Evolutionary Consequences of Learning Strategies in a Dynamic Rugged Landscape

Nam Le, Michael O'Neill, Anthony Brabazon*

ABSTRACT

Learning has been shown to be beneficial to an evolutionary process through the *Baldwin Effect*. Moreover, learning can be classified into two categories: *asocial learning*, e.g. trial-and-error; and *social learning*, e.g. imitation learning. A *learning strategy*, or *learning rule* – a combination of individual and social learning – has been suggested by recent research can be more adaptive than both social and individual learning alone. However, this also leaves open an important question as to how best to combine these forms of learning in different environments. This paper investigates this question under a dynamic rugged landscape (i.e. dynamic NK-landscape). Experimental results show that a learning strategy is able to promote an evolving population better, resulting in higher average fitness, over a series of changing environments than asocial learning alone. The results also show that the population of strategic learners maintains a higher proportion of plasticity – the ability to change the phenotype in response to environmental challenges – than the population of individual learners alone.

KEYWORDS

Baldwin Effect, Phenotypic Plasticity, Evolutionary Algorithms, Dynamic Environment, NK-Landscape

ACM Reference Format:

Nam Le, Michael O'Neill, Anthony Brabazon. 2019. Evolutionary Consequences of Learning Strategies in a Dynamic Rugged Landscape. In *Proceedings of the Genetic and Evolutionary Computation Conference 2019 (GECCO '19).* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3321707.3321741

1 INTRODUCTION

Learning is an important characteristic to enable an organism to adapt its phenotype to the local environment during its own lifetime, and prepare an individual organism for future circumstances. By expressing behavioural plasticity, learning can play a more significant role when the environment becomes dynamic...and so dynamic that the slower evolutionary process cannot encode enough environmental information required for the survival of the species.

Interestingly, evolution and learning can complement each other through the phenomenon called the Baldwin Effect [2, 6, 12, 17], which was first demonstrated computationally by Hinton and Nowlan (henceforth H&N) [7]. Following this success, there have been quite a few important studies on the interaction between learning and evolution, notably in Artificial Life [1], Evolutionary Robotics [21], Evolving Autonomous Artificial General Intelligence [12], and in the NK-Landscape [3, 19]. Apart from some recent papers [13, 17], the influence of learning on evolution has subsequently been little investigated in the field of Evolutionary Computation (EC), despite the fact that many problem domains addressed by EC are inherently dynamic.

Generally, lifetime learning can be classified into two types, namely *Asocial* (or individual) learning (IL) – learning by oneself through direct interaction with the environment, e.g., trial-and-error, and *social* learning (SL) – learning from others, e.g., imitation. SL is considered a form of 'information-parasitism' since an individual can only learn socially from information produced by others. A key open question arising here is how to use SL more effectively. Too much SL can reduce the performance of the whole evolving population [15, 23]. It has been suggested that a learning strategy – a combination of both IL and SL – can show more adaptive evolutionary consequences than either form of learning alone [15, 17, 23].

This paper focuses on studying the effect that different forms of learning strategies may have on the evolutionary process when dealing with dynamic environments. We propose a form of dynamic environment based on the tunably rugged fitness landscape called the NK-landscape [9]. We also propose an algorithm combining evolution with learning strategies to test how the combination of social and asocial learning performs. In the remainder of this paper, we present some background concepts relating to social learning, then some prior research on learning and evolution. We also remind the reader of some the basics of the NK-landscape in the background section. We then present our experimental design, including our proposed dynamic problem and algorithms. Finally, some future directions are indicated.

2 BACKGROUND

2.1 Social Learning

Social learning can be understood as learning that is influenced by the observation of, or the interaction with another organism or its products [5, 8]. Social learning research could be considered a rapidly growing subfield of animal behaviour studies as well as human cultural evolution [22]. Social learning has been observed in organisms as diverse as primates, birds, fruit flies, and especially humans [23]. Although the use of social learning is widespread, understanding when and how individuals learn from others is a significant challenge [23]. Generally, social learning mechanisms include imitation, stimulus and local enhancement, observational conditioning; and amongst them imitation learning has been said a unique ability of highly intelligent animals, including humans [8, 22]. In imitation learning, the observer directly copies the behavior of the observed animal in order to complete a novel task.

*Corresponding Author: Nam Le is with Natural Computing Research & Applications Group, University College Dublin, Dublin, A94-XF34, Ireland.
Email: namlehai90@gmail.com. Tel: +353 83 320 1501.

GECCO '19, July 13–17, 2019, Prague, Czech Republic

ACM ISBN 978-1-4503-6111-8/19/07…$15.00
https://doi.org/10.1145/3321707.3321741

SL, at first glance, seems to be a low-cost, adaptive, mechanism as individual agents can acquire information from others without incurring the cost of trial-and-error learning. Thus, it is plausible to think that SL will result in more effective learning outcomes. Contrary to this belief, it has been found that agents should not learn socially all the time [11, 15, 25]. It is argued that individual learners produce new information about the environment, though at a cost. In contrast, social learners avoid this cost by copying the existing behaviors of others, but do not themselves generate any new information about the environment. Therefore, it is highly likely that social learners will copy outdated information when the environment changes, reducing the average fitness of the population.

Several theoretical models have been proposed to investigate how to use SL effectively [4, 11, 26]. It is said that social learning should be combined with individual learning in a strategic way in order to produce an adaptive advantage. Social learning strategies consist of rules specifying the way an individual relies on social learning by answering three questions:

i. When an individual should learn?
ii. From whom they should learn?
iii. What information should be learned?

The question of *when to copy* covers the decision as to when to seek social information. *Whom to copy* may depend on factors such as the social structure of the population and the ability of the individual to recognise whether other individuals are obtaining higher payoffs. Possibilities include the copying of the most successful individual, copying of kin, or adherence to a social norm by copying the majority. *What to copy* considers which behavior or more specifically what part of that behavior to copy.

There can be three modes of information transmission through social learning [22, 24]. The first is *vertical transmission* - transmission from parents to their children. The second is *oblique transmission* in which cultural traits will be passed to an individual from another individual learner, not its parent, but from the previous generation. The last is horizontal transmission - an observer will learn from a demonstrator in its current generation. In the scope of this paper, we design an algorithm modeling social learning through oblique transmission, which will be described later.

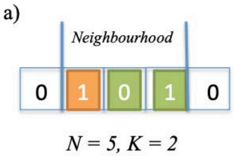
2.2 Learning and Evolution

In 1987, the Cognitive Scientist Geoffrey Hinton and his colleague Kevin Nowlan at CMU presented a classic paper [7] to demonstrate an instance of the Baldwin effect in the computer. It was shown that learning in the form of random search can speed up and *guide* evolution to solve a *haystack* problem.

The model developed by Hinton and Nowlan, though simple, is interesting, opening up the trend followed by a number of studies investigating the interaction between learning and evolution, including [1], [21], [19], [12], [16].

However, most research papers on this topic only study fixed environment, and the relationship between asocial learning and the evolutionary process. We think that it is first worth investigating more about the effect learning might have on evolution in dealing with dynamic environments. Second, we should take social

a)



b)

Lookup Table

Comp	Fitness
000	0.22
001	0.32
010	0.56
011	0.25
100	0.35
101	0.86
110	0.64
111	**0.21**

Figure 1: An example of an NK Model when N = 5 and K = 2. (a) shows how the fitness contribution of each bit depends on K consecutive bits. Therefore there are 2^{k+1} possible allele combinations, each of which is assigned a random fitness in a lookup table as shown in (b). Each bit has such a table created for it. Total fitness of a given sequence is the normalized sum of these values.

learning [8, 16] into consideration, and see how social learning can contribute to the effect that learning may have on evolution.

We give a brief introduction of the NK-model as well as some notable research on learning and evolution in the NK-model in the following section.

2.3 The NK-Landscape

NK-landscape is a tunably rugged fitness landscape which is often studied in theoretical biology and evolutionary computation theory research [9, 10]. In an NK-landscape, there are two landscape parameters, namely N and K. N is the length of each string (genotype) in the landscape. For each string of length N, the fitness of the whole string is the sum of the fitness of each bit (locus) in the string. The fitness of each bit (locus) is called the fitness component for the whole string (genome). Parameter K specifies for one specific bit (locus), how many bits (loci) in its neighborhood contribute to its fitness. These K bits can be chosen randomly, or can be consecutive K bits (in circular fashion). Therefore, this forms the interaction between bits (loci) in a string (genome), creating the ruggedness of the landscape [10]. This is why NK-landscape is called *tunably rugged* because the ruggedness of the landscape can be tuned by changing the interaction graph through parameter K. Figure 1 shows an example of a NK model. Please refer to [9, 10] for more details of the NK Landscape.

The NK-landscape has also been used to investigate the Baldwin Effect, concerning the interaction between evolution and asocial learning. Several notable studies of the Baldwin effect in the NK-model include works by Giles Mayley [19, 20], and some others [3]. Their results, again, demonstrated that the Baldwin effect does occur and the allowance for lifetime learning, in the form of individual learning, helps evolutionary search overcome the difficulty of a rugged fitness landscape.

As mentioned earlier, most research on evolution and learning in NK-landscape are concerned with a fixed environment: There is no variability in the structure and the difficulty of the landscape. Therefore, in this paper we explore the effect that learning strategies might have on evolution when dealing with dynamic NK-landscape. We also combine both asocial and social learning, to see how that interaction can affect the evolutionary process. Experimental design is presented in the following section.

3 EXPERIMENTAL DESIGN

3.1 The Dynamic NK-Landscape

We propose a dynamic version of the NK-model by varying the parameter K – the ruggedness of the landscape over generations. There is a parameter called frequency of change, which controls after how many generations the landscape will change. We initialise a list of Ks and let the algorithm choose K based on some frequency of change. K can be chosen randomly from the list, yet in our experiment we choose K based on a deterministic procedure. We slide through the list of Ks to set the K for the landscape. We allow circular sliding through the list of K, which means that if we reach the last K of the list, the list will start again from the first K. For example, assume the list of K is 0, 2, 4 and initially K is 0. If the frequency of change is 10, at generation 10 K will be 2, at generation 20 K will be 4, at generation 30 K will come back to 0 again, and so on.

By varying K like before, we are imposing some dynamics on the NK-landscape, with increasing and/or decreasing the ruggedness of the landscape over time. The optimal value of the landscape will be changed when the K is changed.

3.2 Experimental Setup

To explore the impact that a learning strategy might have on an evolving population in a dynamic environment we adopt four experimental setups. The first two setups are controls. The first control adopts an evolving population in the absence of any learning (i.e., no asocial or social learning process is present), and the second control is identical to the first control with the addition that an oracle triggers a restart from an randomised population at each point when the environment changes. The third and fourth setups combine evolution with asocial learning, and evolution with a learning strategy respectively. Further details are provided below.

3.2.1 *Experimental Setup I: Evolution alone.*

The first setup evolves a population of individuals without any form of learning. This is the canonical Genetic Algorithms. Every individual has the genome of 20 bits randomly initialised, with only two alleles 0s and 1s. The genotype-phenotype mapping is one-to-one. The fitness of an individual is calculated as the fitness value of the corresponding bit-string in the NK-landscape. At each generation, two individuals are selected, based on fitness-proportionate selection scheme, to produce one child by one-point crossover. Mutation is not used in our experiment and there is no *elitism* permitted.

3.2.2 *Experimental Setup II: Evolution + Restart strategy.*

The second setup only differs from the first setup in that we restart the population initialisation at the generation of change. Though this is an *handcrafted* strategy, we implement this for a special purpose as follows: As the environment changes, the evolutionary process is supposed to lose information about the environment, and may not be able to proceed. We would like to see how evolution with and without restart proceeds through dynamic environments, in order to see if the evolutionary process itself embeds useful information, to ascertain that if learning in combination with evolution adds performance value over evolution process on its own.

The first two experimental setups are served as controls in our experiment. The later two presented below will be concerned with the main hypothesis of this paper.

3.2.3 *Experimental Setup III: Evolution + Asocial Learning.*

In the third setup, we evolve a population of learning individuals. In this simulation, we also allow lifetime learning in the form of asocial learning, in addition to evolutionary algorithm, to update the phenotype of the individual. To allow for lifetime learning we used the same encoding scheme as in [7, 15, 17]: Instead of being fully specified, the genotype now is composed of three alleles '0', '1', and '?'. The allele '?' allows for lifetime learning (or plasticity). Each agent will have 100 rounds of learning during its lifetime. On each round, an individual agent is allowed to do individual learning by changing its allele '?' to either '0' or '1' as the expressed value. Thus, the behaviour of an individual agent is partly specified by its genetic composition, and partly by what it learns in the course of its lifetime.

The evolutionary process is implemented as an evolutionary algorithm similar to the previous experimental setup. At each generation, two individuals are selected from the population as parents to produce one child. The newly-born child is then added into a new population. This process repeats until the new population is completed upon which it replaces the old population of parents, i.e., the process consists of generational replacement without elitism.

When an individual learns, it updates its phenotypic behaviour, and hence its fitness. However, instead of being implemented as a blind random search as in previous work [3, 17], we devise a new learning algorithm as a *hill-climbing* process. The learning algorithm adopted by every individual is presented as Algorithm 1 below.

The above algorithm is relatively self-explanatory. When an individual expresses a new phenotypic behaviour, it checks whether the new behaviour is more adaptive than the current before deciding whether the current phenotype is replaced by the new one. This process helps each agent keep its best behaviour as its current phenotype. Fitness of a phenotype is computed by assigning the fitness value of the corresponding point (a bit-string) in the NK-Landscape.

After lifetime learning, the population goes through the evolutionary process as follows. At each generation, two individuals are selected from the population as parents to produce one child. This process repeats until the new population is filled up and replaces the

Algorithm 1 Learning

```
function LEARNING(ind)
    best_fitness = ind.fitness
    best_phenotype = ind.phenotype
    while ind.learning_attempts < max_attempts do
        ind.learning_attempts+ = 1
        Flip all question marks to get a new phenotype
        best_fitness = compute_fitness(ind.phenotype)
        if best_fitness > ind.fitness then
            ind.fitness = best_fitness
            best_phenotype = ind.phenotype
        end if
    end while
    ind.phenotype = best_phenotype
end function
```

old population of parents. No mutation is employed in the current work as in previous studies [15, 17].

3.2.4 Experimental Setup IV: Evolution + Learning Strategy.

In the last setup, we evolve populations of strategic individuals - individuals that can perform both SL and IL based on a learning rule. In order to implement social learning, first we propose the imitation procedure, with pseudo-code described in algorithm 2 below. This presents the process by which an individual observer imitates the phenotype of its demonstrator. The imitative process starts by extracting the positions of question marks in the phenotype of the observer. For each question mark position, the observer will copy exactly the *trait* from the demonstrator.

Algorithm 2 IMITATION

```
function IMITATION(observer, demon)
    questions = [] comment: question mark array
    for position i ∈ observer.pheno do
        if i =? then
            questions.add(i)
            observer.learning_attempt += 1
        end if
    end for
    for i ∈ questions do
        observer.pheno(i) = demon.pheno(i)
    end for
end function
```

The population now has just one type of individual - strategic individuals that can learn both asocial and socially. We specify the learning strategy for every individual agent as follows: At each generation, an agent first *observes* and learns socially from its demonstrator, then learns asocially on its own until the maximum learning attempt is reached. This scenario can be interpreted as a novice first copies from an expert, then sharpens what it has learned to further its own quality. In the scope of this study, we adopt *oblique transmission* – the individual agent learns from one individual in the previous generation. This also means that there is no social learning at the initial generation. In our current study, all the individuals at each generation have the same demonstrator – the best individual agent in terms of fitness belonging to the previous generation. After the lifetime learning process for each agent, the population goes through selection and reproduction as in EVO+IL.

We use 5 values for K with increasing ruggedness (0, 2, 5, 10, 19). Starting from K = 0 (no ruggedness), each value of K is consecutively chosen after a few generations specified by the frequency of change parameter. We run our experiments through 4 different *frequency* of change. It can be understood that the lower the *frequency* value, the faster the target will change. The environment becomes more dynamic by faster changing, and vice versa. All the parameter settings used in our experiments are described in Table 1 below.

Table 1: Parameter setting

Parameter	Value
N	20
K	0, 2, 5, 10, 19
Genome length	20
Replacement	Generational
Generations	200
Elitism	No
Population size	100
Selection	Fitness-Proportionate selection
Reproduction	Sexual reproduction
Fitness function	Computed in the NK-landscape
Max learning attempts	100
Frequency	5, 10, 20, 40

Please note that we loop through the list of K over 200 generations, the ruggedness of the NK landscape is increased (more difficult) and then decreased (less difficult) depending on the frequency of change and how many times we slide the whole K list.

4 RESULTS, ANALYSIS, AND EXPLANATION

In this section we present comparative analysis of how each of the experimental setups above performs. All the results are averaged over 30 independent runs. The fitness values presented in all the figures below are normalised before analysing, measured by the real fitness of an agent (computed by the NK model) divided by the optimal fitness in the whole NK landscape. Importantly, it should be kindly noted that we **do not alter the original fitness value (computed by the NK model) during the whole course of evolution** in all setups. All individual agents undergo the evolutionary process by their original fitness (computed by the NK model). This fitness normalisation is only for figure presentation, to allow a fair comparative analysis across dynamic landscapes since the optimal target changes as we vary the K parameter. When the NK landscape changes, we have to reinitialise the landscape and the fitness in the lookup table. Thereby we compute the optimal fitness value by exhaustive search just for final statistics.

The main point of our work is to see how our learning strategy (in EVO+Strategy) promotes evolution compared to asocial learning (in EVO+IL). As controls we have also run evolution alone and

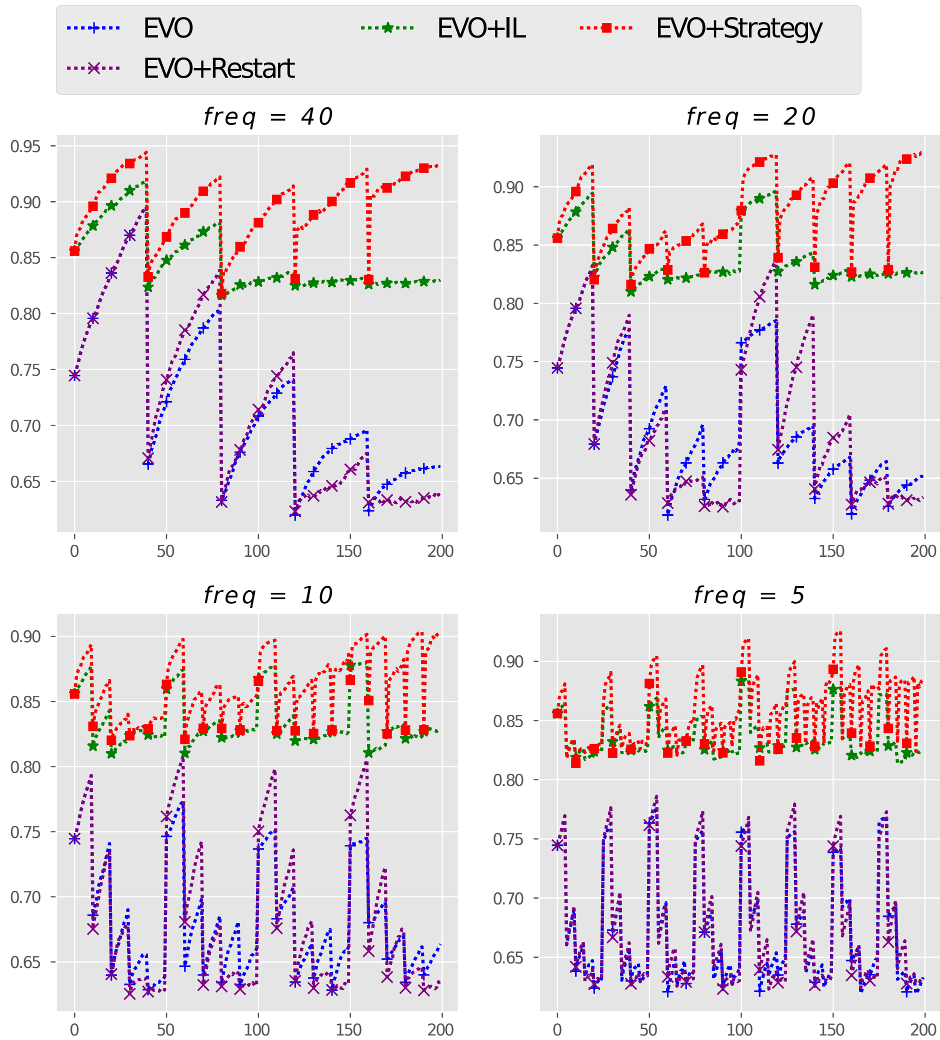


Figure 2: Average Fitness Comparison.

evolution + restart setups with the reason described above. We shortly present the comparison between the EVO, EVO+Restart and EVO+IL populations below.

It can be seen in Figure 2 and 3 that the EVO+IL population outperforms the controls [EVO and EVO+restart] in terms of both best and average fitness. In the restart control we observe that for changes to the less challenging landscapes (i.e. those with lower K) the restart shortly appears to gain a small advantage over the continuously evolving population. But as K increase this advantage disappears. In general, however, our controls behave as expected with both clearly outperformed by the setups that include a learning process.

Conversely, in EVO+IL individual agents are equipped with the ability to learn. It is this learning ability that helps gain more information about the changing landscape. This is why the EVO+IL outperforms both EVO and EVO+Restart populations. Learning is shown to present a clear effect on the foundational evolutionary process.

Focusing on the main point of this study, given the effect of learning on evolution as the NK landscape changes as we have seen, the key and interesting question here is what type of learning can promote an evolving population better, and if so how and why the effect arises?

Overall, it can be seen in Figure 2 and 3 that there is a shared behaviour in the two populations, that when the environment changes there is a drop in performance. Generally, the performance is higher when the ruggedness of the landscape is lower, and vice versa. This is understandable since at the generation of change, the information gained in the previous generations can become incorrect, reducing the fitness of the population. Moreover, when the landscape becomes more rugged (i.e. higher K), it is harder to find the way to the solution because of the increase in *epistasis*, hence the number of peaks in the landscape.

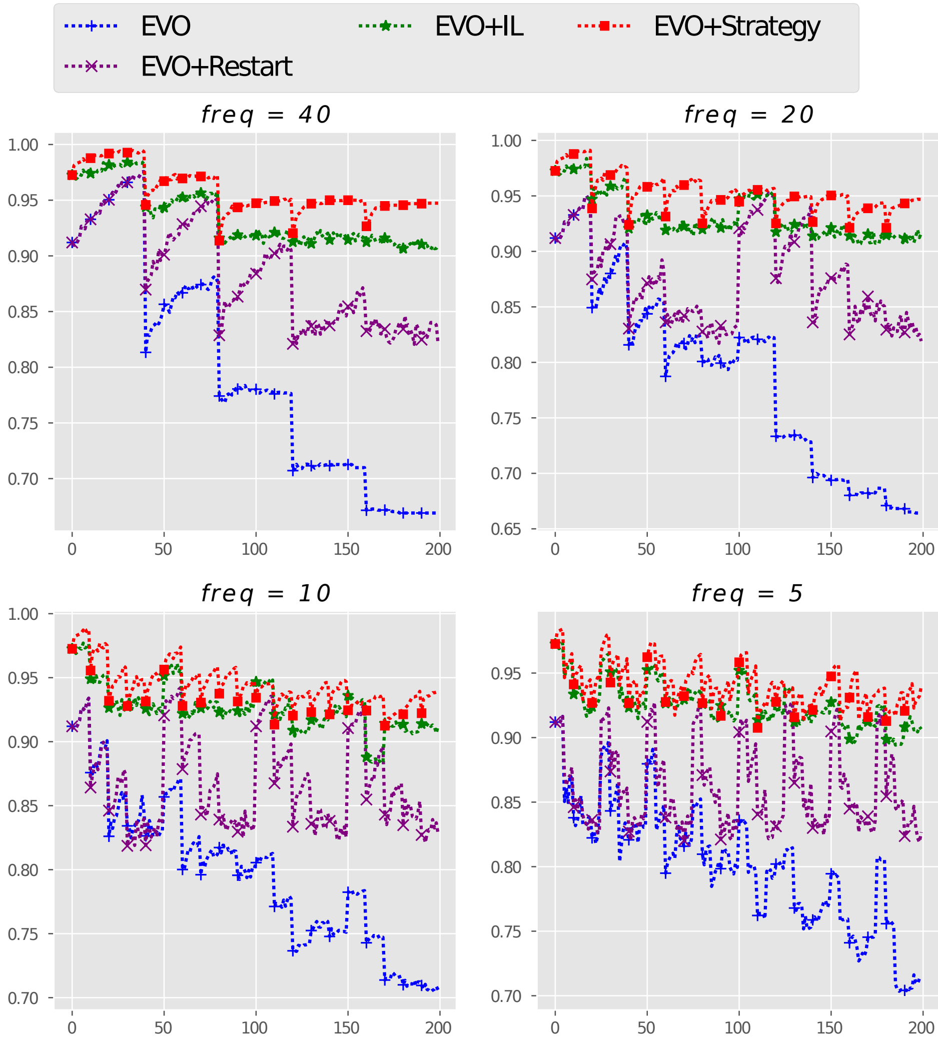


Figure 3: Best Fitness Comparison.

More specifically, it would be more interesting to see how asocial learning and a learning strategy can effect the whole evolving population. By looking at the average fitness over generations, we observe that EVO+Strategy outperforms EVO+IL in all settings and the difference between EVO+IL and EVO+Strategy becomes clearer over generations. Especially, even when the K is higher (higher epistasis), the EVO+Strategy demonstrates the ability to promote the evolving population more than EVO+IL.

Some more interesting questions that requires a deeper explanation arise here as to why EVO+Strategy shows a stronger effect when compared to EVO+IL? What is the theoretical mechanism that drives this? Remember that in EVO+Strategy, we have designed a strategy for each each individual agent that first learns socially, then learns individually on its own. Is that the impact of learning socially before learning individually on evolution that triggers the better ability to deal with the changing NK landscape?

All of these observations here can be explained by the fact that, the individual learning process still does its job – upgrading the current phenotype of an agent towards the target step-by-step through a hill-climbing process. The impact of individual learning on evolution has been shown before. It is reasonable here to think that if an individual agent has a better foundation for individual learning, the agent would get better behaviour at the end of the day.

Remember the nature of our learning strategy mentioned above. Each strategic agent, after birth, first imitates behaviour from the best individual in the preceding generation, then updates its behaviour itself through individual learning. This can be understood in casual language as follows: A *novice* copies a good enough skill from a professional, then makes its own effort to upgrade and sharpen that skill itself. For our problem, this also means that each learning agent is potential to have better phenotypic base before individual learning takes place. Thus, strategic agents – those that can learn from others and on their own - have more advantage over agents that can only learn individually.

One key idea to be noticed here is that agents in EVO+IL learn asocially based on the genetic basis provided by evolution – this is individual learning to update innate foundation. More interestingly, agents in EVO+Strategy learn individually to update what has been learned, or found, individually by other agents from previous generations. This creates a form of cumulative learning process – updating information and knowledge based on what has been found so far in the history.

Therefore, in earlier generations when the evolutionary process plus individual learning alone would need more time to find the best solution, copying from the best seems to give some initial advantage. This is because the imitation process provides a better base for individual learning, compared to individual learning from innate information only when the environmental information is not encoded enough. Over generations, however, when the environment slightly changes and becomes a little bit harder (i.e. a small increase in K), the landscape is not changed to a large degree, all the individual agents in EVO+IL still have a chance to move closer and closer to the target (the solution). There is no huge benefit to copy from any *expert* (the demonstrator). This is why adding social learning before asocial learning does not bring much benefit in these scenarios.

When the environment becomes more difficult to cope with (i.e. higher K), the landscape is changed at a sufficient level of ruggedness and the number of peaks, the individual learning process still updates the phenotypic information to search over the increasingly more complex environment. However, the evolutionary process looses more information over generations (as shown before). Therefore, learning from others, or more precisely, learning from what others have found previously by individual learning is more advantageous than learning from innate information on one's own. This is why EVO+Strategy shows better performance in terms of the average fitness in all cases.

The explanation presented above is still valid in explaining why the best fitness of EVO+Strategy is better than that of EVO+IL in all cases (as shown in Figure 3. Learning from others can provide a better substrate for individual learning to take place. Therefore strategic agents have more chance to move closer to the optimal point compared to agents that learn individually based on innate knowledge, hence the better best fitness (the fitness of the best individual found compared to the optimal in the landscape). More specifically, the EVO+Strategy still can get closer to the solution whereas the EVO+IL cannot when the environment becomes harder (the case when K is getting larger).

Another important thing here is learning from others demonstrates a huge effect on the population as a whole, more than on one single individual agent. This can be simply explained by what has been presented so far, since our explanation for the advantage of strategic learning applies to the whole population, increasing the difference between the average fitness between EVO+Strategy and EVO+IL.

Plasticity: We have equipped every learning agent with *Plasticity* – the ability to express different phenotypic forms in response to environmental challenges. In our encoding scheme, the ability to learn is assigned by a plastic allele $?s$ which can be changed during the lifetime of an individual. Here we hypothesise that in EVO+Strategy there is a bigger proportion of plasticity than in EVO+IL. Looking at Figure 4 we observe that EVO+Strategy maintains a higher percentage of plasticity over generations in all cases. The difference between plasticity between EVO+Strategy and EVO+IL becomes bigger over time. This measurement helps consolidate our explanation so far why EVO+Strategy outperforms EVO+IL.

5 CONCLUSION, FURTHER DISCUSSION AND FUTURE WORK

We have set out to understand the role of learning strategy (i.e. the combination of learning socially and individually) in an evolving population under variable rugged fitness landscape. By proposing a simple version of dynamic NK-landscape, our experimental results illustrate that when the NK-landscape is less rugged and less variable, adding social learning does not bring much benefit to individual learning in helping an evolving population to adapt. However, when the environment becomes harder, imitation learning shows a clear benefit and facilitates individual learning, promoting the adaptation of the evolving population better than is the case under individual learning alone.

The dynamic problem used in this paper is concisely described but captures a wide variety of dynamic settings. While an individual study can of course only directly speak to the problem instances tested, the results in this study are consistent with a more general claim that a combination of social and individual learning can strengthen the performance of evolutionary algorithms when dealing with dynamic optimisation problems. The ability to learn can help individuals to efficiently track and follow the changing target. Indeed, the beneficial effects of social learning extend beyond their 'first order' impact, as an ability to usefully imitate, or learn from others, also provides better foundations for future individual learning to subsequently further improve the socially-learnt information. Future work will look at this aspect and also examine different dynamic optimisation problems.

The same ideas can also be tested on different problem domains of Evolutionary Computation, including Genetic Programming [18] in which the idea of social learning can be employed as copying semantic sub-trees between individuals, and Evolutionary Deep Learning [12] which can be employed to evolve intelligent robots [14]. As occurs in higher-order animals in the biological world, it is plausible that combining evolution, individual and social learning will assist in the creation of better simulated or indeed embodied learning agents.

ACKNOWLEDGMENT

This research is funded by the Science Foundation Ireland under Grant No. 13/IA/1850.

REFERENCES

[1] David Ackley and Michael Littman. 1992. Interactions between Learning and Evolution. In *Artificial Life II, SFI Studies in the Sciences of Complexity*, C. G. Langton, C. Taylor, C. D. Farmer, and Rasmussen S. (Eds.). Vol. X. Addison-Wesley, Reading, MA, USA, 487–509.
[2] J. Mark Baldwin. 1896. A New Factor in Evolution. *The American Naturalist* 30, 354 (1896), 441–451.

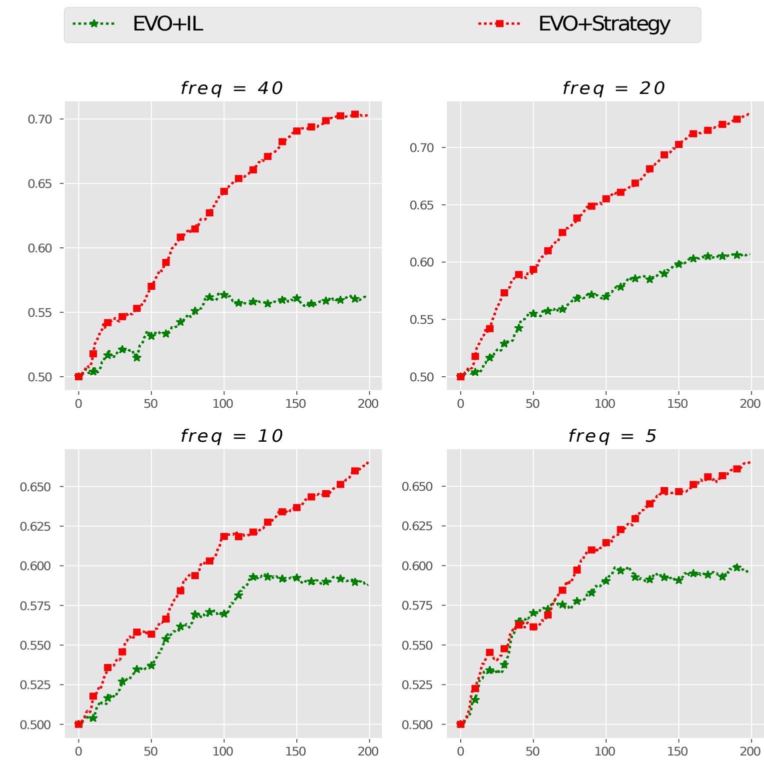


Figure 4: Plasticity Comparison.

[3] Larry Bull. 1999. On the Baldwin Effect. *Artif. Life* 5, 3 (June 1999), 241–246. https://doi.org/10.1162/106454699568764

[4] Marcus W. Feldman, Kenichi Aoki, and Jochen Kumm. 1996. *Individual Versus Social Learning: Evolutionary Analysis in a Fluctuating Environment.* Working Papers. Santa Fe Institute.

[5] C. M. Heyes. 1994. Social learning in animals: Categories and Mechanisms. *Biological Reviews* 69, 2 (may 1994), 207–231. https://doi.org/10.1111/j.1469-185x.1994.tb01506.x

[6] Geoffrey E. Hinton and Steven J. Nowlan. 1986. Adaptive Individuals in Evolving Populations. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, Chapter How Learning Can Guide Evolution, 447–454.

[7] Geoffrey E. Hinton and Steven J. Nowlan. 1987. How Learning Can Guide Evolution. *Complex Systems* 1 (1987), 495–502.

[8] William Hoppitt and Kevin N Laland. 2013. *Social learning: an introduction to mechanisms, methods, and models.* Princeton University Press.

[9] S.A. Kauffman. 1993. *The Origins of Order: Self-organization and Selection in Evolution.* Oxford University Press.

[10] Stuart A. Kauffman and Edward D. Weinberger. 1989. The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology* 141, 2 (nov 1989), 211–245. https://doi.org/10.1016/s0022-5193(89)80019-0

[11] Kevin N. Laland. 2004. Social learning strategies. *Learning and Behavior* 32 (2004), 4–14. Issue 1.

[12] Nam Le. 2019. Evolving Self-taught Neural Networks: The Baldwin Effect and the Emergence of Intelligence. In *10th Symposium on AI & Games, in AISB Convention 2019.* Falmouth, UK.

[13] Nam Le, Anthony Brabazon, and Michael O'Neill. 2018. How the "Baldwin Effect" Can Guide Evolution in Dynamic Environments. In *Theory and Practice of Natural Computing.* Springer International Publishing, 164–175.

[14] Nam Le, Anthony Brabazon, and Michael O'Neill. 2019. The Evolution of Self-taught Neural Networks in a Multi-agent Environment. In *Applications of Evolutionary Computation*, Paul Kaufmann and Pedro A. Castillo (Eds.). Springer International Publishing, Cham, 457–472.

[15] Nam Le, Michael O'Neill, and Anthony Brabazon. 2018. Adaptive Advantage of Learning Strategies: A Study Through Dynamic Landscape. In *Parallel Problem Solving from Nature – PPSN XV.* Springer International Publishing, 387–398.

[16] Nam Le, Michael O'Neill, and Anthony Brabazon. 2019 forthcoming. How Learning Strategies Can Promote an Evolving Population in Dynamic Environments. In *IEEE Congress on Evolutionary Computation, CEC 2019.* IEEE Press, Wellington, New Zealand.

[17] Nam Le, Michael O'Neill, and Anthony Brabazon. forthcoming. The Baldwin Effect reconsidered through the prism of social learning. In *IEEE Congress on Evolutionary Computation, CEC 2018.* IEEE Press, Rio, Brasil.

[18] Nam Le, Hoai Nguyen Xuan, Anthony Brabazon, and Thuong Pham Thi. 2016. Complexity measures in Genetic Programming learning: A brief review. In *2016 IEEE Congress on Evolutionary Computation (CEC).* IEEE. https://doi.org/10.1109/cec.2016.7744087

[19] Giles Mayley. 1996. Landscapes, Learning Costs, and Genetic Assimilation. *Evolutionary Computation* 4, 3 (sep 1996), 213–234. https://doi.org/10.1162/evco.1996.4.3.213

[20] Giles Mayley. 1997. Guiding or Hiding: Explorations into the Effects of Learning on the Rate of Evolution.. In *In Proceedings of the Fourth European Conference on Artificial Life.* MIT Press, 135–144.

[21] Stefano Nolfi, Domenico Parisi, and Jeffrey L. Elman. 1994. Learning and Evolution in Neural Networks. *Adaptive Behavior* 3, 1 (1994), 5–28.

[22] Robert Boyd Peter J. Richerson. 2005. *Not By Genes Alone: How Culture Transformed Human Evolution* (1 ed.). University of Chicago Press.

[23] Luke Rendell, Laurel Fogarty, William J.E. Hoppitt, Thomas J.H. Morgan, Mike M. Webster, and Kevin N. Laland. 2011. Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences* 15, 2 (feb 2011), 68–76.

[24] Peter J. Richerson Robert Boyd. 1985. *Culture and the Evolutionary Process.* Chicago: University of Chicago Press.

[25] Alan R. Rogers. 1988. Does Biology Constrain Culture? *American Anthropologist* 90, 4 (1988), 819–831. https://doi.org/10.1525/aa.1988.90.4.02a00030

[26] Joe Yuichiro Wakano, Kenichi Aoki, and Marcus W. Feldman. 2004. Evolution of social learning: a mathematical analysis. *Theoretical Population Biology* 66, 3 (2004), 249 – 258.