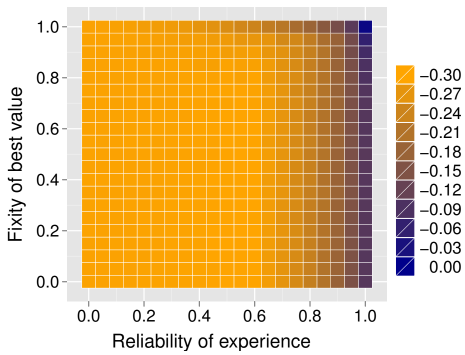
[4] E. Borenstein, M. W. Feldman, and K. Aoki, “Evolution of learning in fluctuating environments: When selection favors both social and exploratory individual learning,” *Evolution*, vol. 62, no. 3, pp. 586–602, 2012.

[5] F. Mery and T. J. Kawecki, “Experimental evolution of learning ability in fruit flies,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 22, pp. 14 274–14 279, Oct. 2002.

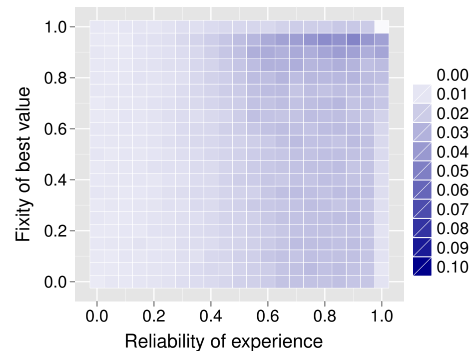
[6] ——, “The effect of learning on experimental evolution of resource preference in drosophila melanogaster,” *Evolution*, vol. 58, no. 4, pp. 757–767, 2004.

[7] D. Chalmers, “The evolution of learning: An experiment in genetic connectionism,” in *Proceedings of the 1990 Connectionist Summer School*. Morgan Kaufmann, 1990, pp. 81–90.

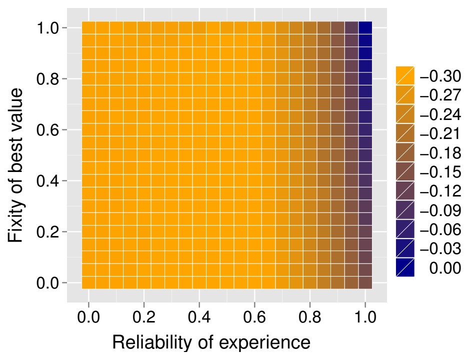
[8] J. Fontanari and R. Meir, “Evolving a learning algorithm for the binary perceptron,” *Network: Computation in Neural Systems*, vol. 2, no. 4, pp. 353–359, 1991.



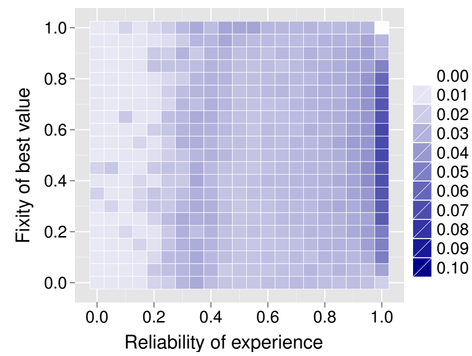
(a) Fitness: No Nurturing



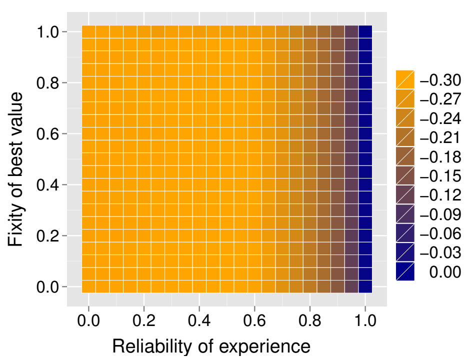
(b) Std Deviation: No Nurturing



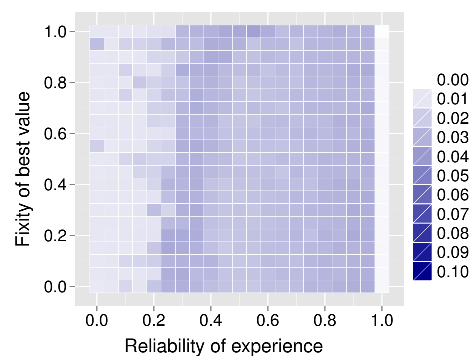
(c) Fitness: Social Learning



(d) Std Deviation: Social Learning



(e) Fitness: Safe Exploration



(f) Std Deviation: Safe Exploration

Fig. 3: The mean fitness for the best of generation individuals in the final few generations in each of the three treatments are shown. The standard deviation for each value is also shown to illustrate the variation of the fitness between evolutionary runs.

[9] S. Nolfi and D. Parisi, "Learning to adapt to changing environments in evolving neural networks," *Adaptive Behavior*, vol. 5, no. 1, pp. 75–97, 1996.

[10] 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, 2012.

[11] A. Soltoggio, P. Dürr, C. Mattiussi, and D. Floreano, "Evolving neuromodulatory topologies for reinforcement learning-like problems," in *IEEE Congress on Evolutionary Computation*, 2007, pp. 2471–2478.

[12] B. E. Eskridge, "Effects of state and action abstraction on development of controllers for concurrent, interfering, non-episodic tasks," PhD Thesis, University of Oklahoma, Norman, OK, 2009.

[13] N. A. Campbell and J. B. Reece, *Biology with Mastering Biology*, 8th ed. Benjamin Cummings, 2007.

[14] 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, to appear.

[15] A. Thornton and T. Clutton-Brock, "Social learning and the development of individual and group behaviour in mammal societies," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 366, no. 1567, pp. 978–987, 2011.

[16] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[17] S. Eliassen, C. Jørgensen, M. Mangel, and J. Giske, "Exploration or exploitation: life expectancy changes the value of learning in foraging strategies," *Oikos*, vol. 116, no. 3, pp. 513–523, 2007.

[18] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI, USA: The University of Michigan Press, 1975.