Adaptive Advantage of Learning Strategies: A Study Through Dynamic Landscape

Nam Le(✉), Michael O'Neill, and Anthony Brabazon

Natural Computing Research and Applications Group,
University College Dublin, Dublin, Ireland
`namlehai90@gmail.com`

Abstract. Learning can be classified into two categories: *asocial* learning, e.g. trial-and-error; and *social* learning, e.g. imitation learning. Theory using mathematical models suggest that social learning should be combined with asocial learning in a strategic way (called *learning rule* or *learning strategy*), and that that combination should be scrutinised under different environmental dynamics, to see how advantageous the learning rule is. More interestingly, learning has been shown to be beneficial to the evolutionary process through the **Baldwin Effect**. This paper investigates the adaptive advantage of social learning when combined with asocial learning under a number of environmental variations. We propose a Dynamic Landscape as well as an algorithm combining both asocial and social learning in order to test our hypotheses. Experimental results show that if each individual in the population is either asocial or social, but not both, the average fitness of the population decreases when the proportion of social learners increases as the environment changes. Moreover, a population consisting entirely of asocial learners outperforms the previous type of population. If every individual agent in the population can perform both asocial and social learning depending on a strategic rule, the evolving population outperforms the two previous populations with respect to average fitness.

Keywords: Social learning · Cultural evolution · Genetic algorithm
Dynamic environment · Baldwin effect

1 Introduction

Evolution and learning are two different ways in which the behavior, and other traits, of organisms can change in order to adapt to environmental variations. Evolution is change at the genetic level of a population, while learning, on the other hand, is change at the phenotypic level of an individual. The idea that the two forms of adaptation interact and complement each other was once proposed by Baldwin [1], called the Baldwin Effect. Hinton and Nowlan (henceforth

A. Auger et al. (Eds.): PPSN 2018, LNCS 11102, pp. 387–398, 2018.
https://doi.org/10.1007/978-3-319-99259-4_31

H&N) presented a computer model to investigate the Baldwin Effect in simulation [2], showing that learning, more specifically asocial learning, facilitates the evolutionary process and enhances the fitness of the population in a Needle-in-a-haystack landscape. Their initial success motivated several further studies, such as [3,4], to show how learning can enhance the evolutionary search.

Generally, learning can be classified into two forms. *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 – are two alternative ways for an individual agent to acquire information from the environment at the phenotypic level. SL has been observed in organisms as diverse as primates, birds, fruit flies, and especially humans [5]. Although the use of SL is widespread, understanding when and how individuals learn from others is a significant challenge. SL is generally less time-consuming, but relies on information produced by others. So when the environment changes, the information from others is likely to be outdated and SL becomes *maladaptive* (not adaptive). On the other hand, IL through trial-and-error is costly, but capable of producing new information when the environment happens to change.

This opens a curious question when an organism should rely on SL rather than IL, and under which environmental condition social learning would evolve, or be adaptive. Several theoretical models have shown that individual agents capable of learning in a strategic way outperform those that are able to learn individually or socially, but not both [6–8].

The main aim of this paper is to investigate, through computer simulation, whether organisms should rely on SL or IL, and what the plausible strategy for an organism could be when the environment changes. We combine evolution and both forms of learning to see how they behave under a dynamic landscape we call Dynamic Needle-in-a-hay-stack. This paper is built upon the success of the previous work in [9]. In the remainder of this paper, we briefly present research on learning and evolution. Social learning and related concepts are briefly introduced. We in turn describe the experiments we use in this paper. Results are analysed and discussed, then the conclusion and some future directions are proposed.

2 Background

2.1 Social Learning

SL has been studied in various disciplines, including Cognitive Biology, Evolutionary Psychology, Behavioral Ecology, Cognitive Science and Robotics. In general, SL covers several mechanisms through which individual organisms learn from others, such as stimulus enhancement, observational conditioning, imitation, and emulation (please refer to [5,10] for the definition of these mechanisms). In this study we focus on one of these mechanisms, namely imitation learning. In this instance of learning, the observer directly copies the behavior of the observed animal in order to complete a novel task.

SL, at first glance, seems to be adaptive at a low cost when 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 [6,11]. It is argued that individual learners produce new information about the environment, though at a cost. Social learners avoid this cost by copying the existing behaviors of others, but do not generate new information. 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 [6–8]. It is said that social learning should be combined with individual learning in a strategic way in order to have an adaptive advantage. Social learning strategies consist of rules specifying the way an individual relies on social learning by answering three questions as follows:

i. When an individual should learn;
ii. From whom they should learn; and
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.

In addition to the *Who* question, the transmission from demonstrators to observers are classified into three types [12]. 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 from the previous generation but differs from its parent. The last is *horizontal transmission* in which an observer learns from a demonstrator in its current generation. In the scope of this paper, we only use *oblique transmission* in our experiments.

2.2 Learning and Evolution in Computer Simulation

H&N presented a classic paper in 1986 [2] to demonstrate an instance of the Baldwin Effect in computer. We discuss this model in detail for clarity. In H&N's model, suppose the task is to find the all-ones target string 111...1 (20 bits). There is only one correct solution – an individual with configuration exactly matched with the target string – which has the fitness of 20. All other configurations are wrong and have the same fitness of 1. This forms a Needle-in-a-haystack landscape whereby an evolutionary search alone cannot find the solution [2].

H&N presented an idea that encodes an individual's genotype whereby one part is fixed by genetic-like information, and the other part is plastic which allows for learning during the lifetime of the individual. Each individual agent has a *genotype*—a string of twenty characters. Each position in a genotype, or *locus*, can have three alternative values: ‘0’, ‘1’, and ‘?’. Each locus is randomly initialized with 25% chance of being assigned a ‘0’, 25% chance of being ‘1’, and 50% chance a ‘?’.

In addition to the above two types of agent (correct or incorrect), there exists another type of agent – called *potential* individual – which will be allowed for life-time learning. The genotype-phenotype mapping is one-to-one and at birth, each individual has its phenotype string identical to its genotype string. An individual is potential only if in its initial genotype, every locus excluding locus with plastic value ‘?’ is matched with corresponding locus in the target string. In case of H&N’s problem, a potential individual could have its initial genotype comprising of only ‘1’ and ‘?’. The allele ‘?’ allows for lifetime learning (or plasticity), over 1000 rounds. 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. After learning, the fitness of that potential individual agent x_i is calculated as:

$$f(x_i) = 1 + \frac{19(1000 - n)}{1000} \tag{1}$$

in which n is the number of trials required to find the correct combination of alleles - the all-one string. It can be inferred from the fitness function that the more trials an agent needs, the lower the fitness it will get. By allowing lifetime learning, H&N showed that learning can create a gradient which facilitates evolution to search for the solution.

Since the success of H&N’s model, there have been a number of studies showing that learning can enhance an evolutionary process, especially when the environment is changing [3,4]. Recently, Le et al. [9] presented a model building on H&N’s simulation, in which they combine evolution, asocial and social learning. It was shown that social learning alone fails to search on Needle-in-a-haystack, but social learning when coupled with individual learning outperforms individual learning alone with respect to average fitness of the population.

In this paper, we build on the success of the previous simulation in [9], in which we propose a dynamic version of H&N’s Needle-in-a-haystack to see how individual learning or social learning behaves under different environmental dynamics. Experimental designs are discussed in the following section.

3 Experimental Design

3.1 Dynamic Needle-in-a-haystack Landscape

We create a dynamic version of H&N’s landscape called Dynamic Needle-in-a-haystack, in which we use two parameters to control the dynamics of the landscape, namely *frequency* and *magnitude* of change. The first parameter tells us after many generations the target (needle) will move to another point in the landscape, while the latter helps determine the likelihood of change for each element of the target. Assume that at a generation g the target is all-one (20 bits of one), *frequency* = 10 and *magnitude* = 0.1 (10%). This informs us that after 10 generations or at generation $g + 10$ the target t = 111...1 (20 bits of 1) is likely to be changed. The magnitude of 0.1 tells us that there are, on average, $20 \times 10\% = 2$ bits in the target that are likely to be modified. For each bit in the target sequence, a random number is generated and then compared with the *magnitude*: if the random value is less than 0.1, the current bit is mutated to its subtraction from 1 (1 becomes 0, and vice versa). Suppose the new target at generation $g + 10$ is $t_1 = 001...1$ (two first bits are changed). There is only one right sequence of bits that exactly matches the new target t_1 and has the fitness of 20. Otherwise, all other configurations are incorrect and get the same fitness of 1. This landscape, again, constitutes a Needle-in-a-haystack, but the needle is moving after a number of generations. That is why we call this landscape *Dynamic Needle-in-a-haystack*.

We also call the period when the environment is unchanged the *interval of stability*. Therefore, the interval of stability has the same value as the dynamic frequency.

3.2 Experiment Setup

In this section, we present the experimental setups used in our paper. It is often said that evolutionary search finds it hard to search in 'Needle-in-a-haystack' landscape. Furthermore, it was claimed that an evolutionary search alone failed in this type of landscape [2]. Le et al. [9] went further to show that an evolutionary search combined with social learning alone also failed to find a solution in a Needle-in-a-haystack. We conduct three experiments with the parameter settings, as shown in Table 1.

Table 1. Parameter setting

Parameter	Value
Original target	111...1 (20 bits of 1 s)
Genome length	20
Replacement	Generational
Generations	50
Population size	1000
Selection	Fitness-Proportionate selection
Reproduction	Sexual reproduction
Mutation rate	0.01
Fitness function	Eq. 1
Maximal learning trials	1000
Frequency	5, 10, 20
Magnitude	0.05, 0.075, 0.1

We run our experiments through 9 different combinations of *frequency* and *magnitude*. For frequencies 5, 10, 20 there will be 10, 5, and 2 times of change in the environment, respectively. In case frequency = 5 or 10, generation 50 will see a change in the environment. It can be understood that the lower the *frequency* value, the faster the target will change; the bigger the value of *magnitude*, the bigger the change of the target. The environment becomes more dynamic or harder to cope with by faster changing and bigger magnitude of change, and vice versa.

Please note that, unlike the so-called memetic algorithm and Lamarckian Evolution, learning in our experiments only happens at the phenotypic level, what an individual learns does not change its genotype. The recombination operators work on the genotypic level, so children may inherit question marks from their parents.

Setup I: A Population of Individual Learners

The first experimental setup is an evolving population of individual learners only (as in H&N model), in which an evolutionary search is combined with IL (denoted EVO+IL). IL performs a local search process by which each '?' allele will guess its value to be '0' or '1' in each learning trial. The evolutionary algorithm in our experiment is a genetic algorithm with crossover and mutation (with mutation rate of 0.01).

Setup II: A Population of Single-Role Learners

In the second experiment setup, we simulate a population of single-role individuals – individuals that are either social learners or individual learners, but not both (denoted EVO+IL&SL). The reason for this experimental design is that we are curious to know how social learning or individual learning would evolve under various environmental dynamics and how they contribute to the average fitness of the population. Unlike EVO+IL, we have two types of individuals in the population now. We add one more bit, or *gene*, called *learning mode*, which is either 0 or 1, onto the genome of each individual. If that value is 0, the individual is likely to learn individually; conversely, if that value is 1, the individual is likely to learn socially. A noteworthy point here is that in our landscape, only potential individuals are able to perform lifetime learning. That means, social and asocial learners are potential individuals with learning mode equal to 1 and 0, respectively. Learning mode is initialised with 50% at 0 and 50% at 1. It should be noted that through recombination, the learning mode of a child is set to be the learning mode of the better individual between its parent. Mutation does not touch the learning mode of the child.

In order to implement social learning, first we propose the imitation procedure, with pseudo-code described in Algorithm 1 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 decide whether to copy exactly the *trait* or a mutated version of that *trait* from the demonstrator.

Algorithm 1. IMITATION

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

At each generation, an asocial agent learns by itself like in EVO+IL model, whereas a social learner imitates its demonstrator. Every individual agent has the same demonstrator that is the best individual in terms of fitness from the previous generation. Because we adopt *oblique transmission*, there is no SL at the initial generation.

Experiment III: A Population of Strategic Individuals

The third setup we evolve a population of *strategic* individuals – individuals that can perform both SL and IL based on a learning rule (denoted EVO+Strategy). Unlike EVO+IL&SL, the population now has just one type of individual - strategic individuals. We specify the learning strategy for every individual agent as follows: At each generation, an agent first looks at its demonstrator (chosen the same as in EVO+IL&SL), and determines whether to learn from that demonstrator or not. If the demonstrator is still *adaptive* in the current generation, the agent imitates the demonstrator based on Algorithm 1; otherwise, the agent learns individually. The demonstrator is said to be *adaptive* in the current environment if its phenotype exactly matches with the target in the current environment. This means every agent determines whether it expresses as an individual learner or as a social learner based on a given rule. After lifetime learning process for each agent, the population goes through selection and reproduction as in EVO+IL.

4 Results, Analyses, and Explanations

First, we look at the average fitness of the population as a measurement of how well each simulation performs. All plots are grouped together, sharing the same labels for x-axis and y-xis as well as the annotation. The results are presented and discussed in a comparative manner below.

A similar trend can be recognised in Fig. 1 that there is a drop in every population with all settings at the generation when the environment begins to change. This is understandable because when the environment changes, a number of adaptive behaviors from previous generations are no longer fit in the current generation, reducing the average fitness of the population. By looking at the behavior of each corresponding line through each row or each column of Fig. 1, we can see another shared behavior that the more dynamic, or difficult, the environment is, the lower the average fitness of the population.

How does each type of population comparatively cope with these environmental dynamics? It is shown that EVO+IL outperforms EVO+IL&SL in all environmental circumstances. More than that EVO+IL&SL shows its inability to track the dynamics of the environment as the average fitness goes down to or much closer to the lowest value of 1 in most cases, except the easiest landscape (when the environment changes most slowly and the magnitude of change is smallest) though it just reaches around 2.5 at the end.

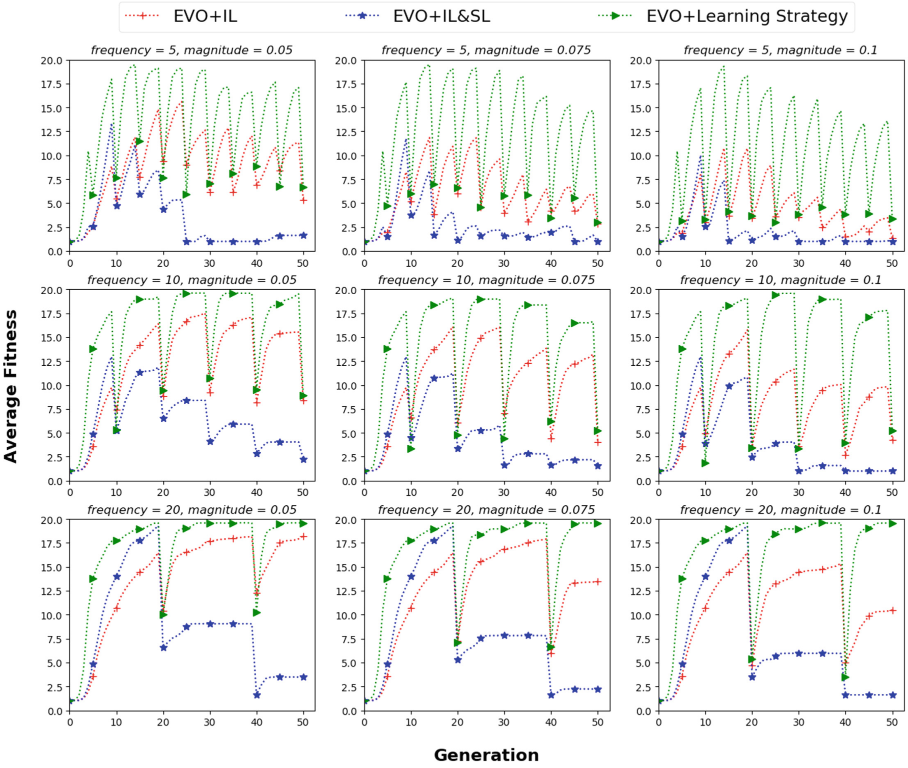

Fig. 1. Average fitness over generations. The red, blue, and green dotted-lines for EVO+IL, EVO+IL&SL, and EVO+Strategy, respectively. (Color figure online)

Remember that EVO+IL&SL evolves a population comprising of both asocial and social learners, so why does it behave very badly and even worse than the population with individual learners alone? We have briefly talked of the fact that SL, though less costly, is 'information parasitism' that is unable to produce new information. When the environment changes, social learners are likely to copy outdated information in case their demonstrators are no longer adaptive in the current generation. In our experiments, the best individual at generation g is set to be the demonstrator for all social learners in generation $g+1$. Assume that the environment changes at generation $g+1$, the demonstrator becomes *maladaptive*. All social learners at generation $g+1$ copy *maladaptive* behavior, thus becoming *maladaptive* and get the same lowest fitness of 1 and reducing the average fitness of the population.

However, the population in EVO+IL&SL is composed of both asocial and social learners at the initial generation. It is said that individual learners through trial-and-error are able to track the dynamics in the environment. So why are individual learners unable to help the population to cope with environmental dynamics? We hypothesise that in EVO+IL&SL the proportion of asocial learner decreases to a very small amount so that there are not enough asocial learners to track the environmental changes.

We know that the imitative process gives social learners an advantage in terms of time and learning trial required to find the correct solution. While asocial learners have to trial-and-error through a number of trials, social learners just need to copy all bits from the correct demonstrator's phenotype to its phenotype. Based on our fitness function (Eq. 1), the more the learning trials needed, the lower the fitness value. Consequently, lower cost gives social learners advantages over asocial learners when the environmental is stable. This argument can be verified by looking at Fig. 1 as at some initial generations, the average fitness of EVO+IL&SL is higher than that of EVO+IL during the interval of stability. Therefore, natural selection will favor social learners during these earlier generations, individual learners become less dominant and are likely to disappear.

We calculate the frequency of asocial and social learners over generations to verify this hypothesis. Figure 2 shows that in EVO+IL&SL the frequency of asocial learner is very low in all cases and tends to go down to zero sooner when the environment becomes harder. This fits with our above hypothesis. In addition to the above analyses, remember that in EVO+IL&SL each individual is initially encoded as either an asocial or a social learner. When social learners are more likely to be favoured by natural selection, more social learner 'genes', or learning mode 1, occur in the reproductive pool, so that offspring produced through sexual recombination are more likely to have the learning mode as social learners. The asocial learner gene becomes less prevalent over time, and in most cases we see it becomes distinct. That is why as the environment changes, the population has less asocial learners to track the environment, hence the average fitness reduces down to the lowest value of 1 in almost all cases.

Conversely, EVO+IL still maintains a higher number of asocial learners than that of EVO+IL&SL to track environmental variations. That is why EVO+IL has a higher average fitness than that of EVO+IL&SL in all cases.

One important point to be extracted here is that SL can give a population advantages when the environment is in a stable interval, whereas IL is much more powerful at the point when the environment changes. That is why we have designed the learning rule in EVO+Strategy to make use of the advantages of both IL and SL. It is easily seen in Fig. 1 that EVO+Strategy outperforms other settings in all cases with respect to average fitness. Moreover, during the interval of stability, EVO+Strategy can reach higher point of average fitness than EVO+IL, especially in the hardest case (frequency = 5, magnitude = 0.1).

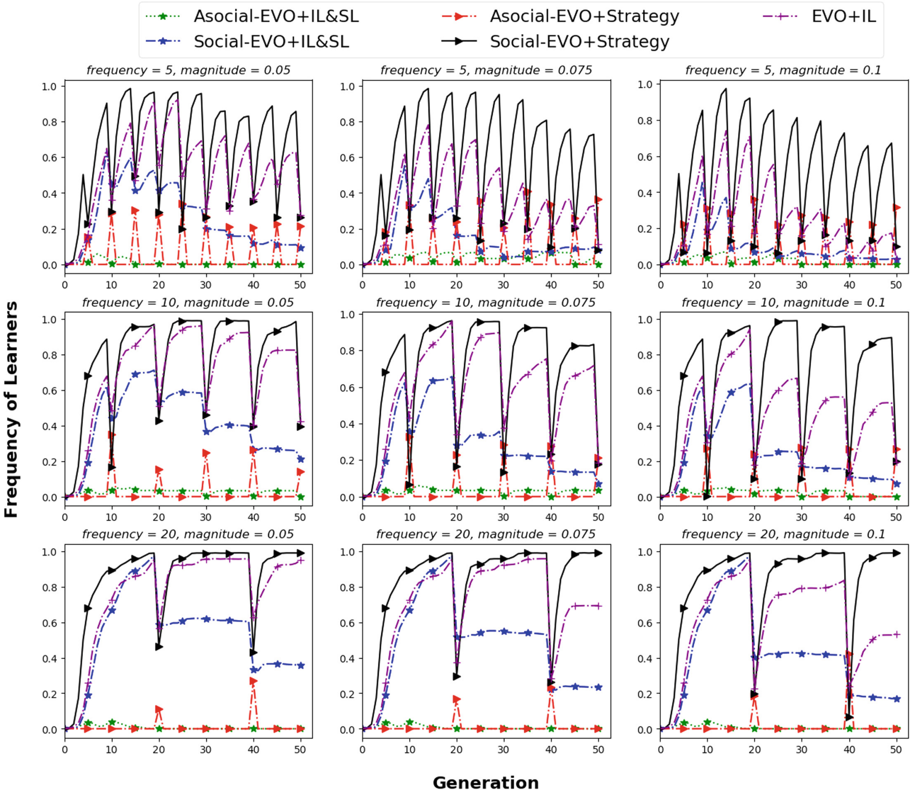

Fig. 2. Frequency of social and asocial learners over generations. Green and blue, red and black lines show the frequency of asocial and social learners in EVO+IL&SL and EVO+Strategy, correspondingly. The purple lines represents the frequency of asocial learners in EVO+IL. (Color figure online)

One plausible explanation for the superiority of EVO+Strategy is that the learning rule used in EVO+Strategy can exploit the advantage of both IL and SL. Remember that every individual agent in EVO+Strategy can perform as either a social learner or an asocial learner. At each generation, each strategic agent checks if its demonstrator is adaptive or not: if adaptive, it imitates the demonstrator; otherwise the agent learns asocially. By this learning rule, EVO+Strategy removes the case that agents will learn maladaptive behaviors like what has been observed in EVO+IL&SL by allowing more asocial learning in case the environment changes. Moreover, when the environment is in a stable interval, the system now allows for more SL to be expressed to make use of the advantage of SL over IL, hence the average fitness is remarkably increased.

We also calculate the frequency of social and asocial learners in EVO+Strategy. It can be observed in Fig. 2 that the frequency of asocial learner is closer to or becomes zero when the environment is stable. However, at the point when the environment changes, the frequency of asocial learners increases. The frequency of social learners, in contrast, goes down at the point when the environment changes, reducing the 'information parasitism' issue and the problem of copying outdated information. Furthermore, when there are some asocial learners to help the population to track the environmental dynamics, the population is shown to maintains a high proportion of social learners and this percentage goes higher in every stable interval. Therefore, the average fitness of the population increases during every interval of stability. This observation fits with the description of the system as well as the behavior and analyses we have given above on the average fitness.

It can also be observed that during the interval of stability, the frequency of asocial learners in EVO+IL is lower than that of social learners in EVO+Strategy in all cases. This gives EVO+Strategy an advantage over EVO+IL to maintain a higher average fitness because social learners can copy correct behaviors at a lower cost compared to asocial learners during an interval of stability.

5 Conclusion and Future Work

We have set out to understand the role social learning may have on the evolutionary process in various environmental dynamics. For the specific landscape and the parameter setting used in this paper, experimental results have empirically shown that social learning is more advantageous when the environment is stable, whereas when the environment happens to change asocial learning is required to track the environment. A learning rule combining both social and asocial learning has been designed and the population with a learning strategy has shown to have a much better adaptive advantage, measured by the average fitness. Several plausible explanations have been presented in this paper for these observations.

In the scope of this paper, we have mainly discussed the adaptive advantage of social learning, asocial learning, and learning strategies in dynamic environments. The evolution of social learning can also been extracted from our results and analyses on the frequency of learners. It is suggested to investigate the question as to how social learning evolves more deeply in future work. It is also recommended that we use different forms of social transmission, such as *vertical transmission* and *horizontal transmission*.

The learning strategy used in this paper is designed by the system designer. In future work, we would like to let the individuals themselves evolve their own learning strategies. One proposal is to create *genes* controlling a learning strategy for each individual agent as was done with EVO+IL&SL. Instead of encoding just one bit for learning mode, we can encode two bits to specify the learning rule, whether to be expressed as social or asocial learning. By doing this, we are able to observe the dynamics of rule changing when the environment changes over time. The motivation for this is that we want to let evolution to optimise the learning strategy for each individual agent, as evolution has done for living organisms, including humans [6].

The paper has empirically verified that social learning should be used in a savvy way to enhance the behavior of a population, complementing to some theoretical findings in the trans-disciplinary research on social learning. Future work will investigate the method and verify the findings in this paper on different types of problems and landscapes, such as NK-landscape [13] or Artificial Life and Robotics domains [3,4].

References

1. Baldwin, J.M.: A new factor in evolution. Am. Nat. **30**(354), 441–451 (1896)
2. Hinton, G.E., Nowlan, S.J.: How learning can guide evolution. In: Belew, R.K., Mitchell, M. (eds.) Adaptive Individuals in Evolving Populations, pp. 447–454. Addison-Wesley Longman Publishing Co., Inc, Boston (1986)
3. Ackley, D., Littman, M.: Interactions between learning and evolution. In: Langton, C.G., Taylor, C., Farmer, J.D., Rasmussen, S. (eds.) Artificial Life II, SFI Studies in the Sciences of Complexity, vol. X, pp. 487–509. Addison-Wesley, Reading (1992)
4. Nolfi, S., Parisi, D., Elman, J.L.: Learning and evolution in neural networks. Adapt. Behav. **3**(1), 5–28 (1994)
5. Heyes, C.M.: Social learning in animals: categories and mechanisms. Biol. Rev. **69**(2), 207–231 (1994)
6. Laland, K.N.: Social learning strategies. Learn. Behav. **32**, 4–14 (2004)
7. Feldman, M.W., Aoki, K., Kumm, J.: Individual versus social learning: evolutionary analysis in a fluctuating environment. Santa Fe Institute, Working Papers (1996)
8. Wakano, J.Y., Aoki, K., Feldman, M.W.: Evolution of social learning: a mathematical analysis. Theor. Popul. Biol. **66**(3), 249–258 (2004)
9. Le, N., O'Neill, M., Brabazon, A.: The Baldwin effect reconsidered through the prism of social learning. In: IEEE Congress on Evolutionary Computation, CEC 2018, 8–13 Jul forthcoming. IEEE Press, Rio (2018)
10. Hoppitt, W., Laland, K.N.: Social Learning: An Introduction to Mechanisms, Methods, and Models. Princeton University Press, Princeton (2013)
11. Rogers, A.R.: Does biology constrain culture? Am. Anthropol. **90**(4), 819–831 (1988)
12. Richerson, P.J., Boyd, R.: Culture and the Evolutionary Process. University of Chicago Press, Chicago (1985)
13. Suzuki, R., Arita, T.: Repeated occurrences of the Baldwin effect can guide evolution on rugged fitness landscapes. In: 2007 IEEE Symposium on Artificial Life. IEEE, April 2007