Interaction Between Evolution and Learning in NK Fitness Landscapes

Graham Todd[1], Madhavun Candadai[2,3] and Eduardo J. Izquierdo[2,3]

[1]Symbolic Systems Department, Stanford University
[2]Cognitive Science Program, Indiana University Bloomington
[3]The Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington
Corresponding email: gdrtodd@stanford.edu

Abstract

Artificial Life has a long tradition of studying the interaction between learning and evolution. And, thanks to the increase in the use of individual learning techniques in Artificial Intelligence, there has been a recent revival of work combining individual and evolutionary learning. Despite the breadth of work in this area, the exact trade-offs between these two forms of learning remain unclear. In this work, we systematically examine the effect of task difficulty, the individual learning approach, and the form of inheritance on the performance of the population across different combinations of learning and evolution. We analyze in depth the conditions in which hybrid strategies that combine lifetime and evolutionary learning outperform either lifetime or evolutionary learning in isolation. We also discuss the importance of these results in both a biological and algorithmic context.

Introduction

There are two approaches to generating adaptive behavior: evolution and learning. The robustness and flexibility of life on earth demonstrates that there is an advantage to combining the two. Baldwin (1896) argued that characteristics acquired during the individual's lifetime can alter the course of evolution, despite not being inherited, and simulation work in Artificial Life (ALife) has been consistent with this observation. Hinton and Nowlan (1987) were the first to demonstrate how learning alters evolution in an idealized simulation model. This work has been extended in several different directions, including by analzying dynamic environments (Suzuki and Arita, 2000, 2004), neural networks (Todd and Miller, 1991; Nolfi and Parisi, 1996; Littman, 1996), and ontogenetic development (Belew, 1990; Sendhoff and Kreutz, 1999). Of particular relevance to our contribution is the work of Hinton and Nowlan (1987) and its adaptation to the NK-landscape framework developed by Kauffman and Levin (1987). While previous ALife research has shown that combining lifetime and evolutionary learning can outperform either type in isolation, only very specific instantiations of lifetime and evolutionary learning have been studied in detail.

In the field of Artificial Intelligence (AI), meanwhile, the majority of work has focused on the use of learning and evolution as optimization techniques to tune the parameters of large neural systems: either evolution-inspired population search algorithms (Husbands et al., 1997; Stanley et al., 2019), or gradient descent as an individual learning technique in the form of supervised learning (Mitchell et al., 1997) or reinforcement learning (Sutton and Barto, 2018). Notably, Ackley and Littman (1991) explored early on the potential for combining the two and demonstrated the benefits of the hybrid approach in a discrete reinforcement learning task. More recently, the hybrid approach has bloomed within the deep learning and deep reinforcement learning literature (Fernando et al., 2017; Houthooft et al., 2018; Khadka and Tumer, 2018; Pourchot and Sigaud, 2018; Leite et al., 2020). However, these systems are more complex than necessary to systematically examine the trade-off between learning and evolution. Furthermore, unlike Hinton and Nowlan (1987)'s stochastic lifetime learning, the learning used to train artificial neural networks in supervised and reinforcement learning involves estimating the direction of steepest (de)ascent. In living systems, lifetime learning could be anywhere on the spectrum between steepest ascent and stochastic hill-climbing. Understanding the interaction between learning and evolution will require understanding the effect of different levels of determinism in individual learning.

Following the Modern Synthesis, the large majority of simulation work developed to understand the interaction between learning and evolution does not consider the inheritance of acquired characteristics, with a few exceptions (Gruau and Whitley, 1993; Whitley et al., 1994). However, to understand the interaction between evolution and learning in the broader perspective of AI, we must also consider a Lamarckian perspective, as algorithms selected for AI need not follow biological constraints. We also note that while there may not seem to be clear biological motivation for Lamarckian inheritance, in the broader context of evolving populations, there are some ways for acquired characteristics to be effectively inherited, such as behavioral inheritance through social learning and language-based information transmission (Jablonka et al., 1998).

We extend previous work in four important directions. First, we focus our research on a balanced trade-off between learning and evolution. That is, for each population we fix the total number of learning events and systematically vary how those events are distributed between individual lifetime learning and distributed evolutionary learning. Second, we examine the effect of task-difficulty on the optimal trade-off between learning and evolution. Third, we examine both Lamarckian and non-Lamarckian conditions as a way to assess their effect on the interaction between learning and evolution. Finally, we compare the effect of steepest ascent versus stochastic ascent for individual learning on its interaction with evolution.

Methods

In this section we describe our idealized model of lifetime and evolutionary learning. Following previous work (Curran et al., 2007a,b), we model the *population* as a set of individual genotypes and corresponding phenotypes moving in a fitness landscape. Specifically, we modeled a group of 50 individuals exploring the problem space through lifetime and evolutionary learning. In this section, we describe the NK fitness landscape, the lifetime learning forms examined, the forms of inheritance examined, and the method by which the total number of optimization events was held constant.

NK fitness landscape We follow previous work in modeling evolution and learning in tunably rugged NK-landscape first developed by Kauffman and Levin (1987). Each landscape is determined by the number of dimensions (N) and the number of epistatic interactions for each dimension (K). Each dimension corresponds to a single locus in the genotype and contributes some amount to the overall fitness. The contribution of a specific dimension, however, also depends on the values at K other dimensions. In this way, the parameter K determines the "smoothness" of the fitness landscape. The simplest problem space, $K = 0$, contains a single global optimum, but as K increases, the problem space becomes more rugged and difficult. When K is at its highest possible ($K = N - 1$), the problem is effectively random. For the simulations presented below, we fixed the dimensionality ($N = 15$) and varied systemically the number of epistatic interactions across the full range (K between 0 and 14). For each problem space, scores were normalized to run between 0 and 1, with 0 corresponding to the worst possible solution and 1 corresponding to the best solution as determined by an exhaustive search of the landscape. Following previous work (Lazer and Friedman, 2007; Barkoczi and Galesic, 2016), we elevated the scores to the power of 8 to widen the distribution of solutions. Due to the variability of different instantiations of each NK landscape, each condition was tested on the same set of 5,000 landscapes and starting conditions.

Lifetime learning We consider two distinct types of learning: *lifetime learning events* at the level of the individual and *evolutionary learning events* at the level of the population. During a lifetime learning event, each individual in the population independently attempts to improve its fitness by examining its local area in the NK landscape and altering its phenotype, with no consideration of the rest of the population. We primarily examine two types of lifetime learning: *stochastic hill-climbing* and *steepest hill-climbing*. For stochastic hill-climbing, an individual flips a random bit of its phenotype (i.e. takes a single "step" in the landscape). If the resulting phenotype corresponds to a higher fitness than the current fitness of the individual, then the change is kept. Otherwise, the original phenotype is retained. For steepest hill-climbing, an individual instead examines all N single bit-flips of the phenotype and keeps the change that results in the highest fitness (or the original phenotype, if all changes lead to worse fitness). We also briefly analyze intermediate hill-climbing approaches in which an individual examines some number of single-bit flips of the phenotype between one (purely stochastic) and N (full steepest-ascent). In an intermediate case, each individual randomly selects some number of random genes, observes the fitness of each single bit-flip, and either keeps the change that gives the highest fitness, or its original phenotype if none are better.

Evolutionary learning During an evolutionary learning event, a new population of individuals is generated from the current population through the typical genetic algorithm processes of selection, recombination, and mutation. First, each individual is ranked according to their phenotypic fitness (we note that in conditions with no learning events, the phenotype is simply the same as the genotype). A portion of the highest-fitness individuals referred to as the *elite* (top 10%) have their genotypes passed directly to the new population. The remainder of the new individuals are created by selecting two "parents" from the population based on their fitness rank. The two parent genotypes are randomly combined to form the new genotype. Finally, each locus of the new genotype individually undergoes a bit-flip mutation with a probability of $1/N = 0.0667$. Each phenotype for the new generation of individuals is set equal to the new genotype.

Inheritance We distinguish here between *Darwinian* (in this sense meaning that non-genetic traits are not inherited) and *Lamarckian* inheritance. The Darwinian case proceeds exactly described above – during learning events individuals make alterations to their phenotypes, while reproduction events operate over the underlying and unchanging genotypes. In the Lamarckian case, however, changes made to the phenotype during learning events are actually reflected in the individual's genotype and go on to affect the behavior of future reproduction events directly. While the Lamarckian condition has no direct parallels in biology, the combination of individual and distributed search methods is still of some algorithmic interest.

Learning event scheduling In all experiments, we use a simple scheduling regime to holding the total number of events constant (e=100) while varying the proportion of learning and evolution. At one extreme, we have a population that only learns, and hence has 100 learning events. At the other, we have a population that only evolves, and hence has 100 reproduction events. In between, there are many possibilities that mix learning and evolution to different degrees by varying the number of learning events preceding each reproduction event. For instance, in the $4:20$ schedule there are 4 learning events preceding every reproduction event, leading to 20 reproduction events. Altogether, the following schedules were evaluated: 0:100, 1:50, 4:20, 9:10, 19:5, 49:2, 100:0. Importantly, in every schedule the total number of learning events (e) remains constant, allowing comparisons across different schedules. It should be noted that the number of numerical operations does *not* necessarily remain constant across experimental trials (as stochastic individual learning involves fewer comparisons than steepest individual learning). While this might have an effect on analyses of algorithmic performance, our primary interests are more theoretical.

Results

While previous work has demonstrated the advantages of combining lifetime and evolutionary learning over either in isolation, we systemically explore the conditions in which these advantages manifest. Specifically, we begin by analyzing the performance of populations across different proportions of individual and evolutionary learning. Afterwards, we analyze in turn the effects of task difficulty, type of inheritance, and type of lifetime learning.

Performance across different combinations of lifetime and evolutionary learning

We begin by discussing results for the Darwinian inheritance condition with steepest-ascent hill-climbing, while varying the proportion of learning and evolution events. This is partly motivated by biology: in a biological setting, each evolutionary learning event corresponds to the birth of a new generation of individuals and so the number of lifetime learning events is roughly analogous to each individual's lifespan. Further, there is a tradeoff associated with the selection of a lifespan. A longer lifespan allows for individuals to better learn to exploit their environments, while a shorter lifespan allows for a greater number of novel genetic strategies to be explored within a fixed period of time. In Figure 1, we see the best fitness of the population over time for different proportions of learning and evolution and a fixed task difficulty (K=6). In all settings, performance tends to increase over time and nears optimal performance after 100 events. In hybrid settings involving both evolutionary and lifetime learning events, we observe gradual increases in fitness punctuated by sharp decreases with every evolutionary learning event. This is unsurprising, given that evolutionary learning events do not record any of the phenotypic improvements made during prior lifetime learning events. Nevertheless, fitness quickly recovers from each of these dips. Indeed, the two extreme trajectories of exclusively lifetime learning (the orange line) and of exclusively evolutionary learning (the purple line), while smooth, do not achieve the optimal performance overall. This indicates that hybrid approaches (or intermediate lifespans) allow individuals to better reach the optimal areas of the fitness landscape.

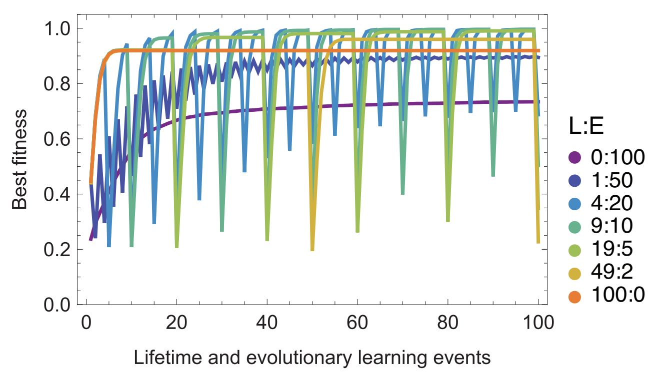

Figure 1: Performance across different proportions of lifetime and evolutionary learning, as function of total learning events (held constant). Different proportions shown in different colors (e.g. pure evolution in purple, pure learning in orange, hybrid in blue). All populations used Darwinian inheritance, steepest hill-climbing, and a landscape with $K = 6$.

Performance across task difficulty

Does the hybrid approach combining lifetime and evolutionary learning always outperform the approaches that consider only one form of learning? Are there some conditions for which certain combinations of lifetime and evolutionary learning work better than other ones? We address these questions in the context of task difficulty by examining the performance of the different conditions analyzed in the previous section across fitness landscapes with different levels of epistatic interactions ("ruggedness", or K), and present results in Figure 2. We observe that the ruggedness of the landscape has a significant effect on the population's performance. In simple environments with small K values, all strategies perform equally well. However, as the landscape becomes more rugged, we observe a greater disparity in performance between the extreme strategies and the optimal hybrid strategy. Indeed, the difference in performance between extreme strategies and hybrid strategies becomes increasingly pronounced as the task difficulty increases. This indicates that such hybrid strategies may be most helpful in complex environments. In the case presented in Figure 2, the optimal proportion of lifetime learning events to evolutionary learning events was 4:20, and the optimal proportion does not seem to vary as a function of task difficulty.

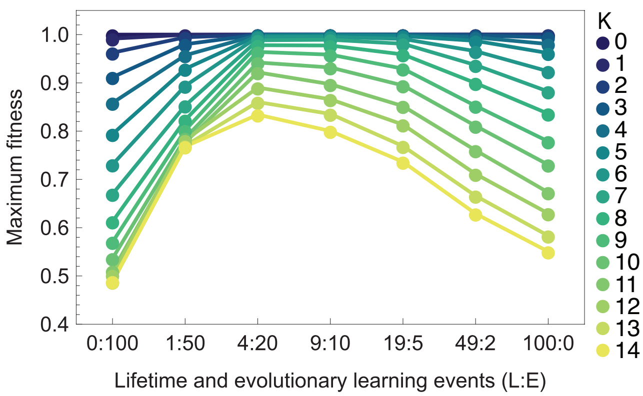

Figure 2: Performance across task difficulty. Maximum fitness reached by an individual in the population as function of different proportions of lifetime and evolutionary learning. Different task difficulties (K) shown as different colored traces. Darwinian inheritance. Steepest gradient ascent individual learning

Effect of inheritance on learning

Next we examine the effect of Lamarckian inheritance, in which individuals change their genotypes directly through learning. The motivations for this are twofold, one biological and one algorithmic. First, while genes do not acquire characteristics during an individuals lifetime, there are other mechanisms of inheritance that allow for information acquired during a lifetime to be passed to offspring, such as social learning or external information encoded in the environment. Second, the performance of hybrid strategies is also of interest to the field of AI, in which the choice of which characteristics to have inherited by members of a population is an open one. In Figure 3 we see that regardless of the distribution between learning and reproduction events, the best overall fitness steadily increases over time. This follows from the fact that individuals are not forced to start learning over again with each evolutionary learning event – they get to “keep” some of the experience they gained through individual learning. Overall, we find that while pure learning slightly outperforms pure evolution, hybrid strategies still often outperform both extremes, as with the Darwinian Inheritance condition. This offers some indication that algorithmic optimization strategies may benefit from combining localized and distributed search methods.

Interestingly, the absolute fitness scores for individuals in the Darwinian and Lamarckian inheritance conditions are

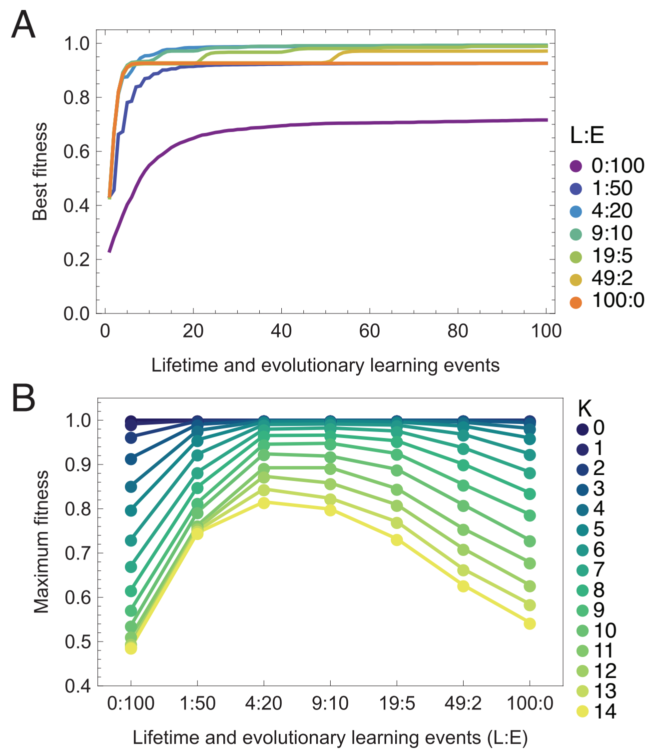

Figure 3: Effect of inheritance on learning. **Panel A**: Performance across different proportions of lifetime and evolutionary learning, as function of total learning events (held constant). Different proportions shown in different colors. Intermediate task difficulty ($K = 6$). **Panel B**: maximum fitness reached by an individual in the population as function of different proportions of lifetime and evolutionary learning. Different task difficulties (K) shown in different colors. **Both:** Lamarckian inheritance. Steepest-ascent hill-climbing individual learning

roughly equivalent. This is somewhat surprising, given that individuals in the Lamarckian condition are able to accumulate learning improvements through multiple generations. One possible explanation is that individuals in the Darwinian inheritance condition have a selective pressure to be near areas of steep incline in the landscape, which would allow them to reach a high fitness with relatively few individual learning events and make up for the apparent advantage gained by individuals in the Lamarckian inheritance condition.

Effect of the type of lifetime learning

Finally, we examine the effect of different types of individual learning events on population performance. While in many settings it is appropriate to view individual learning as a process of identifying directions of steepest gradient ascent or descent, in others a more stochastic model is a better fit. For instance, while many modern optimization algorithms move the weights in the exact direction of the computed gradient, the gradient is often calculated over a subset of all data, leading to a stochastic estimate. In Figure 4, we show the best fitness achieved by the population for different schedules and task difficulties (as in Figure 2) using stochastic hill-climbing for individual learning events. In the Darwinian inheritance condition, we see that hybrid strategies no longer outperform extreme strategies for any settings of K. Instead, there is a slight tendency for more learning to perform better. However, hybrid strategies *do* still outperform mixed strategies in the Lamarckian inheritance condition, though the difference is less than for steepest-ascent hill-climbing. What accounts for this difference? One possibility is that the ability for individuals in the Darwinian inheritance condition to take advantage of areas in the landscape of steep ascent is negated when individuals must improve themselves stochastically and cannot reliably exploit these directions. The individuals in the Lamarckian inheritance condition are potentially less affected by stochastic hill-climbing because they rely less on identifying such areas of steep ascent.

To further explore the effect of different types of lifetime learning, we systemically varied the number of genes examined during lifetime learning between the stochastic case and the steepest case for a single task difficulty (K=6), and present the results in Figure 5. We see that in the Darwinian inheritance condition, increasing the number of genes considered caused hybrid strategies to outperform extreme strategies, with more genes considered corresponding to a greater difference in performance. Effectively, we see the smooth emergence of the "inverted-U" shape that characterizes our earlier experiments. By contrast, in the Lamarckian condition, we observe that hybrid strategies outperform their extreme counterparts regardless of the number of genes considered during lifetime learning, though the number of genes considered does correlate with the performance difference. In both conditions, the overall best performance unsurprisingly increases with the number of genes considered.

Conclusion

We have shown that for simple NK-fitness landscapes, the best performance is not achieved by either individual hill-climbing or distributed evolutionary selection, but rather by a mixture of the two. Further, the benefit of these hybrid strategies becomes more pronounced as the landscape becomes more difficult. In the Darwinian inheritance condition, we also observe that the type of lifetime learning can have a dramatic impact on its interaction with evolutionary learning. The results from the Darwinian inheritance condition offer some insight into the mechanisms of optimization underlying biological agents, while the Lamarckian inheritance condition offers some potential motivation for novel algorithms in artificial intelligence and machine learning.

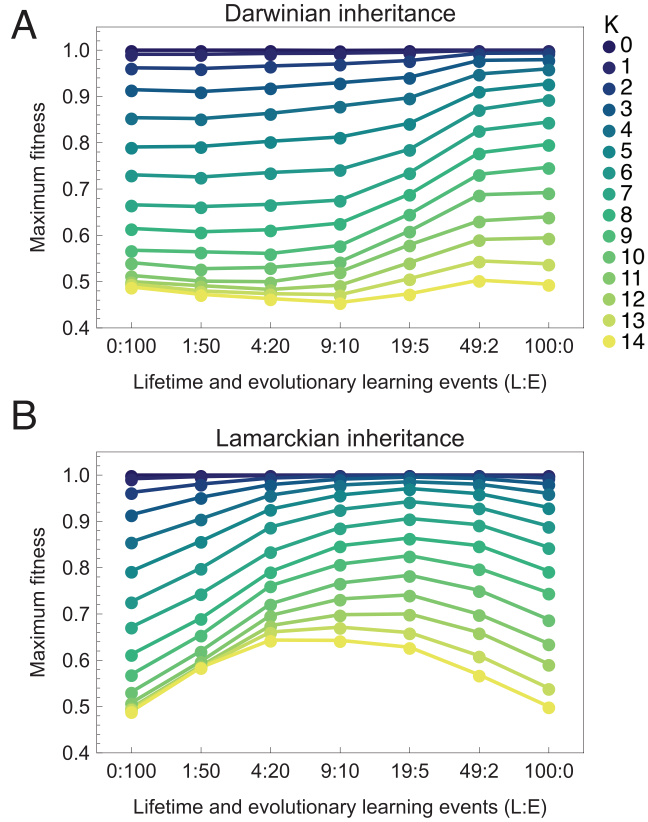

Figure 4: Effect of the type of lifetime learning. **Panel A**: maximum fitness reached by an individual in the population as a function of different proportions of lifetime and evolutionary learning. Different task difficulties (K) shown as different colored traces. Darwinian inheritance. Stochastic hill-climbing individual learning. **Panel B**: the same, but for Lamarckian inheritance.

A question of primary importance concerning our results is *why?* That is, why do hybrid strategies perform better than pure learning or pure evolution? There are many potential explanations, but perhaps a reasonable one is that hybrid strategies do a better job of balancing exploitation and exploration in search. As we modeled it, lifetime learning involves taking advantage of local improvements in fitness, and spends no time exploring potentially suboptimal changes. Evolutionary learning, which operates primarily through mutation and recombination, involves substantially more exploration, with only the selection step acting to exploit advantages. As has been observed in many contexts, neither pure exploitation nor exploration achieves optimal results. Thus, it may be the case that hybrid strategies serve to offer a more balanced search strategy that is ultimately more successful, especially as task difficulty increases. Future work could involve validating this hypothesis by measuring how often each strategy performs exploitative or exploratory updates.

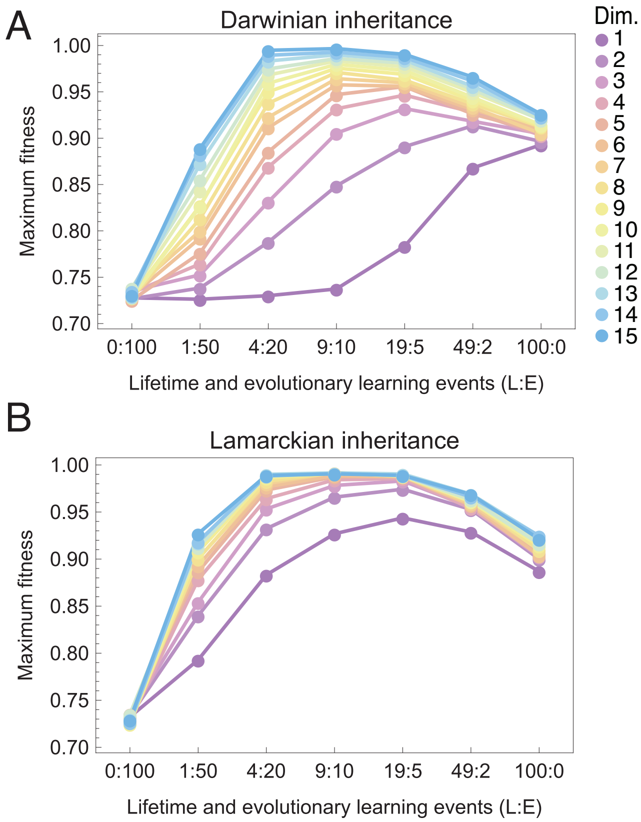

Figure 5: Further effect of type of lifetime learning. **Panel A**: maximum fitness reached by an individual in the population as a function of different proportions of lifetime and evolutionary learning. Intermediate task difficulty (K=6). Darwinian inheritance. Different number of genes considered during lifetime learning shown as different colored traces. **Panel B**: the same, but for Lamarckian inheritance.

Another promising area of future research is to investigate further modifications of other relevant search parameters, such as the number of parents involved in genetic reproduction, the mutation rate, and whether individual learning events are able to "explore" or perform non-optimal updates. Another reasonable extension is the exploration of more complex fitness landscapes through the use of more powerful algorithms and models. For instance, deep neural networks could be used to investigate the trade-off between individual learning and evolutionary learning for simulated walking agents. Another path is to investigate whether hybrid strategies offer benefits over pure strategies for multi-task learning and optimization. In either case, it will be worthwhile to explore both the biologically-motivated Darwinian inheritance condition and the algorithmically-motivated Lamarckian inheritance condition.

Data availability

The simulation code and data files are publicly available in our research group's GitHub account: github.iu.edu/EASy/ToddALife2020.

Acknowledgements

We are thankful for the reviewers' comments. This material is based upon work supported by the National Science Foundation under Grant No. 1845322.

References

Ackley, D. and Littman, M. (1991). Interactions between learning and evolution. *Artificial life II*, 10:487–509.

Baldwin, J. M. (1896). A new factor in evolution. *The american naturalist*, 30(354):441–451.

Barkoczi, D. and Galesic, M. (2016). Social learning strategies modify the effect of network structure on group performance. *Nat. Commun.*, 7:13109.

Belew, R. K. (1990). Evolution, learning, and culture: Computational metaphors for adaptive algorithms. *Complex Systems*, 4(1):11–49.

Curran, D., O'Riordan, C., and Sorensen, H. (2007a). An analysis of the effects of lifetime learning on population fitness and diversity in an nk fitness landscape. In *European Conference on Artificial Life*, pages 273–283. Springer.

Curran, D., O'Riordan, C., and Sorensen, H. (2007b). Evolving cultural learning parameters in an nk fitness landscape. In *European Conference on Artificial Life*, pages 304–314. Springer.

Fernando, C., Banarse, D., Blundell, C., Zwols, Y., Ha, D., Rusu, A. A., Pritzel, A., and Wierstra, D. (2017). Pathnet: Evolution channels gradient descent in super neural networks. *arXiv preprint arXiv:1701.08734*.

Gruau, F. and Whitley, D. (1993). Adding learning to the cellular development of neural networks: Evolution and the baldwin effect. *Evolutionary computation*, 1(3):213–233.

Hinton, G. E. and Nowlan, S. J. (1987). How learning can guide evolution. *Complex systems*, 1(3):495–502.

Houthooft, R., Chen, Y., Isola, P., Stadie, B., Wolski, F., Ho, O. J., and Abbeel, P. (2018). Evolved policy gradients. In *Advances in Neural Information Processing Systems*, pages 5400–5409.

Husbands, P., Harvey, I., Cliff, D., and Miller, G. (1997). Artificial evolution: A new path for artificial intelligence? *Brain and cognition*, 34(1):130–159.

Jablonka, E., Lamb, M. J., and Avital, E. (1998). 'lamarckian'mechanisms in darwinian evolution. *Trends in Ecology & Evolution*, 13(5):206–210.

Kauffman, S. and Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology*, 128(1):11–45.

Khadka, S. and Tumer, K. (2018). Evolutionary reinforcement learning. *arXiv preprint arXiv:1805.07917*.

Lazer, D. and Friedman, A. (2007). The network structure of exploration and exploitation. *Adm. Sci. Q.*, 52(4):667–694.

Leite, A. J., Candadai, M., and Izquierdo, E. J. (2020). Reinforcement learning beyond the bellman equation: Exploring critic objectives using evolution.

Littman, M. (1996). Simulations combining evolution and learning. *Adaptive Individuals in Evolving Populations: Models and Algorithms*, 26:465–477.

Mitchell, T. M. et al. (1997). Machine learning.

Nolfi, S. and Parisi, D. (1996). Learning to adapt to changing environments in evolving neural networks. *Adaptive behavior*, 5(1):75–98.

Pourchot, A. and Sigaud, O. (2018). Cem-rl: Combining evolutionary and gradient-based methods for policy search. *arXiv preprint arXiv:1810.01222*.

Sendhoff, B. and Kreutz, M. (1999). A model for the dynamic interaction between evolution and learning. *Neural Processing Letters*, 10(3):181–193.

Stanley, K. O., Clune, J., Lehman, J., and Miikkulainen, R. (2019). Designing neural networks through neuroevolution. *Nature Machine Intelligence*, 1(1):24–35.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Suzuki, R. and Arita, T. (2000). Interaction between evolution and learning in a population of globally or locally interacting agents. In *Proceedings of the Seventh International Conference on Neural Information Processing*, volume 738, page 743.

Suzuki, R. and Arita, T. (2004). Interactions between learning and evolution:: The outstanding strategy generated by the baldwin effect. *Biosystems*, 77(1-3):57–71.

Todd, P. M. and Miller, G. F. (1991). Exploring adaptive agency ii: Simulating the evolution of associative learning. In *Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats*, pages 306–315.

Whitley, D., Gordon, V. S., and Mathias, K. (1994). Lamarckian evolution, the baldwin effect and function optimization. In *International Conference on Parallel Problem Solving from Nature*, pages 5–15. Springer.