Simulated Trust: Towards robust social learning

Dieter Vanderelst[1], René Ahn[2] and Emilia Barakova[2]

[1]University of Antwerp, Belgium, [2]Einhoven University of Technology, The Netherlands
dieter_vanderelst@emailengine.org

Abstract

Social learning is a potentially powerful learning mechanism to use in artificial multi-agent systems. However, findings about how animals use social learning show that it is also possibly detrimental. By using social learning agents act based on second-hand information that might not be trustworthy. This can lead to the spread of maladaptive behavior throughout populations. Animals employ a number of strategies to selectively use social learning only when appropriate. This suggests that artificial agents could learn more successfully if they are able to strike the appropriate balance between social and individual learning. In this paper, we propose a simple mechanism that regulates the extent to which agents rely on social learning. Our agents can vary the amount of trust they have in others. The trust is not determined by the performance of others but depends exclusively on the agents' own rating of the demonstrations. The effectiveness of this mechanism is examined through a series of simulations. We first show that there are various circumstances under which the performance of multi-agents systems is indeed seriously hampered when agents rely on indiscriminate social learning. We then investigate how agents that incorporate the proposed trust mechanism fare under the same circumstances. Our simulations indicate that the mechanism is quite effective in regulating the extent to which agents rely on social learning. It causes considerable improvements in the learning rate, and can, under some circumstances, even improve the eventual performance of the agents. Finally, some possible extensions of the proposed mechanism are being discussed.

The Ecology Of Social Learning

Throughout the animal kingdom individuals exploit information that has been gathered by others. Animals from invertebrates (reviewed in Leadbeater and Chittka, 2007; Leadbeater et al., 2006; Fiorito, 2001) to great apes and humans (e.g. Tomasello, 1999; Whiten et al., 2007; Bonnie et al., 2006) exhibit forms of social learning[1]. The widespread use of social learning among taxa is caused by its enormous ecological advantages in many circumstances (see for example Kendal et al., 2005; Coolen et al., 2005; Bonnie and Earley, 2007, and references therein). Evolution favored social learning because it might allow individuals to be flexible and adaptive learners while avoiding the dangers associated with individual exploration (Boyd and Richardson, 1988; Zentall, 2006). Ecologists typically stress the fact that individuals benefit from copying behavior from others because it saves them the costs of asocial learning (Laland, 2004). Indeed, Zentall (2006) remarked that the behavior of others has often already been shaped by its consequences and might therefore be assumed to be safe to copy.

Unsurprisingly, social learning comes in many flavors. Various forms of social learning have been identified (Zentall, 2006) and the underlying mechanisms range from fairly simple to utterly complex (Noble and Todd, 2002). However, when studying the dynamics and ecological properties of social learning one can ignore the differences in implementations and consider underlying exchange of information only (Coussi-Korbell and Fragaszy, 1995). This made it possible to evaluate the advantages of social learning in theoretical studies focusing on the game-theoretic aspects.

This line of theoretical research, supported by empirical findings in animal behavior, has shown that the advantage of social learning is by no means universal. Social learning is advantageous only if one takes certain precautions (Laland, 2004; Galef and Laland, 2005). The fundamental problem is that social learning can support the spread, acquisition and the persistence of maladaptive behavior (Giraldeau et al., 2002). This is because social learners re-use information gathered by others but do not collect new information themselves. Therefore, they are implicitly assuming that the information they gather from others is reliable. There are several circumstances under which this assumption does not hold (see Giraldeau et al., 2002; Laland, 2004; Leadbeater and Chittka, 2007, for reviews and references). Second hand information can be, amomg others, incomplete, outdated, biased or utterly wrong.

Instead of animals relying on social learning whenever they can, evidence clearly shows that they are somewhat reluctant to use social information unless there is a good reason to do so (Galef and Laland 2005, see Laland et al. 2005 for a short discussion of a striking example in sticklebacks).

[1]Here, on theoretical grounds, taken to include the use of public information. See Bonnie and Earley (2007) for a discussion.

Animals (including humans, see Koenig and Harris 2005) employ certain selection strategies to control the copying of behavior. This allows them to use social learning in an intelligent fashion, avoiding its potential pitfalls. Because of this, examples in which social learning leads to maladaptive behavior are rather scarce in the literature on animal behavior. The most clear examples of social learning supporting maladaptive behavior are obtained under experimental circumstances where the, usual adequate, strategies fail (e.g. Laland and Williams, 1998; Pongrcza et al., 2003). Such experiments can uncover the strategies adopted by animals.

Laland (2004) found that guppies were induced to take a longer, less efficient, route to a feeding site if others were doing this also. In contrast, single guppies learned quickly to take a shorter route. In the context of the experiment the socially transferred behavior was clearly maladaptive. However, in natural circumstances, choosing the same route as others is a good strategy since it protects against predation by forming shoals. The guppies' strategy to conform leads them to adopt longer routes but this is clearly an advantageous strategy if considered in the ecological context in which it evolved.

Opposed to the conformity bias observed in guppies, some experimental results show a selective use of social learning. Capuchin monkeys do not resort to social learning if the problems they are challenged with (e.g. opening a box) are easy to solve. In contrast, when faced with a difficult task they will copy the behavior of others more frequently (see Laland, 2004, for references and a discussion). Presumably, monkeys are more willing to use the potentially flawed social information if asocial learning is costly. This shows that these animals do not assume a priori that social information is correct and reliable (and thus worthwhile to copy). Instead they adopt a trade-off between learning socially and individually taking into account possible costs and gains. See Laland (2004) for more examples of selective social learning in animals.

While the literature on animals shows relatively few instances of maladaptive social learning under natural circumstances, humans, who rely far more on social learning (Tomasello, 1999) than any other animal, provide many more examples (Boyd and Richerson, 2006). Obvious candidates are the social transfer of tobacco and drug use, reducing fertility and endangering fetus development. But also other, less dramatic, socially transferred behavior could reduce fitness in humans.

Social Learning in Artificial Agents

Recently, different authors have begun to explore the use of social learning as a way of instruction in artificial multi-agent systems (e.g. Acerbi et al., 2007; Pini et al., 2007; Belpaeme et al., 2007; Noble and Franks, 2002; Alissandrakis et al., 2004).

In a multi-agent setting, artificial agents could search concurrently for a solution for a given problem (e.g. how to pick up food). Once a single agent has found a solution, this innovation could be copied by others and could propagate through the population. In this way, social learning could drastically reduce the total number of learning trials needed for a population of artificial agents to solve a problem (Pini et al., 2007). Innovations in groups of animals are known to spread in the same way (e.g. Bonnie et al., 2006; Bonnie and Earley, 2007; Leadbeater and Chittka, 2007).

Though this argument rightfully assumes that social learning has attractive properties, it was also argued above that this is certainly not true in all circumstances. In fact, as said, animal behavior data suggest that social learning should be only engaged in sparsely and with great caution (Laland, 2004; Galef and Laland, 2005; Leadbeater and Chittka, 2007).

Reasoning by analogy, we can hypothesize that a successful learning strategy for artificial agents should strike a careful balance between different types of learning. This suggests that the learning performance of artificial agents can be improved by mechanisms that restrict social learning to circumstances under which it is appropriate.

Experimental Setup

We have investigated the question how agents can balance social and individual learning by simulating a very simple world with a number of agents. The agents in this world have been equipped with a mechanism that regulates the extent to which they rely on social learning. The fundamental risk in social learning is to act on untrustworthy information. Therefore, we equip agents with the possibility to change the level of trust they have in the demonstrations of others. This in turn determines their reliance on social learning.

We investigate the learning behavior of the agents by comparing their performance in simulations for various conditions. In all conditions we consider two populations of agents that have the same cognitive architecture. The first population is born before the second one, and has therefore already acquired some level of experience in the simulated world when the second population is initiated. The experimental conditions modeled differ in two important respects: (1) the protective trust mechanism employed and (2) whether both populations must learn the same task, or different tasks.

The Trust Mechanism

All agents have the same cognitive architecture (schematically represented in figure 1) and operate in a world in which a limited number of percepts (situations) can arise. Agents can respond to each percept using one of limited set of actions. Once this action is performed, the world returns a reward to the agent. The agents learn both individually and socially which action to perform in response to each percept.

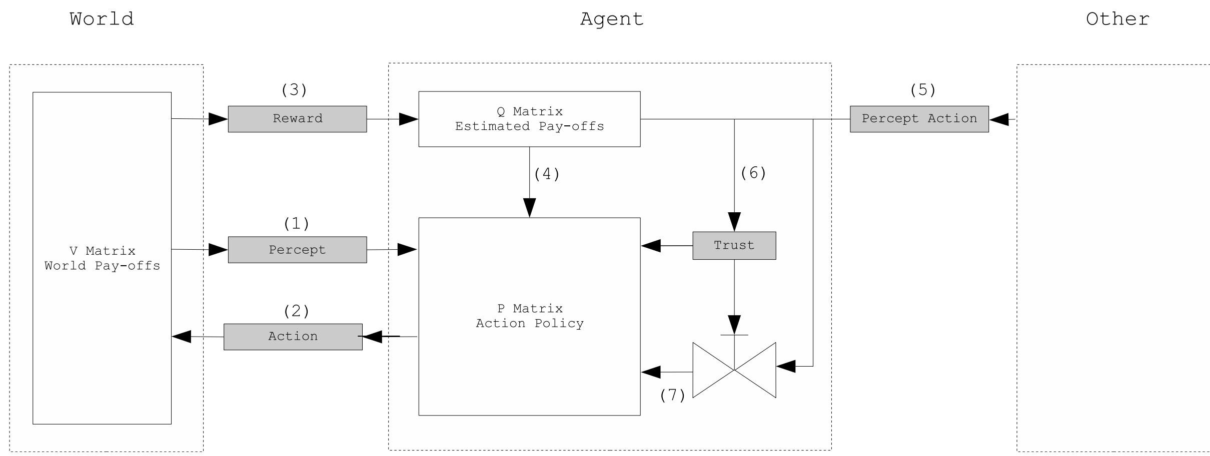


Figure 1: The cognitive architecture of the agents in the simulations and their relationship with the environment.

The behavior of our agents can be captured by a few simple rules. When learning individually this is what happens (the numbers correspond to the ones in figure 1):

- An $Agent$ is confronted (1) with a randomly chosen percept p.
- The $Agent$ chooses (2) an action a with which to respond to the percept p based on its policy **P**. The policy **P** defines the probability of an action a given a percept p.
- The world responds (3) to this action with the appropriate reward as given by the world pay-off function **V** .
- Based on this returned (3) reward, the estimated pay-off Q_{pa} for choosing the action given the percept is adapted.
- The $Agent$ updates (4) its policy **P**, effecting incremental changes to the probabilities for the various actions given the percept p, based on the changed estimates of the pay-offs.

When learning socially, this sequence of events take place:

- An $Agent$ observes (5) what an other agent perceives (percept) and how it reacts (action).
- Based on its own estimated pay-offs **Q** for the given percept, the $Agent$ updates (6) its trust in the observed other.
- The $Agent$ updates (7) its policy **P** for the given percept *dependent* on the trust it has in the other.

So, while learning individually, the agent builds an estimate Q_{pa} of the rewards obtained by executing each of the actions a when confronted with a percept p. This estimate determines its action policy.

During social learning, the agent copies the behavior of other agents whose actions it observes. However, the *extent* to which the behavior of others influences the agents' own policy, depends on the level of trust. The more an agent trusts the other, the more given observations will change its action policy **P**. Therefore, the trust level, which changes over time, regulates the extent to which agents rely on social learning.

Our agents increase the trust they have in others if the perceived behavior is in line with their own estimates of the rewards. If an agent perceives another responding to a percept with an action which itself thinks to be rewarding, the level of trust will rise. So, the more an agent sees others perform according to what it itself thinks is a rewarding policy, the more it will trust and copy them.

Simulations

Methods: Agents & World

In this section we describe the algorithm and settings of the simulations in detail.

The simulated world contains a fixed number of possible percepts. Agents can select one of small number of actions to respond to a given percept.

When an agent is confronted with a certain percept p it performs an action a. Subsequently, it receives a reward from the world. This reward is given a value V_{pa} stored in a matrix **V**. The values V_{pa} characterize the properties of the interaction between the agents and the world. In the present simulations, for any given percept p only one of the values $V_{p.}$ is set to 1 (see table 1 for examples). The others are set to -1. The action a for which $V_{pa} = 1$ determines which action an agent should perform when observing the percept p.

Each artificial agent has the same cognitive architecture (schematically represented in figure 1). These are the three central structures:

- a matrix **P** containing the current action policy,

- a matrix **Q** containing the pay-offs as estimated by the agent,
- a value T reflecting the trust level of the agent in others.

The matrix **P** gives, for each percept p and action a, the chance of an agent choosing this action a when confronted with this percept p (see equation 1). Agents are supposed to learn the optimal policy **P** that goes with the rewards as specified by matrix **V**. Matrix **P** is initialized with random values between 0 and 1 with the constraint that each row must sum to 1.

The matrix **Q** contains an estimate of the matrix **V** that is progressively constructed by the agent over the course of a simulation. This matrix is initialized containing only zeros.

The level of trust T of an agent is given by a value between 0 and 1. At the start of the simulation T is 1 which signifies that initial trust is total[2].

$$\mathbf{P} = \begin{pmatrix} P(a_1|p_1) & \dots & P(a_n|p_1) \\ \dots & \dots & \dots \\ P(a_1|p_m) & \dots & P(a_n|p_m) \end{pmatrix} \tag{1}$$

$$\mathbf{Q} = \begin{pmatrix} Q_{p_1a_1} & \dots & Q_{p_1a_n} \\ \dots & \dots & \dots \\ Q_{p_ma_1} & \dots & Q_{p_ma_n} \end{pmatrix} \tag{2}$$

In these simulations time is represented by an integer. At each time tick all agents are updated one by one (in a random order). In each cycle of the model each agent performs a single individual learning trial and may perform several social learning trials. This reflects the assumption that social learning is cheaper than individual learning. In the presented simulations, social learning does not restrict an agent's opportunity to learn individually. This will capture most biological (see Laland, 2004, for a discussion) and artificial situations to a certain extent. So, in our simulations, social learning is modeled as an additional learning method besides individual learning.

At each tick of the model all agents learn individually. Each agent is presented with a random percept. The agent selects one of the possible actions to respond to the percept. The chance $P(a|p)$ is given by the agent's matrix **P**.

After selecting an action p the world returns a reward V_{pa}. Based on this reward, the estimated pay-off Q_{pa} is updated. The update is governed by equation (3). In this equation α_Q is a step size parameter for updating the estimated pay-off matrix **Q**.

$$\Delta Q_{pa} = \alpha_Q(V_{pa} - Q_{pa}) \tag{3}$$

After updating the estimated pay-off matrix **Q**, the policy **P** is updated according to equation (4). The parameter α_I is the individual learning speed. Equation (4) augments the chance of picking action a given a percept p for which the estimated pay-off is currently the largest. Of course, it also decreases the chance of picking any of the other actions. This form of updating action policies is known in the literature on reinforcement learning as pursuit learning (Sutton and Barto, 1998).

$$\begin{cases} for\ a\ = arg\ max_a\ Q_{pa} : \Delta P(a|p) = \alpha_I(1 - P(a|p)), \\ \qquad\qquad \forall a' \neq a : \Delta P(a'|p) = \alpha_I(0 - P(a'|p)). \end{cases} \tag{4}$$

After updating its policy **P**, an agent stores p and a for later consultation by other agents during social learning.

After all agents have learned individually, all agents may perform several social learning trials. Whether they do so or not depends on the specific settings of the simulation (see later).

To learn socially, an agent randomly selects an agent to learn from and consults its latest action and percept. Social learning is modeled as a two-stage process. First, the observing agent updates its trust level. It consults the action a' and the percept p' stored by the observed other. The trust is updated based on the agents **Q** and **P** matrices according to equation (5). In this equation α_T is the step size for updating T. The trust values are constrained to lie between 0 and 1.

$$\Delta T = \begin{cases} -\alpha_T & \text{if } \sum_a [P(a|p) \times Q_{pa}] \leq Q_{p'a'}, \\ \alpha_T & \text{if } \sum_a [P(a|p) \times Q_{pa}] > Q_{p'a'}. \end{cases} \tag{5}$$

Second, after updating the trust level, the observing agent updates its value P_{pa} according to equation (6) with T denoting the trust level the agent has in the observed other. The parameter α_S is the step size governing social learning.

$$\begin{cases} for\ a\ : \Delta P(a|p) = \alpha_S \times T \times (1 - P(a|p)), \\ \forall a' \neq a : \Delta P(a'|p) = \alpha_S \times T \times (0 - P(a'|p)). \end{cases} \tag{6}$$

Selecting an agent to learn from socially is done in the following way. Each agent randomly selects a single agent to learn from and social learning is done as specified above. This is repeated 20 times. An agent can by chance choose an agent to learn from that it has choosen in a previous repetition. However, note that the behavior of this agent might have somewhat changed in the meantime because it has learned socially as well.

As experimenters we evaluate an agents policy by calculating the expected performance E according to equation (7).

$$E_i = \sum_p \sum_a P(a|p) \times V_{pa} \tag{7}$$

Note that in the current simulations no influence of the spatial distribution of the agents was incorporated.

[2]This is by no means essential for the behavior of the model. Results similar to the ones reported in the next section, were obtained by setting T, where appropriate, initially to 0.

	Actions							
	Values 1 ($\mathbf{V}_1$)				Values 2 ($\mathbf{V}_2$)			
	Actions				Actions			
Percepts	1	2	3	4	1	2	3	4
1	**1**	-1	-1	-1	-1	-1	-1	**1**
2	-1	**1**	-1	-1	-1	-1	**1**	-1
3	-1	-1	**1**	-1	-1	**1**	1	-1
4	-1	-1	-1	**1**	**1**	-1	-1	-1

Table 1: The two $\mathbf{V}_{pa}$ matrices in the form of tables used in the reported simulations.

Experimental Simulations

We ran various simulations to explore the properties of the model and to investigate whether and under what circumstances the trust mechanism protects learners against acquiring faulty information.

In these simulations, the agents are required to learn the profitable policies in a world where there are 4 possible percepts with 4 possible actions each. For each percept only one action has a good outcome ($V_{pa} = 1$).

Because we also want to experiment with situations where different populations need to perform different tasks, we need to define two different types of interaction with the world. This is done through two different reward matrices V_{pa} as given in Table 1.

All simulations are run for 200 time ticks. The simulations consist of two stages. First, an initial population (Population 1) of 21 agents is trained. After 50 ticks, another 21 agents (Population 2) are added to the population.

In each simulation Population 1 and 2 can either learn socially ($\alpha_S = 0.1$), individually ($\alpha_I = 0.1, \alpha_Q = 0.1$) or both ($\alpha_S = 0.1, \alpha_I = 0.1, \alpha_Q = 0.1$). If a learning strategy is not being used, the corresponding learning rate α is set to 0. If agents use social learning, they select 20 agents to learn from. This means that these agents have 20 social learning opportunities for each individual learning opportunity.

Simulations also differ with respect to the update of the trust value. Trust values could either be updated ($\alpha_T = 0.1$) or not ($\alpha_T = 0$).

An overview of the settings of the simulations can be found in table 2.

Simulation Results

The results of some of our simulations are plotted in figure 2. Figure 2(a) shows how performance changes over time, while figure 2(b) gives an insight into the development of trust values (where appropriate).

First, we want to demonstrate that our setting is indeed one where social learning can be advantageous. To this end we have simulated a situation where two populations need to perform the same task. Population 1 only learns individually, and population 2 also learns socially. The result are shown in simulation 1 of figure 2(a). As we can see from these results, when Population 2 is introduced into a population of reasonably instructed agents, social learning allows it to quickly catch up with them. Population 2 learns more rapidly than Population 1 by using both individual and social learning and catches up with them in about 20 time ticks.

In Simulation 2 we consider a situation where both populations learn individually and socially. This simulation shows that social learning is not advantageous under all circumstances. At the beginning of the simulation, the performance of Population 1 is actually hampered by the use of social learning. Population 1 learns slower in simulation 2 (with social learning) than in simulation 1 (without social learning). The reason for this is off course that, in simulation 2, the social learning process is also copying erroneous information.

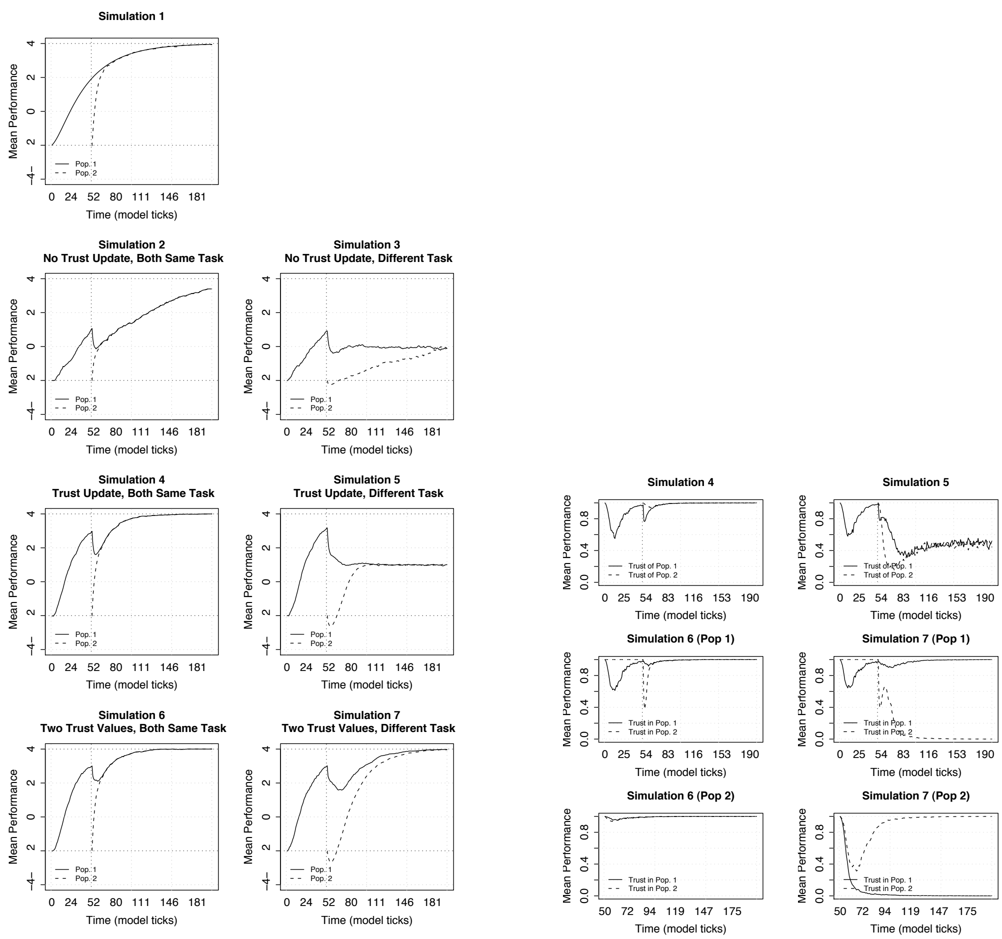
In simulation 3 we considered a situation where population 1 and population 2 have to learn different policies. As is to be expected, here the learning performance is even worse than in simulation 2. After the introduction of Population 2, the performance of Population 2 is actually decreased because it copies the flawed demonstrations of Population 1. Also, in contrast to what happens in simulation 2, Population 1 is now unable to regain its original level of performance because the more population 2 learns, the higher its faulty influence. In the end, the behavior of the two populations converges to a trade-off between the two optimal policies which is optimal for neither of them. The cause for the suboptimal performance in simulations 2 and 3 is that agents copy others even when these are not performing very well or even when they demonstrate a faulty policy. This is exactly what the trust mechanism is supposed to prevent.

The remaining simulations do incorporate different versions of the proposed trust mechanism. In simulation 4, we have copied the situation of simulation 2, but now both populations have a trust mechanism. As can be seen the mechanism is clearly advantageous to both populations.

To better understand what happens here, we plotted the dynamic behavior of the trust values in figure 2(b). Initially, Population 1 is trusting others (of the same population). However, agents quickly discover that demonstrations are not trustworthy. They respond by decreasing their trust level a bit (from 1.0 to 0.6). This allows agents to attain a performance level, through individual learning, at which their demonstrations are accurate enough to be trusted again. After about 20 ticks, the trust levels of the agents start to rise again re-enabling social learning to its full extent. A similar sequence of events is repeated at the introduction of Population 2. The trust levels are reduced to about 0.8 after which they rise again to 1.0. The trust mechanism makes sure that agents perform well enough before they start relying on social learning (again). This causes social learning to be used only if adequate. This significantly increases the learning speed (compare Population 1 in Simulation 1 and 4).

	Population 1				Population 2			
Simulation	Ind.	Soc.	Trust Update	**V**	Ind.	Soc.	Trust Update	**V**
Simulation 1	Yes	No	No	$\mathbf{V}_1$	Yes	Yes	No	$\mathbf{V}_1$
Simulation 2	Yes	Yes	No	$\mathbf{V}_1$	Yes	Yes	No	$\mathbf{V}_1$
Simulation 3	Yes	Yes	No	$\mathbf{V}_1$	Yes	Yes	No	$\mathbf{V}_2$
Simulation 4	Yes	Yes	Yes	$\mathbf{V}_1$	Yes	Yes	Yes	$\mathbf{V}_1$
Simulation 5	Yes	Yes	Yes	$\mathbf{V}_1$	Yes	Yes	Yes	$\mathbf{V}_2$
Simulation 6	Yes	Yes	Yes	$\mathbf{V}_1$	Yes	Yes	Yes	$\mathbf{V}_1$
Simulation 7	Yes	Yes	Yes*	$\mathbf{V}_1$	Yes	Yes	Yes*	$\mathbf{V}_2$

Table 2: The parameter settings in the seven simulations. When social or individual learning is used by a population in a given simulation the corresponding learning rate α is set to 0.1. $\mathbf{V}_1$ & $\mathbf{V}_2$ are given in table 1. *: agents store a separate trust value T for each population.



(a) (b)

Figure 2: The results of the simulations across 50 runs. Subfigure (a): The mean performance level in simulations 1-7. The vertical line denotes the time tick at which Population 2 is introduced in the model. The lower horizontal line gives the expected performance of an agent with a randomized **P** matrix (being -2). The upper horizontal line gives the maximum performance an agent can attain (being 4). Subfigure (b): The mean trust level in simulations 4-7. The vertical line denotes the time tick at which Population 2 is introduced in the model.

In simulation 5, we recreate the situation of simulation 3, but now with a trust mechanism. As one can see, the trust mechanism is not capable of completely solving the problems arising in simulation 3: the final performance of the agents in simulation 5 is slightly better than in 3, but still sub-optimal.

In the plot of the trust levels in figure 2(b), it can be seen that the trust levels converge to 0.5. This is caused by having half of the agents demonstrating a faulty policy and half a correct one. The agents cannot improve their performance because they cannot discriminate between trustworthy and untrustworthy agents.

Simulations 6 and 7 are identical to 4 and 5 but for the introduction of a separate trust value for each population. Every agent is equipped with two T values. This means that each agent can have a different level of trust in the members of Population 1 versus the members of Population 2.

Maintaining seperate trust values for the two populations has only a negligible effect in a situation in which both agents have to learn the same task (simulation 4). In simulation 6 the dip in the performance associated with the introduction of Population 2 is somewhat shallower than in simulation 4. Otherwise the results of simulation 4 and 6 are fairly similar. However, being able to discriminate between different types of agents allows the agents in simulation 7 to perform much better than in simulation 5. Now, both populations are able to achieve a perfect score.

Discussion

The results indicate that the proposed trust mechanism is capable of regulating the extent to which agents rely on social learning. Equipping our agents with the mechanism boosts their performance in situations where social learning is potentially disadvantageous (i.e situations in which demonstrations are untrustworthy).

Interestingly, the trust mechanism, as it is proposed in this paper, is a biologically plausible strategy. Humans, but also animals (e.g. Cheney and Seyfarth, 1988), learn more if they trust the source of information (See Carpenter and Call, 2007, for additional references). Koenig and Harris (2005) report experiments in which children from the age of 4 learned the names of novel objects from people who have shown to be trustworthy earlier in the experiment. They do not endorse names supplied by people who earlier misnamed known objects (e.g. naming a ball as a shoe). So, while adapting trust is a strategy that is, as yet, not widely studied in animal behavior, some empirical findings support that it is indeed being used. Further research might discover more instances in which trust is a important factor in human and animal social learning.

It is important to note explicitly that the presented trust mechanism differs from social learning strategies that seek to copy high performant demonstrators. For example, Schlag (1998) proposed that social learning agents (animals) should copy others if they are performing better than they are themselves (*copy-if-better*). However, this requires agents to be able to assess the performance of others, which might not be easy to do (Laland, 2004), especially for artificial agents. The form of trust introduced in the current paper does not require agents to evaluate the performance of others. Instead, agents trust others if they act in the same way as they would given the same percept. Simulation 7 serves as a demonstration of the difference between acting based on trust or the performance of others. In the second phase of the simulation (after tick 50), Population 1 is clearly performing better than Population 2. Nevertheless, Population 2 quickly looses its initial trust in Population 1 and stops copying its behavior. In contrast, Population 2 has more trust in itself. If performance would dictate social learning, all agents should be copying Population 1.

Finally, we think that much of the strength of the proposed mechanism lies in the fact that it can be extended in various interesting ways. We list two of the extensions we consider the most interesting.

First, a fundamental feature of the proposed trust mechanism is that it generalizes over all percept-action pairs. This is to say, an agent that learns to trust another by observing its response to a given percept p, also trusts the others response to all other percepts p'. This behavior is in concordance with the findings in children reported by Koenig and Harris (2005). Indeed, it is hard to see what would be the function of a trust mechanism that does not generalizes across stimuli. In the current simulations, this property of the model is not fully exploited. Generalizing across stimuli might enable agents, just like their biological counterparts, to learn socially about significant but rare stimuli. Some stimuli might not occur frequently enough for agents to learn individually from these instances. However, by observing how other agents, that are judged trustworthy, react to the stimuli, agents could assemble enough learning trials to associate a proper response with these stimuli. One possible extension of the presented work could explore the behavior and the value of the model under such circumstances

Another interesting extension, already hinted at in simulation 7, would be to increase the number of trust values that agents maintain. In the extreme case, an agent could have a trust value associated with each other agent in the population. Trust levels associated with individual agents would enable agents to form trust networks directing the flow of information that is spread through social learning (See Coussi-Korbell and Fragaszy, 1995, for a seminal paper on directed social learning). Also, agents could learn which individuals' behavior is worthwhile to copy (see Dautenhahn and Nehaniv, 2007).

In conclusion, we presented an extendable mechanism that allows agents to regulate their reliance on social learning. The mechanism to boost the performance of agents in multi-agent settings that incorporate social learning. Importantly, the mechanism does not require agents to be able to judge whether the actions of observed demonstrators have a favorable outcome.

References

Acerbi, A., Marocco, D., and Nolfi, S. (2007). Social facilitation on the development of foraging behaviors in a population of autonomous robots. In Costa, F. A., Rocha, L. M., Costa, E., Harvey, I., and Coutinho, A., editors, *ECAL*, volume 4648 of *Lecture Notes in Computer Science*, pages 625–634. Springer.

Alissandrakis, A., Nehaniv, C., and Dautenhahn, K. (2004). Towards robot cultures? learning to imitate in a robotic arm test-bed with dissimilarly embodied agents. *Interaction Studies*, 5:3–44.

Belpaeme, T., de Boer, B., and Jansen, B. (2007). The dynamic emergence of categories trough imitation. In Dautenhahn, K. and Nehaniv, C. L., editors, *Imitation in Animals and Artifacts*. MIT Press.

Bonnie, K. E. and Earley, R. L. (2007). Expanding the scope for social information use. *Animal Behaviour*, 74(2):171–181.

Bonnie, K. E., Horner, V., Whiten, A., and de Waal, F. B. M. (2006). Spread of arbitrary conventions among chimpanzees: a controlled experiment. *Proceedings in Biological Science*, 274(1608):367–372.

Boyd, R. and Richardson, P. J. (1988). An evolutionary model of social learning: the effect of spatial and temporal variation. In Zentall, R. R. and Galef, B. j., editors, *Social Learning: Psychological and Biological Perspectives*, pages 29–48. Erlbaum, Hillsdale, NJ.

Boyd, R. and Richerson, P. J. (2006). *The Innate Mind: Culture and Cognition*, chapter Culture, Adaptation, and Innateness. Oxford University Press.

Carpenter, M. and Call, J. (2007). The question of what to imitate: inferring goals and intentions from demonstrations. In Dautenhahn, K. and Nehaniv, C. L., editors, *Imitation in Animals and Artifacts*. MIT Press.

Cheney, D. and Seyfarth, R. (1988). Assessment of meaning and the detection of unreliable signals in vervet monkeys. *Animal Behaviour*, 36:477–486.

Coolen, I., Dangles, O., and Casas, J. (2005). Social learning in noncolonial insects? *Current Biolology*, 15(21):1931–1935.

Coussi-Korbell, S. and Fragaszy, D. M. (1995). On the relationship between social dynamics and social learning. *Animal Behaviour*, 50(6):1441–1453.

Dautenhahn, K. and Nehaniv, C. (2007). An agent-based per- spective on imitation. In Dautenhahn, K. and Nehaniv, C. L., editors, *Imitation in Animals and Artifacts*. MIT Press.

Fiorito, G. (2001). Socially guided behaviour in non-insect invertebrates. *Animal Cognition*, 4(2):69–79.

Galef, B. and Laland, K. (2005). Social learning in animals: Empirical studies and theoretical models. *Bioscience*, 55(6):489–499.

Giraldeau, L.-A., Valone, T. J., and Templeton, J. J. (2002). Potential disadvantages of using socially acquired information. *Philosophical Transactions of the Royal Society of London, Series B*, 357(1427):1559–1566.

Kendal, R., Coolen, I., van Bergen, Y., and Laland, K. (2005). Trade-offs in the adaptive use of social and asocial learning. *Advances in the Study of Behavior*, pages 333–379.

Koenig, M. A. and Harris, P. L. (2005). The role of social cognition in early trust. *Trends in Cognitive Sciences*, 9(10):457–459.

Laland, K., Coolen, I., and Kendal, R. (2005). Why not use public information? *Science*, 308:354–355.

Laland, K. N. (2004). Social learning strategies. *Learning & Behavior*, 32(1):4–14.

Laland, K. N. and Williams, K. (1998). Social transmission of maladaptive information in the guppy. *Behavioral Ecology*, 9(5):493–499.

Leadbeater, E. and Chittka, L. (2007). Social learning in insects–from miniature brains to consensus building. *Current Biology*, 17(16):R703–R713.

Leadbeater, E., Raine, N. E., and Chittka, L. (2006). Social learning: ants and the meaning of teaching. *Current Biology*, 16(9):R323–R325.

Noble, J. and Franks, D. W. (2002). Social learning mechanisms compared in a simple environment. In Proceedings of the Eighth International Conference on Artificial Life.

Noble, J. and Todd, P. M. (2002). Imitation or something simpler? modelling simple mechanisms for social information processing. In *Imitation in Animals and Artifacts*. MIT Press, Cambridge, MA.

Pini, G., Tuci, E., and Dorigo, M. (2007). Evolution of social and individual learning in autonomous robots. Ecal Workshop: Social Learning in Embodied Agents.

Pongrcza, P., dm Miklsia, Kubinyia, E., Toplb, J., and Csnyia, V. (2003). Interaction between individual experience and social learning in dogs. *Animal Behavior*, 65(3):595–603.

Schlag, K. H. (1998). Why imitate, and if so, how? a boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78(1):130–156.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press, Harvard.

Whiten, A., Spiteri, A., Horner, V., Bonnie, K. E., Lambeth, S. P., Schapiro, S. J., and de Waal, F. B. M. (2007). Transmission of multiple traditions within and between chimpanzee groups. *Current Biology*, 17(12):1038–1043.

Zentall, T. R. (2006). Imitation: definitions, evidence, and mechanisms. *Animal Cognition*, 9(4):335–353.