

Mechanisms of Social Learning in Evolved Artificial Life

Alberto Bartoli¹, Marco Catto², Andrea De Lorenzo¹, Eric Medvet^{1,2}, and Jacopo Talamini²

¹Machine Learning Lab, Department of Engineering and Architecture, University of Trieste, Italy

²Evolutionary Robotics and Artificial Life Lab, Department of Engineering and Architecture, University of Trieste, Italy
bartoli.alberto@units.it, marco.catto@hotmail.it, andrea.delorenzo@units.it, emedvet@units.it, jacopo.talamini@phd.units.it

Abstract

Adaptation of agents in artificial life scenarios is especially effective when agents may evolve, i.e., inherit traits from their parents, *and* learn by interacting with the environment. The learning process may be boosted with forms of *social learning*, i.e., by allowing an agent to learn by combining its experiences with knowledge transferred among agents. In this work, we tackle two specific questions regarding social learning and evolution: (a) from whom learners should learn? (b) how should knowledge be transferred? We address these questions by experimentally investigating two scenarios: a simple one in which the mechanism for evolution and learning is easily interpretable; a more complex and realistic artificial life scenario in which agents compete for survival. Experimental results show that social learning is more profitable when (a) the learners learn from a small set of good teachers and (b) the knowledge to be transferred is determined by teachers experience, rather than learner experience.

Introduction

Artificial agents adapt to the environment according to different mechanisms and along different time scales: *evolution* is a long-term process in which agents inherit traits from their parents, whereas *learning* is a short-term process in which agents adapt by interacting with the environment autonomously or through knowledge transfer among agents, i.e., through *social learning*. Combinations of individual and social learning with evolution have been shown to be successful for adaptation (Le et al., 2020).

In this work, we tackle two specific questions regarding social learning in evolved artificial life: (a) from whom learners should learn? (b) how should knowledge be transferred? We address these questions by investigating two artificial life scenarios: one in which agents have a fixed life-span and do not compete among themselves; another in which agents compete for survival by moving and collecting food in a toroidal environment. In the first, simpler scenario, social learning occurs by directly copying *traits* from the teachers to each learner. In the more complex scenario, a learner attempts instead to match the *behavior* of its teachers on a set of situations, i.e., perceptions of the environment provided by sensors.

We analyzed a number of different configurations in terms of size and composition of the set of teachers and of the way knowledge is transferred from teachers to learners. Our experiments indicate that social learning effectiveness depends on teachers size and knowledge transfer. Concerning the former factor, social learning is more effective when the learner learns from a set of good teachers, rather than from just the best teacher: moreover, with more than one teacher, learning is more robust to imperfect trait transfer. Concerning the latter factor, social learning results in more successful learners when they attempt to reproduce teachers behaviors on teachers, rather than learners, experience.

Related works

Learning and evolution interplay Lifetime learning in evolution is a long-standing topic that has been treated, among the firsts, by Baldwin (1896): in his seminal paper, the author claims that learning is not only allowed by evolutionary conditions, but that it also influences the overall adaptation of a species over the time. In other words, Baldwin (1896) claims that, while evolution takes place at the genotype level and learning takes place at phenotype level, there is an interplay between them. A recent overview on the literature on the so called “Baldwin effect”, and a possible different interpretation of the phenomenon, has been proposed by Le (2019b).

In an early application of the lifetime learning paradigm combined with evolution, Hinton and Nowlan (1987) considered a simple artificial life scenario where the individual employs an ANN containing many potential connections. The reproductive fitness of each organism is the number of correct connections, meaning that pure evolutionary search can only discover the optimal configuration by randomly exploring the space of possible configurations. The authors show that specifying some individual traits as inherited and others as learned, the search may be more efficient.

The effectiveness of learning combined with evolution in the survival of the species has been showed in a more complex artificial life scenario by Ackley and Littman (1991), where each agent is embodied by an ANN, and lifetime

adaptation occurs through Reinforcement Learning (RL).

Neuroevolution Evolutionary optimization has been shown to be an effective mechanism of adaptation when the agent consists (also) of an ANN. In (Yao, 1999; Floreano et al., 2008), both the architecture of the network and its weights can be subjected to the optimization. Le (2019a) proposes a strategy based on evolving self-taught ANN for a multi-agent scenario. In (Keesing and Stork, 1991; Nolfi et al., 1994) evolution and lifetime learning are employed on an ANN, and here the learning strategy adopts backpropagation to update the network parameters. Soltoggio et al. (2018) propose an Evolved Plastic ANNs framework that introduces plasticity in ANNs, hence allowing evolution to optimize more than just the weights of the ANN.

Social learning Animals do not learn only by themselves, i.e., without involving social interactions, but they learn also from the other members of the same species using different strategies (Laland, 2004), which are referred as social learning strategies (Heyes, 1994).

Van Schaik and Burkart (2011) observe that, not only social learning is possible in animal species, but it is even preferred to individual learning, as it allows to transfer knowledge in a more efficient way. It seems also that the species that practice social learning have a more advanced set of cultural skills, and tend to be more responsive to the evolutionary selection (van der Post et al., 2016). This belief is supported by the results obtained in the mathematical model designed by Wakano et al. (2004).

Some studies investigate, also comparatively, both individual and social learning (Mesoudi et al., 2016). Heiner et al. (2018) consider learning occurring in a group of physical robots, and two different scenarios, where each robot is controlled by an ANN. They show that social learning is an efficient strategy to reduce the time required for finding the optimal solution. Le et al. (2018) takes into account a simulated scenario, and finds that also social learning gives the “Baldwin effect”, i.e., that it improves the overall evolutionary process of a population of individuals. A comparison between individual learning and social learning, as alternative forms of adaptation, is given by Feldman et al. (1996) and Bullinaria (2017), where the authors provide empirical results on the power of social learning in a rather simple artificial life scenario. Borenstein et al. (2008) observe that social and individual learning can evolve independently of each other as alternative adaptation mechanisms. Annunziato and Pierucci (2003), and more recently Le et al. (2020), have investigated the relationship between social learning and individual learning, and the combination of the two learning strategies in relative complex artificial life scenarios, where agents are controlled by ANNs.

The experiments performed by Marriott and Chebib (2014) in an artificial life scenario of resource gathering

show that there seems to be an evolutionary pressure to promote individuals with the ability to learn, rather than preferring individuals with a higher fitness. Moreover social learning seems to support an adaptation trend that is decoupled from the one of evolution (Marriott and Chebib, 2014; Chebib and Marriott, 2016), a finding that is similar to the results achieved by Borenstein et al. (2008). Jolley et al. (2016) propose different criteria for choosing the agent to teach the other agents in social learning, being either a parent, the fittest individual, the oldest individual, a random individual or another young agent, however they do not consider learning from multiple teachers, which is a research question that we address in this work. Relevant social learning aspects, such as the choice of whom to learn from in a population, and what aspects of culture are transmitted, have not been completely understood yet, and are open to future research (Marriott et al., 2018). In the present study, we attempt to answer a few of these questions in the context of artificial life.

Simple scenario

We consider a simple scenario inspired by the results obtained by Le et al. (2018) and the *Hinton and Nowlan model* (Hinton and Nowlan, 1987). In this model, a population of agents evolves on the long-term evolutionary time scale, whereas the short-term life-time time scale is instantaneous.

An *agent* is defined by a genotype \mathbf{g} and a phenotype \mathbf{p} . The *genotype* is a fixed-length sequence $\mathbf{g} = (g_1, \dots, g_l) \in \{0, 1, ?\}^l$ of l genes that can assume values, called *alleles*, in $\{0, 1, ?\}$. The *phenotype* is a fixed-length sequence $\mathbf{p} = (p_1, \dots, p_l) \in \{0, 1\}^l$ of l traits.

The phenotype is determined from the genotype according to a genotype-phenotype mapping function $m : \{0, 1, ?\}^l \rightarrow \{0, 1\}^l$ that we use, in this model, to represent learning (explained below). The mapping function is applied as soon as the individual is generated. Similarly, the *fitness* of an individual is calculated as soon as the genotype is mapped to the phenotype. The fitness is the ratio of 1s in the phenotype, i.e., $f(\mathbf{p}) = \frac{1}{l} |\mathbf{p}|_1 = \frac{1}{l} |\{1 \leq i \leq l : p_i = 1\}|$. It follows that the optimal phenotype \mathbf{p}_{opt} , i.e., the one with the maximum fitness $f(\mathbf{p}_{\text{opt}}) = 1$, is unique in $\{0, 1\}^l$.

The population of agents changes in the evolutionary time scale according to an evolutionary algorithm (EA) in which fitter agents have more chances to reproduce and survive. Algorithm 1 shows the EA that drives the evolution of the population. It is a standard EA that evolves a fixed-size population of n_{pop} individuals, initially set randomly, for n_{gen} generations, with a non-overlapping generational model—i.e., the population of the i -th generation does not contain individuals of the $(i - 1)$ -th generation. Individuals are selected for reproduction with a tournament selection of size n_{tour} : once the genotypes $\mathbf{g}_1, \mathbf{g}_2$ of two individuals (parents) are selected, the genotype of the new individual is obtained by first applying the uniform crossover to $\mathbf{g}_1, \mathbf{g}_2$ and then

the gene-wise mutation to the resulting genotype. The gene-wise mutation consists in changing, with a probability p_{mut} , each gene to a random allele. We remark that genetic operators operate on \mathbf{g} , rather than \mathbf{p} : the evolution does not hence explicitly propagate learned traits, that is, it is a Darwinian, rather than Lamarckian, evolution.

```

1  $P \leftarrow \text{initialize}(n_{\text{pop}})$ 
2 foreach  $i \in \{1, \dots, n_{\text{gen}}\}$  do
3    $P' \leftarrow \emptyset$ 
4   foreach  $j \in \{1, \dots, n_{\text{pop}}\}$  do
5      $\mathbf{g}_1 \leftarrow \text{selectTournament}(P, n_{\text{tour}})$ 
6      $\mathbf{g}_2 \leftarrow \text{selectTournament}(P, n_{\text{tour}})$ 
7      $\mathbf{g}' \leftarrow \text{mutate}(\text{crossover}(\mathbf{g}_1, \mathbf{g}_2))$ 
8      $P' \leftarrow P' \cup \{\mathbf{g}'\}$ 
9   end
10   $P \leftarrow P'$ 
11 end

```

Algorithm 1: The EA of the simple scenario.

We consider 3 different variants for the genotype-phenotype mapping function m that represent the cases of no-learning (EVO), individual learning (IL), and social learning (SL). In the variants representing the cases with learning, the ? allele in the genotype represents learnability, i.e., if $g_i = ?$, then that i -th trait can be learned, otherwise it is determined statically from the corresponding i -th gene.

No learning (EVO) In this variant, no learning occurs and learnable genes are mapped to 0:

$$\mathbf{p} = m_{\text{EVO}}(\mathbf{g}) = (m'_{\text{EVO}}(g_1), \dots, m'_{\text{EVO}}(g_l)), \quad (1)$$

where:

$$m'_{\text{EVO}}(g_i) = \begin{cases} 1 & \text{if } g_i = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Individual learning (IL) This variant represents a two-steps process in which (i) the agent performs a number n_{iter} of agent-environment interactions by randomly varying its learnable traits and (ii) the phenotype is set to the one corresponding to best interaction in terms of fitness. Formally, a set $\{\mathbf{p}_1, \dots, \mathbf{p}_{n_{\text{iter}}}\}$ of realizations of the agent phenotype is first randomly generated, with:

$$\mathbf{p}_j = (m'_{\text{IL}}(g_1), \dots, m'_{\text{IL}}(g_l)), \quad (3)$$

and:

$$m'_{\text{IL}}(g_i) = \begin{cases} U(\{0, 1\}) & \text{if } g_i = ? \\ g_i & \text{otherwise,} \end{cases} \quad (4)$$

where $U(S)$ is a random element in the set S . Then, the agent phenotype is set to the best realization in the set:

$$\mathbf{p} = m_{\text{IL}}(\mathbf{g}) = \arg \max_{\mathbf{p}_j \in \{\mathbf{p}_1, \dots, \mathbf{p}_{n_{\text{iter}}}\}} f(\mathbf{p}_j). \quad (5)$$

Social learning (SL) In this variant the agent learns trait-wise by imitating the most common value for each trait among a given collection T of teachers. Formally, given the collection $T = \{\mathbf{p}'_1, \mathbf{p}'_2, \dots\}$ of the phenotypes of the teachers, \mathbf{p} is obtained as:

$$\mathbf{p} = m_{\text{SL}}(\mathbf{g}, T) = (m'_{\text{SL}}(g_1, T), \dots, m'_{\text{SL}}(g_l, T)), \quad (6)$$

with:

$$m'_{\text{SL}}(g_i, T) = \begin{cases} g_i & \text{if } g_i \neq ? \\ \hat{p}_i & \text{otherwise,} \end{cases} \quad (7)$$

where \hat{p}_i is the most common value among the multiset of i -th traits of the teacher phenotypes in T , or $\hat{p}_i = 0$ if $T = \emptyset$.

We explored a few alternatives for selecting the collection T of teachers from a population P , with the aim of answering the first of the two main research questions of this paper, i.e., from whom learners should learn? In each alternative, since learning involves teachers phenotypes, we used the population of the previous generation for composing T : at the first generation no learning occurs, hence the mapping is that of Equations (1) and (2). In one alternative, T is composed of the phenotype of an individual chosen randomly in P : we denote this SL variant with SL-Random. In the other alternatives, T is composed of the phenotypes of the n fittest individuals in P : we denote these SL variants with $\text{SL-}\frac{n}{|P}| \%$ and experimented with a few values for n . Significant cases of $\text{SL-}\frac{n}{|P}| \%$ are the one with $n = 1$, where the learner learns only from the best individual, and the one with $n = |P|$, where the learner learns from all the individuals: we denote these two variants with SL-Best and SL-All, respectively.

Experiments and discussion

We performed several experiments by varying the main parameters of the model presented in the previous section: in particular, we experimented with 3 values for the genotype/phenotype length l , 3 for the population size, and 6 for the number n of teachers in social learning. Table 1 shows the values for all the parameters. In the random initialization of the population, each gene was set to a random allele with probability 0.25, 0.25, and 0.5 respectively for 0, 1, and ? : we set a larger probability for ? to let the learning have a large impact on adaptation.

For each combination of l, n_{pop} , we experimented with each one of the 8 variants (EVO, IL, and 6 SL) for the genotype-phenotype mapping function; for each of the resulting $3 \cdot 3 \cdot 8 = 72$ treatments, we executed 10 evolutionary runs of the EA.

Table 2 presents the results of the experiments in terms of two indexes. In the rightmost half, the table shows the value f_{final}^* of the fittest individual in the last generation. In the leftmost half, it shows the generation $i_{0.9}^*$ at which, during the evolution, the fitness of the best individual first exceeded the value $0.9f_{\text{final}}^*$, i.e., 90% of the fitness of the best individual at the last generation of the same run. Intuitively, the

Table 1: Parameters of the simple scenario.

Parameter	Value
Geno./pheno. length l	100, 500, 1000
Generations n_{gen}	400
Pop. size n_{pop}	100, 500, 1000
Tour. sel. size n_{tour}	5
Mut. prob. p_{mut}	0.01
IL iterations n_{iter}	50
SL n. of teachers n	$1, \frac{n_{\text{pop}}}{10}, \frac{n_{\text{pop}}}{4}, \frac{n_{\text{pop}}}{3}, \frac{n_{\text{pop}}}{2}, n_{\text{pop}} $

former index captures the *effectiveness* of the adaptation and the latter captures the *efficiency*, i.e., how fast a given adaptation degree is achieved. For both indexes, the table shows the mean value across the 10 repetitions.

The foremost finding, although not a straightforward answer to the research questions of this study, that arises from the figures in Table 1 is that SL leads to a more efficient and more effective adaptation than both IL and no learning. This result is substantially in agreement with previous experimental studies on social learning in similar settings, as, e.g., Marriott et al. (2018), where, as here, SL outperforms IL and IL outperforms no learning.

Concerning, more specifically, the first research question of this paper, Table 2 shows that there are some differences, yet not sharp, among different variants of social learning. SL-Random is in general moderately less effective and efficient than the other SL variants. SL-Best appears slightly less efficient than the other SL- $p\%$, with the exception of the case with large population; moreover, it also obtains a final best fitness that is always lower than the other SL- $p\%$. SL-All looks slightly less efficient than the other remaining SL- $p\%$, but no differences are visible in terms of effectiveness. Summarizing, Table 2 suggests that the adaptation is favored when SL occurs with the learner that learns from more than one individuals.

To better understand the results of these experiments, we show in Figure 1 how the fitness f^* of the best individual and the number $|g^*|_?$ of ?s in its genotype vary during the evolution for the case with $n_{\text{pop}} = 100$ and $l = 500$ —the plots for the other combinations of n_{pop}, l are qualitatively similar. The second index measures, intuitively, the genetic proneness of an individual to learn, in brief, individual *learnability*. Since in these experiments the evolution is, in general, fast, Figure 1 shows generation with a logarithmic scale for better highlighting differences in efficiency among variants.

Two observations can be made based on Figure 1. First, the plot of f^* during the evolution confirms the differences in efficiency among the considered variants.

Second and foremost, there is a neat difference in the final number $|g_{\text{final}}^*|_?$ of ?s among the variants. For EVO and IL, $|g_{\text{final}}^*|_?$ is very low, approximately 0 for the former and

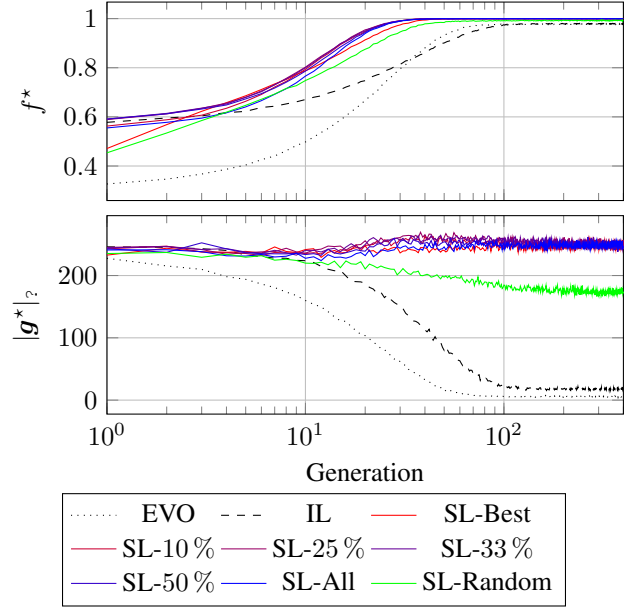


Figure 1: Fitness f^* (top) and number $|g^*|_?$ of ?s in the genotype (bottom) of the best individual during the evolution, for $n_{\text{pop}} = 100$ and $l = 500$ and in the simple scenario.

10 for the latter. For all the SL variants with the exception of SL-Random, $|g_{\text{final}}^*|_? \approx 250$, that is, much larger than EVO and IL. Finally, for SL-Random $|g^*|_?$ tends to slowly decrease during the evolution, reaching a final value of $|g_{\text{final}}^*|_? \approx 175$. Since $|g^*|_?$ represents the individual learnability, this finding means that in this simple scenario evolution favors learnability at the expense of genetic quality, when learning can compensate lack of genetic quality. In other words, there is no evolutionary reason for having a perfect genotype if good traits can be obtained by learning from other individuals. Moreover, the lower value of $|g_{\text{final}}^*|_?$ for SL-Random suggests that the better the teachers, the more this compensation actually applies.

We assessed the statistical significance of the results presented above on $i_{0,9}^*$ and f_{final}^* by means of the Wilcoxon signed-rank test with correction. The analysis outcome concerning $i_{0,9}^*$ suggest that EVO, IL, and SL-Random are respectively significantly different ($p < 0.005$) from all the adaptation strategies. Interestingly also SL-Best is significantly different ($p < 0.007$) from all the other strategies. On the other side there is no significant difference among SL-10%, SL-25%, and SL-33%, and between SL-50% and SL-All. The outcome concerning f_{final}^* confirms the efficiency test results: EVO, IL, and SL-Random are significantly different ($p < 0.008$) from all the adaptation strategies. There is no significant difference among the other SL strategies.

Imperfect social learning In order to better investigate the effect of social learning on evolved learnability, we con-

Table 2: Experimental results with the simple scenario.

Variant	n_{pop}	Adaptation efficiency $i_{0.9}^*$									Adaptation effectiveness f_{final}^*								
		100			500			1000			100			500			1000		
		l	100	500	1000	100	500	1000	100	500	1000	100	500	1000	100	500	1000	100	500
EVO		14	41	100	14	40	59	14	40	59	1.00	0.98	0.89	1.00	0.98	0.93	1.00	0.98	0.93
IL		14	45	75	13	44	64	9	34	52	1.00	0.98	0.91	1.00	0.98	0.94	1.00	0.99	0.95
SL-Random		8	22	50	7	21	32	7	21	33	1.00	0.99	0.91	1.00	0.99	0.97	1.00	0.99	0.97
SL-Best		5	19	47	5	18	29	5	18	29	1.00	1.00	0.93	1.00	1.00	0.98	1.00	1.00	0.98
SL-10 %		5	16	50	4	14	22	7	19	29	1.00	1.00	0.98	1.00	1.00	0.99	1.00	1.00	0.99
SL-25 %		6	16	54	4	15	23	7	20	30	1.00	1.00	0.98	1.00	1.00	0.99	1.00	1.00	0.99
SL-33 %		5	16	57	4	15	23	7	20	31	1.00	1.00	0.98	1.00	1.00	0.99	1.00	1.00	0.99
SL-50 %		6	17	59	4	15	24	7	20	31	1.00	1.00	0.98	1.00	1.00	0.99	1.00	1.00	0.99
SL-All		6	17	64	6	17	26	8	21	32	1.00	1.00	0.98	1.00	1.00	0.99	1.00	1.00	0.99

ducted a further set of experiments with a modified SL variant that represents an *imperfect transfer of knowledge* from the teachers to the learner. In this variant, the genotype-phenotype mapping function is the same of Equations (6) and (7), but, the most common i -th trait \hat{p}_i among the teachers is determined on a subset of the teachers traits $\{p'_1, \dots, p'_{|T|}\}$ obtained by removing each p'_j with a probability 0.5. In other words, in this SL variant, only half of the teachers knowledge concerning this trait is actually transferred. This constitutes a form of imperfect transmission of knowledge, rather than of genetic material Borg et al. (2011).

Figure 2 shows (only for modified SL variants) how f^* and $|g^*|_?$ vary during the evolution for the case with $n_{\text{pop}} = 100$ and $l = 500$, similarly to Figure 1.

According to Figure 2, two consequences of the imperfect transfer of knowledge can be seen. First, there is a larger diversity in efficiency among SL variants, whereas effectiveness remains the same. Second, and more interesting, $|g_{\text{final}}^*|_?$ is no more approximately the same for all the SL variants. In particular, for all the variants where the set T of teachers is composed by more than one individuals, $|g_{\text{final}}^*|_?$ is ≈ 250 , whereas it is ≈ 10 for single-teacher variants SL-Best and SL-Random. In other words, social learning cannot compensate low genetic quality if one can learn only from one teacher with imperfect transfer knowledge: interestingly, this holds regardless the single teacher being the best one or a random individual.

Artificial life scenario

The simple scenario presented in the previous section allows to compare different forms of social learning and to easily interpret the outcome, in terms of interaction of learning and evolution. However, due to its simplicity it is not representative of any concrete realization of a multi-agent scenario. With respect to the aim of the present study, the limitations of the simple scenario are two-fold. First, concerning the

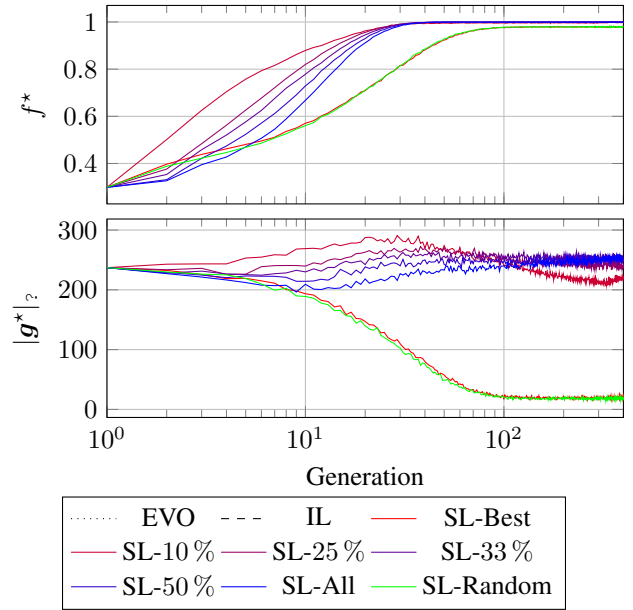


Figure 2: Fitness f^* (top) and number $|g^*|_?$ of ?s in the genotype (bottom) of the best individual during the evolution, for $n_{\text{pop}} = 100$ and $l = 500$ in the simple scenario, only SL variants with the imperfect knowledge transfer.

world model: (a) agents do not interact, nor compete, among themselves and (b) their life-span is instantaneous. Second, concerning the learning mechanism: (c) knowledge can be transferred by directly copying the traits.

In order to extend the generality of our findings, we studied a more complex artificial life scenario. Inspired by Le et al. (2020), we define a 2D-world, discrete in time and space, where agents move in a toroidal environment and strive for survival by collecting food, i.e., non-moving items located in the environment.

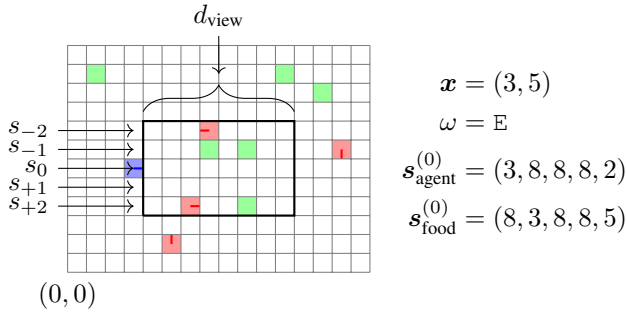


Figure 3: Schematic representation of how the agent (blue square) perceives the environment, i.e., distances to other agents (red squares) and food items (green squares) that are in its view range (thick rectangle). Corresponding values are shown on the right, assuming here $d_{\text{view}} = 7$ and $n_s = 2$.

Agent An agent is defined by its state, genotype, and phenotype: the state may change over the time, the genotype never changes during the agent life, the phenotype may change only once during the agent life. The agent *state* $(\mathbf{x}, \omega, a, e)$ consists of its position $\mathbf{x} \in \{0, \dots, w\}^2$, where w is the world width, its orientation $\omega \in \{\text{N}, \text{E}, \text{S}, \text{W}\}$, its age $a \in \mathbb{N}$, and its energy $e \in \mathbb{N}$.

At each time step the agent takes an action in $A = \{\circlearrowright, \circlearrowleft, \uparrow, \emptyset\}$, i.e., rotate clockwise, rotate counter-clockwise, move ahead, do nothing. The agent determines the action according to a policy (detailed below) that takes as input the agent sensing of the environment $\mathbf{s} = [\mathbf{s}_{\text{agent}}^{(0)} \ \mathbf{s}_{\text{food}}^{(0)} \ \mathbf{s}_{\text{agent}}^{(-1)} \ \mathbf{s}_{\text{food}}^{(-1)}]$, with each $\mathbf{s}_*^{(k)} \in [0, d_{\text{view}}]^{2n_s+1}$. Each sensor reading $s_i \in \mathbf{s}_*^{(k)}$, $i \in \{-n_s, \dots, +n_s\}$ corresponds to a sensor perceiving the row of world positions starting on the i -th position on the agent left-side ($i < 0$), front ($i = 0$), or right-side ($i > 0$) and going in the direction of agent orientation: the reading is $s_i \leq d_{\text{view}}$ if the closest agent (for $\mathbf{s}_{\text{agent}}^{(k)}$) or food item (for $\mathbf{s}_{\text{food}}^{(k)}$) is closer than d_{view} positions to the row first position, or $s_i = d_{\text{view}} + 1$ otherwise. Readings in $\mathbf{s}_*^{(0)}$ correspond to distances at the current time step; readings $\mathbf{s}_*^{(-1)}$ correspond to the previous time step. Figure 3 graphically represents how the agent perceives the environment.

The policy of the agent is an artificial neural network (ANN) with $4(2n_s + 1)$ inputs, 4 outputs, no hidden layer, and the Rectifier Linear Unit (ReLU) as activation function. Once the input layer is set to \mathbf{s} , the agent takes the action of A corresponding to the largest network output. The ANN parameters $\theta = \mathbf{p} \in [-1, 1]^l$, with $l = 4(4(2n_s + 1) + 1)$, constitute the *phenotype* of the agent.

The agent *genotype* is a vector $\mathbf{g} = [\mathbf{g}_w \ \mathbf{g}_t]$, with $\mathbf{g}_w \in [-1, 1]^l$ and $\mathbf{g}_t \in \{0, 1\}^l$: element of \mathbf{g}_t resemble the function of question marks in the simple scenario. The genotype determines the phenotype, possibly as the result of a learn-

ing procedure, as detailed below.

World evolution At the beginning of a simulation, $n_{\text{agent}}^{\text{init}}$ agents and $n_{\text{food}}^{\text{init}}$ food items are randomly placed in the environment, i.e., a $w \times w$ grid. Each agents initial age and energy are set to 0 and e^{init} , respectively; genotype \mathbf{g} is set randomly in the proper domain and phenotype is initially set to $\mathbf{p} = \mathbf{g}_w$.

On the short-term life-time time scale, the world evolves because agents move and consume food. At each time step each agent takes an action based on its policy and sensing. The agent results in a change of the agent state at the next time step: besides the trivial change in position \mathbf{x} and orientation ω , if at the updated position there is a food item, the agent energy e is incremented by 10 and the food item is removed from the world. Moreover, regardless of the action, age a is incremented by 1 and energy e is decremented by 1.

In order to conserve the overall energy in the world, a new food item is added if:

$$10n_{\text{food}}^{(k)} + \sum_{j=1}^{j=n_{\text{agent}}^{(k)}} e_j^{(k)} < 10n_{\text{food}}^{\text{init}} + \sum_{j=1}^{j=n_{\text{agent}}^{\text{init}}} e_j^{\text{init}}, \quad (8)$$

where $n_{\text{food}}^{(k)}$ and $n_{\text{agent}}^{(k)}$ are the current number of food items and agents and $e_j^{(k)}$ is the energy of the j -th agent.

On the long-term evolutionary time scale, the world evolves because agents die and are born, as follows. If an agent energy becomes negative, the agent dies and is removed from the world. If an agent energy exceeds a threshold e_{repr} , the agent (parent) generates a new agent (child), i.e., it asexually reproduces (as in Annunziato and Pierucci (2003)). Upon reproduction, the parent energy is decreased by $\frac{1}{2}e^{\text{repr}}$. The child position is set to a random position adjacent to the parent position, its age is set to 0, and its energy to $\frac{1}{2}e^{\text{repr}}$. The child genotype $\mathbf{g}' = [\mathbf{g}'_w \ \mathbf{g}'_t]$ is a mutated copy of the parent genotype: \mathbf{g}'_w is obtained by applying Gaussian mutation to \mathbf{g}_w , with a parameter σ_{mut} , and \mathbf{g}'_t by applying bit-flip mutation to \mathbf{g}_t , with a probability of p_{mut} . The child phenotype is initially set to $\mathbf{p}' = \mathbf{g}'_w$.

As in the simple scenario, we consider a few variants of the model here described that differ in the social learning mechanism, i.e., in how and when the agent phenotype is updated during its life. Since we are interested mainly in social learning, in this second scenario we do not consider a variant corresponding to individual learning.

No learning (EVO) In this variant, no learning occurs. The agent phenotype is never updated and remains the one set at the agent birth, i.e., $\mathbf{p} = \mathbf{g}_w$.

Social learning by direct transfer (SLD) This variant resembles the SL that we defined in the simple scenario. Given a collection T of teachers, the phenotype \mathbf{p} of the learning

agent is set to:

$$\begin{aligned} \mathbf{p} &= m_{\text{SLD}}(\mathbf{g}, T) \\ &= (m'_{\text{SLD}}(g_{w,1}, g_{t,1}, T), \dots, m'_{\text{SLD}}(g_{w,l}, g_{t,l}, T)), \end{aligned} \quad (9)$$

with:

$$m'_{\text{SLD}}(g_{w,i}, g_{t,i}, T) = \begin{cases} g_{w,i} & \text{if } g_{t,i} = 0 \\ \bar{p}_i & \text{otherwise,} \end{cases} \quad (10)$$

where \bar{p}_i is the mean value of the i -th elements of teacher phenotypes in T . We remark that all agents have the same ANN architecture.

Based on the experimental findings of the simple scenario, we considered here only one way for constructing T from the current population. We set T to the subset of the 25% agents with the largest ratio f between the number of collected food items $\frac{1}{10}(e - a)$, since birth, and age a . In other words, f is the *fitness* of the agent and the teachers are the current fittest agents.

We make this form of learning occur just upon the agent birth.

Social learning by experience sharing (SLE) In this variant there is no direct transfer of ANN parameters. Instead the learner updates its parameters to match the behavior of the teachers on a set of cases. To this end, each agent state includes also a set S_{exp} of up to n_{exp} sensor readings s , with $s \in [0, d_{\text{view}}]^{4(2n_s+1)}$, that represents the *experience* of the agent in terms of sensor readings. At each time step, the current s is added to S_{exp} with probability 0.5: then, if $|S_{\text{exp}}| > n_{\text{exp}}$ one random item is removed from S_{exp} .

Given a collection T of teachers and a sequence S of sensor readings, the learning works as follows. First, for each input $s_i \in S$, (i) the action $z_{i,j}$ of each j -th teacher is determined by applying the corresponding ANN on s_i , hence obtaining a sequence Z_i , then (ii) the most common output $\hat{z}_i \in Z_i$ is obtained and associated to s_i . This step results in a sequence E of pairs (s, \hat{z}) . Second, the learning agent modifies its \mathbf{p} , i.e., its ANN parameters, by performing a backpropagation on E , where the target ANN output is set by one-hot encoding each \hat{z} . We perform backpropagation with learning rate 0.0001, 100 iterations, and with mean square error on the network output as cost function.

We explored two variants of the SLE procedure here described by varying the way the collection S is built: (a) from the learner experience S_{exp} , by selecting a random subset of n_{learn} items, or (b) from the overall teachers experience $\bigcup_T S_{\text{exp}}$ by selecting a random subset of n_{learn} items. In both cases, T is constructed as in SLD. We denote the two cases as SLE-L and SLE-T, respectively.

Since the learning agent needs to collect some experience before being able to apply SLE-L, in this variant the agent performs learning as soon as its experience S_{exp} reaches the size n_{exp} (i.e., on average at $2n_{\text{exp}}$ age). Differently, SLE-L occurs upon birth, as for SLD.

Table 3: Parameters of the artificial life scenario.

Parameter	Value	Parameter	Value
Word width w	500	Repr. energy e_{repr}	10
Time steps n_{steps}	10^5	Energy th. e_{thr}	10 000
N. of runs n_{runs}	10	Pop. size $n_{\text{agents}}^{\text{init}}$	1000
Dist. of view d_{view}	10	N. of teachers $ T $	$\frac{n_{\text{agent}}}{4}$
N. of sensors n_s	7	Learning exp. n_{exp}	1000

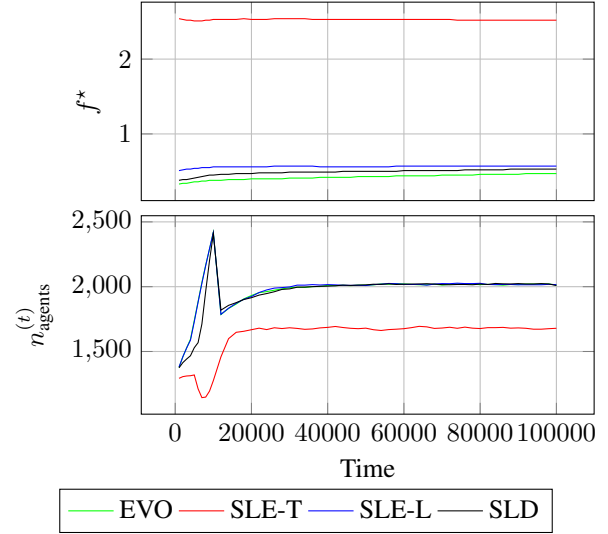


Figure 4: Fitness f^* of the best individual (top) and population size (bottom) during the simulation of different adaptation strategies in the artificial life scenario.

Experiments and discussion

We performed n_{runs} simulations of n_{steps} time steps for each one of the four variants (EVO, SLD, SLE-L, and SLE-T) with the parameters of Table 3. Figure 4 shows the results of the experiments for all the variants, averaged on n_{runs} simulations. We plot the best individual fitness f^* (above) and the population size $n_{\text{agents}}^{(k)}$ where the 2 plots refers to respectively to the best individual fitness, and the number of agents, over the simulations time steps.

From the best individual fitness plot in Figure 4 it can be seen that SLE-T results in a larger best fitness f^* . Among the remaining variants, SLE-L is slightly more efficient than EVO and SLD, but equally effective, since all these variants converge to the best individual fitness $f^* = 0.5$. On the other hand, EVO, SLE-L, and SLD result in larger populations, whose size reaches ≈ 2000 individuals after 30 000 time steps. For these three variants there is an initial spike in the population size, that we motivate as follows: (a) the initial energy of the agents is almost high enough to allow them to reproduce (b) the initial abundance of food allows the agent to collect food by just performing small movements in the surrounding environment (c) most of the first

generations agents have a policy that is not suitable for survival, thus reducing competition for the few active ones.

Differently the number of agents in the population of SLE-T increases after 7000 time steps, which corresponds to the peak of the spike in the other variants population plot, converging to populations of 1700 agents for the remaining time steps. The convergence observed in all the variants means that agents are adapting to survive, and that the children agents are good enough to compete with their parents, thus increasing the competition for the food. In particular, it seems that the agents adapting through SLE-T, are the best performing ones in terms of best individual fitness, with maximum value $f^* = 2.71$, but also in terms of efficiency, since in all the runs the highest individual fitness is reached at the beginning of the simulation.

This experimental findings provide an intuitive answer to our research question: it is more profitable to choose the learning samples from the teachers experience, rather than from the learners one.

Concluding remarks and future work

In this work, we investigated two relevant aspects of social learning in artificial life: from which teachers to learn and which is the best way to transfer knowledge from teachers to learners. To answer these questions, we assessed the effectiveness of social learning, when it is interacting with evolution, in two different scenarios: a simple one and a more complex one. In the former, we varied the composition of teachers set, whereas in the latter we experimented with different knowledge transfer mechanisms. Experimental results in the first scenario show a higher benefit when teaching is performed by a set of good teachers rather than by a single excellent one. As a possible explanation, learning from multiple sources rather than just one favors generality of learned traits. Concerning the second scenario, results highlight how learning improves when what to teach is decided by the teachers rather than by learner.

We believe that our results are promising and that our findings can help to better understand the potential and implications of social learning. As future work, we plan to investigate the case where knowledge transfer can occur only among individuals that are physically close and, hence, how mobility of agents shapes social learning, possibly favoring mobility-related skills more than other skills.

References

- Ackley, D. and Littman, M. (1991). Interactions between learning and evolution. *Artificial life II*, 10:487–509.
- Annunziato, M. and Pierucci, P. (2003). The emergence of social learning in artificial societies. In *Workshops on Applications of Evolutionary Computation*, pages 467–478. Springer.
- Baldwin, J. M. (1896). A new factor in evolution. *The american naturalist*, 30(354):441–451.
- Borenstein, E., Feldman, M. W., and Aoki, K. (2008). Evolution of learning in fluctuating environments: when selection favors both social and exploratory individual learning. *Evolution*, 62(3):586–602.
- Borg, J. M., Channon, A., Day, C., et al. (2011). Discovering and maintaining behaviours inaccessible to incremental genetic evolution through transcription errors and cultural transmission. In *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*, pages 101–108. MIT Press.
- Bullinaria, J. A. (2017). Imitative and direct learning as interacting factors in life history evolution. *Artificial Life*, 23(3):374–405.
- Chebib, J. and Marriott, C. (2016). Modeling the evolution of gene-culture divergence. In *Proceedings of the Artificial Life Conference 2016 13*, pages 500–507. MIT Press.
- Feldman, M. W., Aoki, K., and Kumm, J. (1996). Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropological Science*, 104(3):209–231.
- Floreano, D., Dürr, P., and Mattiussi, C. (2008). Neuroevolution: from architectures to learning. *Evolutionary intelligence*, 1(1):47–62.
- Heinerman, J., Bussmann, B., Groenendijk, R., Van Krieken, E., Slik, J., Tezza, A., Haasdijk, E., and Eiben, A. (2018). Benefits of social learning in physical robots. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 851–858. IEEE.
- Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biological Reviews*, 69(2):207–231.
- Hinton, G. E. and Nowlan, S. J. (1987). How learning can guide evolution. *Complex systems*, 1(3):495–502.
- Jolley, B. P., Borg, J. M., and Channon, A. (2016). Analysis of social learning strategies when discovering and maintaining behaviours inaccessible to incremental genetic evolution. In *International Conference on Simulation of Adaptive Behavior*, pages 293–304. Springer.
- Keesing, R. and Stork, D. G. (1991). Evolution and learning in neural networks: the number and distribution of learning trials affect the rate of evolution. In *Advances in Neural Information Processing Systems*, pages 804–810.
- Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior*, 32(1):4–14.
- Le, N. (2019a). Evolution and self-teaching in neural networks: another comparison when the agent is more primitively conscious. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 2040–2043.
- Le, N. (2019b). Organic selection and social heredity: The original baldwin effect revisited. In *The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, pages 515–522. MIT Press.

- Le, N., Brabazon, A., and O'Neill, M. (2020). Social learning vs self-teaching in a multi-agent neural network system. In *23rd International Conference on the Applications of Evolutionary Computation*.
- Le, N., O'Neill, M., and Brabazon, A. (2018). The baldwin effect reconsidered through the prism of social learning. In *2018 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE.
- Marriott, C., Borg, J. M., Andras, P., and Smaldino, P. E. (2018). Social learning and cultural evolution in artificial life. *Artificial life*, 24(1):5–9.
- Marriott, C. and Chebib, J. (2014). The effect of social learning on individual learning. In *Artificial Life Conference Proceedings 14*, pages 736–743. MIT Press.
- Mesoudi, A., Chang, L., Dall, S. R., and Thornton, A. (2016). The evolution of individual and cultural variation in social learning. *Trends in ecology & evolution*, 31(3):215–225.
- Nolfi, S., Parisi, D., and Elman, J. L. (1994). Learning and evolution in neural networks. *Adaptive Behavior*, 3(1):5–28.
- Soltoggio, A., Stanley, K. O., and Risi, S. (2018). Born to learn: the inspiration, progress, and future of evolved plastic artificial neural networks. *Neural Networks*, 108:48–67.
- van der Post, D. J., Franz, M., and Laland, K. N. (2016). Skill learning and the evolution of social learning mechanisms. *BMC evolutionary biology*, 16(1):166.
- van Schaik, C. P. and Burkart, J. M. (2011). Social learning and evolution: the cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):1008–1016.
- Wakano, J. Y., Aoki, K., and Feldman, M. W. (2004). Evolution of social learning: a mathematical analysis. *Theoretical population biology*, 66(3):249–258.
- Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE*, 87(9):1423–1447.