Theoretical Population Biology 76 (2009) 189–196

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Simulated Trust: A cheap social learning strategy

Dieter Vanderelst [a,*], René M.C. Ahn [b], Emilia I. Barakova [b]

[a] Active Perception Lab, University Antwerp, Antwerp, Belgium

[b] Designed Intelligence Group, Eindhoven University of Technology, Eindhoven, The Netherlands

ARTICLE INFO

Article history:
Received 3 March 2009
Available online 19 July 2009

Keywords:
Social learning
Agent-based model
Directed social learning

ABSTRACT

Animals use heuristic strategies to determine from which conspecifics to learn socially. This leads to directed social learning. Directed social learning protects them from copying non-adaptive information. So far, the strategies of animals, leading to directed social learning, are assumed to rely on (possibly indirect) inferences about the demonstrator's success. As an alternative to this assumption, we propose a strategy that only uses self-established estimates of the pay-offs of behavior. We evaluate the strategy in a number of agent-based simulations. Critically, the strategy's success is warranted by the inclusion of an incremental learning mechanism. Our findings point out new theoretical opportunities to regulate social learning for animals. More broadly, our simulations emphasize the need to include a realistic learning mechanism in game-theoretic studies of social learning strategies, and call for re-evaluation of previous findings.

© 2009 Elsevier Inc. All rights reserved.

Throughout the animal kingdom, individuals exploit information that has been gathered by others. Animals ranging from invertebrates (Reviewed in Leadbeater and Chittka, 2007; Leadbeater et al., 2006; Fiorito, 2001) to great apes and humans (Tomasello, 1999; Whiten et al., 2007; Bonnie et al., 2006, for example) exhibit forms of social learning.[1] The widespread use of social learning among the taxa is caused by its enormous ecological advantages under many circumstances (See for example Kendal, 2005; Coolen et al., 2005; Bonnie and Earley, 2007, and references therein). Evolution favored social learning because it allows 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 physiological mechanisms range from fairly simple to thoroughly complex (Noble and Todd, 2002). However, when studying the dynamics and ecology properties of social learning one can ignore the differences in implementations and consider the exchange of information only (Coussi-Korbell and Fragaszy, 1995). This makes it possible to evaluate the advantages of social learning in theoretical studies focusing on the game-theoretic and computational aspects (e.g. Voelkl and No, 2008; Schlag, 1998).

This theoretical line of research, supported by empirical findings in animal behavior, has shown that the advantage of social learning is by no means universal. In contrast to intuition, social learning is advantageous only if one takes certain precautions (Laland, 2004; Galef and Laland, 2005; Ihara, 2008). Social learning can support the spread, the 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 circumstances under which this assumption does not hold (Giraldeau et al., 2002; Laland, 2004; Leadbeater and Chittka, 2007; Laland et al., 2005; Galef and Laland, 2005). Second hand information can be a.o. incomplete, outdated, biased, utterly wrong or already overexploited by others.

Theoretical and empirical studies have identified a number of strategies in animals and humans to determine when and from whom to learn socially (for a review see Laland, 2004). These strategies allow to use social learning in a more adaptive fashion, avoiding its potential pitfalls.

In this paper we focus on the type of strategies that regulate from which conspecifics animals learn socially (directed social learning strategies or 'whom' strategy according to Laland, 2004). These strategies are heuristics that allow animals to determine which animals have information that is interesting to copy. One class of these heuristics try to evaluate the value of the

* Corresponding author.
E-mail address: dieter.vanderelst@ua.ac.be (D. Vanderelst).

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

0040-5809/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2009.07.001

demonstrator's example indirectly through properties that are assumed indicative such as the age of the demonstrator (Coussi-Korbell and Fragaszy, 1995) or its social status (Ihara, 2008). A second class of strategies depend on a direct evaluation of the pay-offs of the demonstrator's actions. *Copy-If-Better* and *Copy-The-Most-Successful-Behavior* are examples of this class of strategies (see Laland, 2004). The first class of strategies will suffer if the correlation between the adaptive value of an individual's behavior and its personal traits is weak. The second class of strategies requires animals to be able to evaluate the outcome of a demonstrator's actions directly. A feat that might not be easy to accomplish for non-human learners (Laland, 2004).

In this paper, we propose a new type of strategy that focuses on the pay-offs of actions rather than on secondary traits of the demonstrator. In addition, the strategy does not require individuals to assess the pay-off received by others. In this way, the strategy reaps the benefits of both classes of strategies leading to directed social learning while avoiding their respective setbacks. We will refer to this new strategy as *Copy-if-similar*.

Summarized, the *Copy-if-similar* strategy can be stated as follows:

> An animal should trust whoever behaves like it would behave itself under similar circumstances. This is, an animal i should increase the trust it has in an animal j if and only if j repeatedly exhibits behavior in a situation x_1 that is considered to be adaptive by i. The trust i places in j should generalize across all situations x_n. The amount of trust an animal i places in j should be proportional to the extent to which it learns socially from agent j.

Critically, our *Copy-if-similar* strategy exploits the opportunities that arise when a new type of behavior is acquired through a series of learning experiences instead of in a one-shot learning. Incremental learning allows animals, at each point in time, to exploit their limited knowledge of the problem at hand to select the demonstrators that are most informative to them. This makes our simulations different from the one-shot learning models typically used (e.g. Schlag, 1998; Noble and Franks, 2002) when evaluating social learning strategies. Incremental learning might be the rule rather than the exception when learning complex or novel types of behavior. For example, Ottoni et al. (2005) report that it takes typically up to three years before capuchin monkeys master nutcracking through social learning.

In what follows, we present the *Copy-if-similar* strategy and evaluate its benefits in simulations. For this we compare it directly to an implementation of the *Copy-If-Better* strategy proposed by Schlag (1998) as a strategy that is guaranteed to lead to a high pay-off (given certain assumptions are fulfilled).

1. Setup

We have investigated the question of how animals can direct social learning by modeling a simple environment with a number of agents. The agents in this environment 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 the agents with the possibility to change the level of trust they have in each of the demonstrators. The level of trust in a demonstrator in turn determines the agent's reliance on the demonstrator's actions for social learning. The level of trust should be regarded upon as a theoretical construct that could incorporate a range of psychological mechanisms allowing animals to direct their social learning to certain members of the local population.

We investigate the learning behavior of the agents by evaluating their performance in simulations under various conditions while comparing the *Copy-if-similar* strategy with a version of the *Copy-If-Better* strategy adapted to incremental learning.

In all conditions, we consider two populations of agents that have the same cognitive architecture. The first population enters the simulations before the second one, and has therefore already acquired a high level of experience in the simulated environment when the second population is initiated.

2. Methods

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

In the simulations, time is represented by an integer. At each time step, 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 performs 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. So, in our simulations, social learning is a mechanism that acts in addition to individual learning rather than instead of it. This is, individual learning and social learning are compatible.

The behavior of the agents can be captured by a few simple rules. When learning individually the following sequence of events takes place (the numbers between square brackets correspond to those in Fig. 1):

- Step 1: An agent i is confronted with a randomly chosen percept p drawn from a limited set of percepts [1].
- Step 2: The agent i chooses an action a with which to respond to the percept based on its policy P_i [2]. The matrix P_i gives, for each percept p and action a, the chance of an agent i choosing action a when confronted with the percept p. Actions which have been highly associated with the percept through learning have a higher chance of being selected.
- Step 3: The environment responds to this action with the appropriate reward V_{pa} as given by the environment pay-off matrix V [3].
- Step 4: Based on the returned reward [3], the estimated pay-off Q_{ipa} for choosing the action a given the percept p by the agent i is adapted according to Eq. (1). In this equation α_Q is a parameter governing the speed with which Q_{ipa} is updated.

$$\Delta Q_{ipa} = \alpha_Q (V_{pa} - Q_{ipa}). \tag{1}$$

- Step 5: The agent i updates the value $P(a|p)$ in its policy matrix P_i [4], effecting incremental changes to the probabilities for the various actions a given the percept p, based on the changed estimates of the pay-offs. The updating of $P(a|p)$ is done according to Eq. (2). In the literature on reinforcement learning, this form of updating action policies is known as pursuit learning (Sutton and Barto, 1998).[2] In Eq. (2) a^* is the action for which the current estimated pay-off is the highest. This is, $a^* = max_a\, Q_{ipa}$. After updating its policy P_i, an agent stores p, a and the reward V_{pa} for consultation by other agents during social learning.

$$\begin{cases} \text{for } a = a^* : \Delta P_i(a|p) = \alpha_I (1 - P_i(a|p)), \\ \text{for } a \neq a^* : \Delta P_i(a|p) = \alpha_I (0 - P_i(a|p)). \end{cases} \tag{2}$$

[2] Notice that the equations governing the simulations are agnostic with respect to the scale of V.

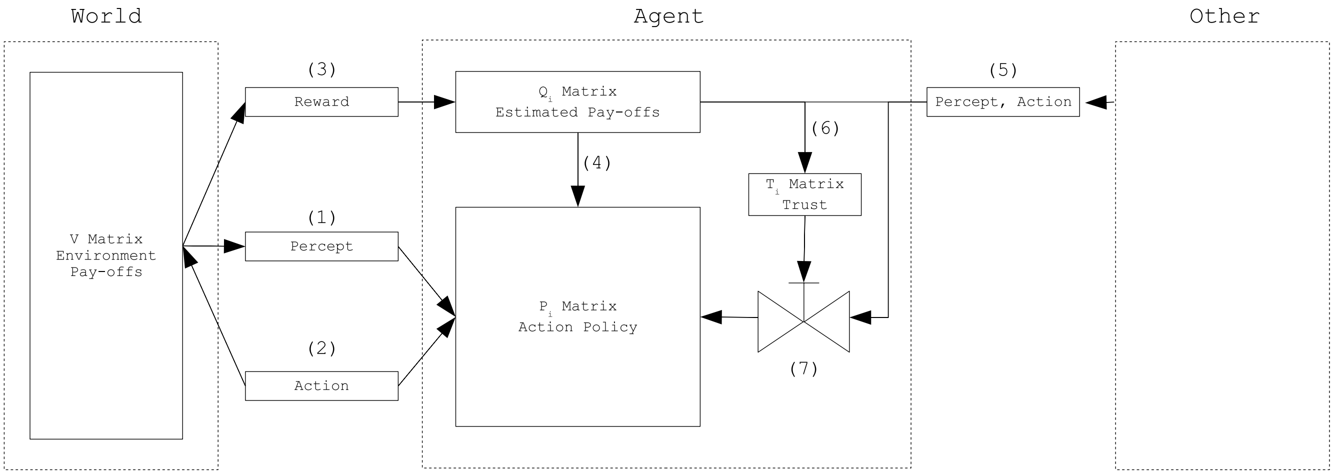

Fig. 1. The cognitive architecture of the agents in the simulations and its relationship with the environment. Each agent consists of three matrices: Q, P and T.

After all agents have learned individually, they perform a number of social learning trials during which they serially sample the behavior of several other agents. When learning socially, the following sequence of events take place (the numbers correspond to the those in Fig. 1.):

- Step 1: An agent i consults the latest percept p and the action a stored by another agent j during individual learning. This is analogous to perceiving in what situation j finds itself (percept) and how it reacts (action) [5]. Agent j is chosen randomly from the set of agents currently in the simulation.
- Step 2: Based *on its own estimated pay-offs* $Q_{ip.}$ for the given percept, the agent i updates its trust T_{ij} in the other j [6]. See Eq. (3). Eq. (3) increases the trust of agent i in j if j chooses an action in response to p agent i currently thinks to have a higher pay-off than the average expected pay-off. Trust values are constrained to the range [0, 1]. In Eq. (3), a' denotes the action demonstrated by agent j. So, $Q_{ipa'}$ is i's estimate of the pay-off for the action a' chosen by j.

$$\Delta T_{ij} = \begin{cases} \alpha_T & \text{if } \sum_a [P_i(a|p) \times Q_{ipa}] \le Q_{ipa'}, \\ -\alpha_T & \text{if } \sum_a [P_i(a|p) \times Q_{ipa}] > Q_{ipa'}. \end{cases} \tag{3}$$

Under the *Copy-If-Better* strategy, the trust in another agent is updated based on the reward received by j when it executed action a in response to percept p. Therefore, when evaluating the *Copy-If-Better* strategy, Eq. (3), which is used under the *Copy-If-Similar* strategy, is replaced by Eq. (4). Notice that in this equation the term $Q_{ipa'}$ is replaced by a term V_{jpa} referring to the actual reward received by the observed agent j when performing action a in response to percept p. Eq. (4) increases the trust of agent i in agent j if the expected pay-off of agent i, using its current policy, for the percept p is lower than the observed pay-off of agent j for this percept.

$$\Delta T_{ij} = \begin{cases} \alpha_{T_{ij}} & \text{if } \sum_a [P_i(a|p) \times Q_{ipa}] \le V_{jpa}, \\ -\alpha_{T_{ij}} & \text{if } \sum_a [P_i(a|p) \times Q_{ipa}] > V_{jpa}. \end{cases} \tag{4}$$

In Eqs. (3) and (4), α_T is a step size parameter governing the size of the trust update.
- Step 3: The agent updates its policy P_i for the given percept depending on the trust T_{ij} it has in the other [7] according to Eq. (5). The parameter α_S is the step size governing the speed of social learning.

Table 1
The environment pay-off matrices used in the reported simulations.

Actions								
	Values 1 (V_1)				Values 2 (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

$$\begin{cases} \text{for } a = a^* : \Delta P_i(a|p) = \alpha_S \times T_{ij} \times (1 - P_i(a|p)), \\ \text{for } a \neq a^* : \Delta P_i(a|p) = \alpha_S \times T_{ij} \times (0 - P_i(a|p)). \end{cases} \tag{5}$$

As can be deduced from the equations above, under the *Copy-if-similar* strategy, 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 it thinks to be rewarding, the level of trust it has in this agent will rise.

All matrices P_i are initialized with random values between 0 and 1 with the constraint that each row must sum to 1. The matrices Q_i, containing estimates of the matrix V that are progressively constructed by the agents over the course of a simulation, are initialized containing only zeros. At the start of the simulation the matrices T_i contain only ones signifying that initially trust is total. However, initiating the matrices T_i with zeros leads to similar results (results not shown).

As experimenters we evaluate an agent's policy by calculating the expected performance E according to Eq. (6).

$$E_i = \sum_p \sum_a P_i(a|p) \times V_{pa}. \tag{6}$$

3. Results

3.1. Simulation parameters

Both population 1 and 2 contain 40 agents. Population 2 enters the simulation after time tick 50. Simulations are run for 200 steps. Parameters α_Q, α_S, α_I and α_T are set to 0.1. During each time step of the model, each agent performs one individual learning trial and five social learning trials. We used 4 percepts and 4 actions in the simulations reported below. The environment pay-off matrices V used in the reported simulations are given in Table 1.

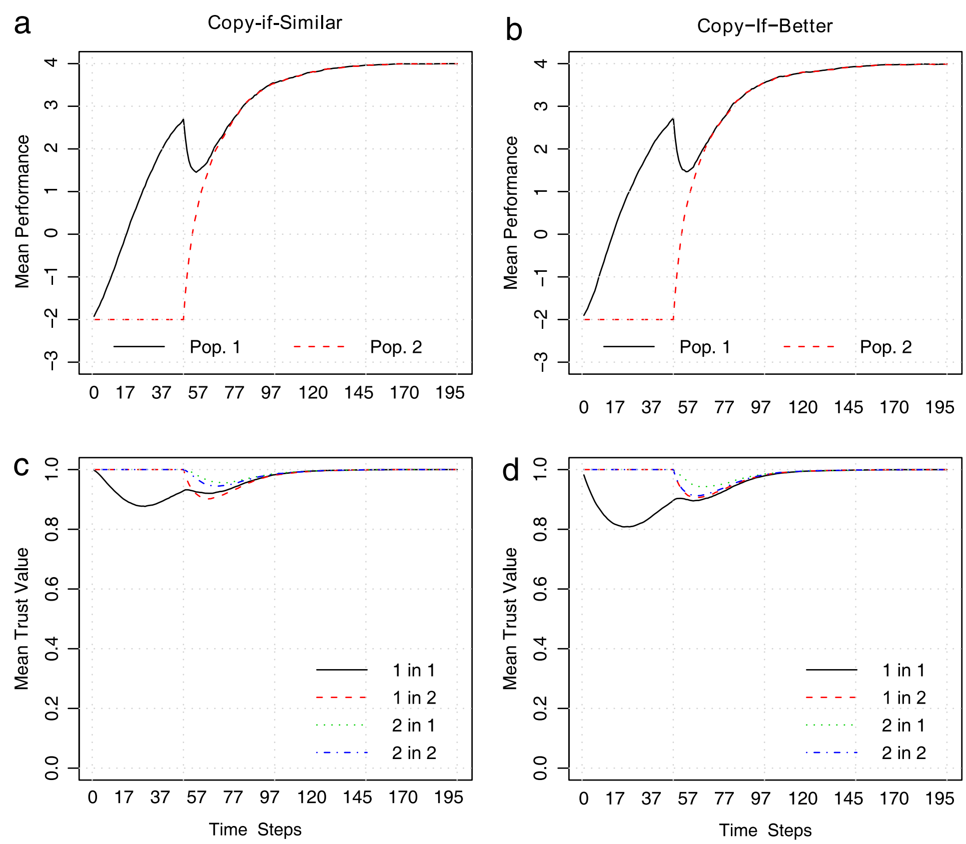

Fig. 2. Results of simulation 1 (averaged across 50 runs). The top panels (a–b) depict the mean performance of the agents using either the *Copy-if-similar* or the *Copy-If-Better* strategy to regulate their trust in others. The bottom panels (c–d) show the mean trust ($\bar{T}$), for both strategies, each population has in each other population as a function of time. The legend annotation '1 in 2' means the trust population 1 has in population 2, etc.

3.2. Simulation 1: Identical learning tasks

In the first simulation the *Copy-if-similar* strategy is directly compared with the *Copy-If-Better* strategy under favorable circumstances. The V matrix to be learned by both populations is given as V_1 in Table 1. The results of simulation 1 are depicted in Fig. 2.

Simulation 1 is conceived to be a situation that favors social learning. Social learning allows population 2 to learn faster than population 1 (and to catch up with it). The main conclusion to be drawn from the results is that both strategies lead to a high reliance on social learning. Ultimately, the performance of both populations reaches the maximum. So, under favorable circumstances, both strategies lead to a similar learning trajectory.

3.3. Simulation 2: Different learning tasks

In the second simulation, the second population we introduce is required to learn a different task than the first population. The pay-offs for population 1 are governed by V_1 while the pay-offs for population 2 are governed by V_2 (see Table 1). This models a situation in which animals are confronted with unreliable demonstrators. While unlikely to occur in reality under this form, this situation poses a good test case for any strategy that directs social learning to reliable demonstrators.

From the results (Fig. 3) it is clear the *Copy-if-similar* strategy can cope with this situation while the *Copy-If-Better* strategy rely heavily on the demonstrations by population 1 (because of the high yields). However, for population 2 this is not adaptive. The *Copy-if-similar* strategy allows the agents to assess the quality of the demonstrator based on their past experience. This enables them to attach proper weights to the demonstrations.

3.4. Simulation 3: Modeling the trust-matrix as an adjacency matrix

In the two simulations above, we have assumed that there is an equal probability for an agent to observe every other agent. Agents learned how much weight they should attach to the observations. However, this implementation requires animals to maintain a set of weights corresponding to every other animal in the population. This might pose somewhat of a memory load. More importantly, agents were required to be able to perceptually distinguish between individual members of the population.

The need for individual recognition can be resolved in two ways, either by assuming that agents can distinguish between a few different, meaningful, *classes* of fellow agents. This was simulated by Vanderelst et al. (2008). Alternatively, it is not required that the trust matrix is explicitly memorized by the animals. This information could be partially offloaded to physical behavior of animals. Animals could try to spend more time in the vicinity of animals when found trustworthy. In this way, the proximity of animals to each other would encode the trust they have in each other. This could lead to self-reinforcing biases in their observations. This second solution will be explored in this section of the paper.

To investigate the performance of the strategies with this adaptation, we modeled the trust matrix as an adjacency matrix, giving the chance for each agent i to observe an agent j during a social learning trial. In simulation 1 and 2, an agent i had an equal probability to observe each other agent j. The trust value determined how much weight i should place on the demonstrations of j. In contrast, in simulation 3, we take T to be a matrix of chances giving for

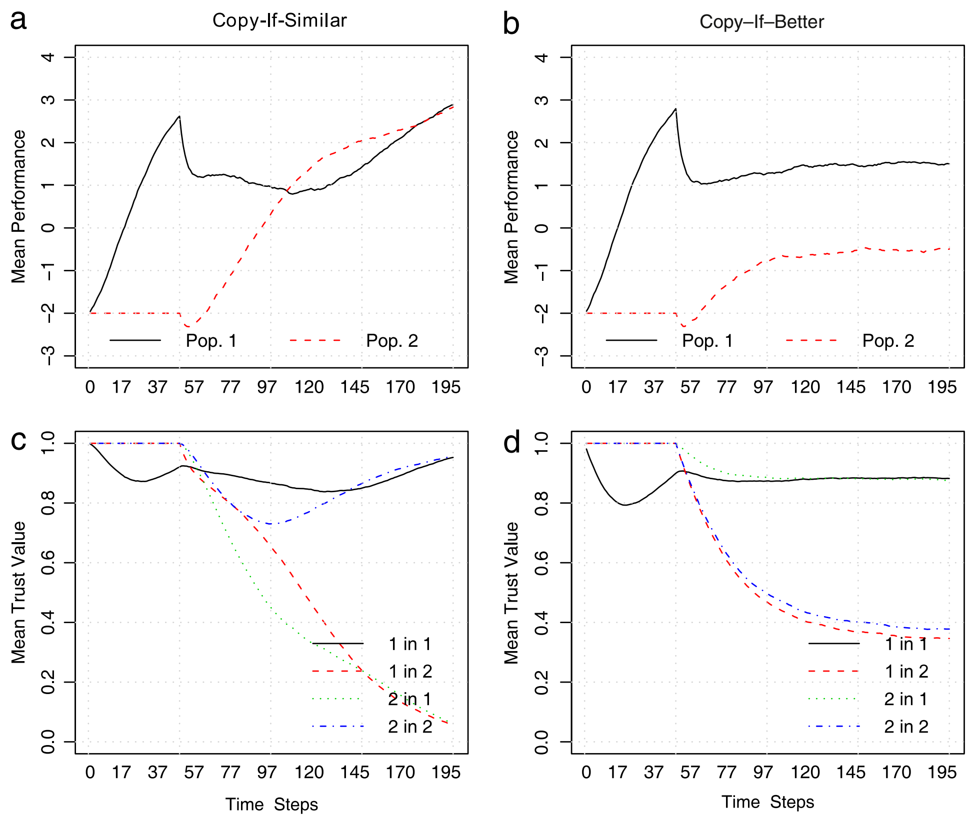

Fig. 3. Results of simulation 2 (averaged across 50 runs). The top panels (a–b) depict the mean performance of the agents using either the *Copy-if-similar* or the *Copy-If-Better* strategy to regulate their trust in others. The bottom panels (c–d) show the mean trust ($\bar{T}$), for both strategies, each population has in each other population as a function of time. The legend annotation '1 in 2' means the trust population 1 has in population 2, etc.

each agent i the chance of observing j. For this purpose Eq. (5) was reduced to Eq. (7).

$$\begin{cases} \textit{for } a^* : \Delta P_i(a|p) = \alpha_S \times (1 - P_i(a|p)), \\ \text{for } a' \neq a^* : \Delta P_i(a'|p) = \alpha_S \times (0 - P_i(a'|p)). \end{cases} \tag{7}$$

Also, since T contains probabilities, additional equations are needed to control the correct updating of T in function of ΔT_{ij}. This is, ΔT_{ij} as given by Eqs. (3) and (4) must be transformed into a series of $\Delta T'_{ij}$ for every agent j such that the sum $\sum_{j=1}^{n} T_{ij}$ is equal to one at all time. If ΔT_{ij} is positive, T is updated according to $\Delta T'_{ij}$ given by Eq. (8). When ΔT_{ij} is negative, Eq. (9) is used.

$$\text{if } \Delta T_{ij} > 0 \begin{cases} \text{for } j: \ \Delta T'_{ij} = \alpha_T(1 - T_{ij}) \\ \text{for } j' \neq j: \ \Delta T'_{ij} = \alpha_T(0 - T_{ij}) \end{cases} \tag{8}$$

$$\text{if } \Delta T_{ij} < 0 \begin{cases} \text{for } j: \ \Delta T'_{ij} = \alpha_T(0 - T_{ij}) \\ \text{for } j' \neq j: \ \Delta T'_{ij} = \alpha_T(1 - T_{ij}). \end{cases} \tag{9}$$

To summarize, in simulation 3, the algorithm outlined under 'Methods' is altered in two ways: (1) The agent j from which an agent i learns socially, is no longer chosen randomly but according to the probability T_{ij} in matrix T. (2) The update of the policy P_i is no longer attenuated by the level of trust agent i has in agent j (Eq. (5) is replaced by Eq. (7)).

The results in Fig. 4 show that the *Copy-if-similar* strategy still copes very well under these circumstances while the *Copy-If-Better* strategy fails. As in simulation 2, the agents of population 2 place too much confidence in population 1. Because a slight bias in T is reinforced, all agents quickly converge to a situation where they uniquely observe members of population 1.

Representing the trust matrix containing the chances for each agent to observe another, as the spatial distribution of agents is not trivial. The trust matrix does not satisfy the requirements of a distance matrix: it is not symmetrical and the triangle inequality is not guaranteed. In fact, the trust matrix, being an adjacency matrix, can be thought of as a directed graph. Fig. 5 illustrates this by plotting a directed graph based on an arbitrary matrix T using the Fruchterman and Reingold's algorithm (1991) implemented by Butts et al. (2008).

Kruskal's Non-metric 2D Multidimensional Scaling (Kruskal, 1964; Venables and Ripley, 2002) was employed to assess whether the trust matrix could be partially fitted using a spatial representation of the agents. In effect, this algorithm tries to place the agents on a 2D plane such that the ordinal distance between each agent i and agent j was inversely related to the chance of i observing j. This is, a monotonic non-linear relation between T_{ij} and the spatial distance between agents was imposed. Because the scaling algorithm assumes a symmetrical distance matrix, a symmetrical matrix S was derived from T as $S = 1 - \frac{1}{2}(T + T^T)$. A spatial representation was fitted to S for each time step of the 50 replications of simulation 3. The mean chance for an agent i to observe an agent j in function of the fitted distance between them was calculated. The results are depicted in Fig. 6.

From Fig. 6 it can be seen that the trust matrix T can, to a certain extent, be represented as distances between agents: agents trusting each other more were placed closer to each other by the scaling algorithm. Moreover, mean initial stress of the fitted solutions (using a random spreading of the agents in space) was 58.83 (sd: 6.51). The mean final stress was 35.99 (sd: 3.48). Stress R is calculated using Eq. (10) (Kruskal, 1964). In this equation, $f(S_{ij})$ is the distance fitted to the trust value S_{ij}.

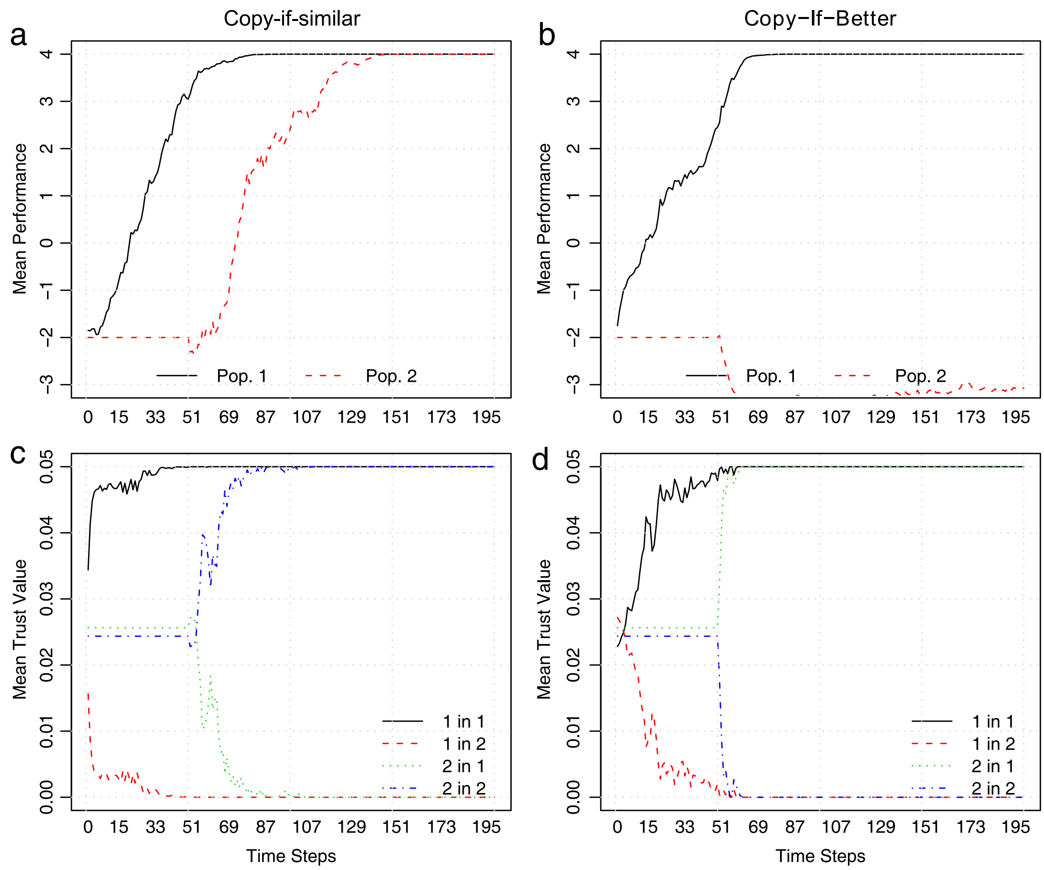

Fig. 4. Results of simulation 3 (averaged across 50 runs). The top panels (a–b) depict the mean performance of the agents using either the *Copy-if-similar* or the *Copy-If-Better* strategy to regulate their trust in others. The bottom panels (c–d) show the mean trust ($\overline{T}$), for both strategies, each population has in each other population as a function of time. The legend annotation '1 in 2' means the trust population 1 has in population 2, etc.

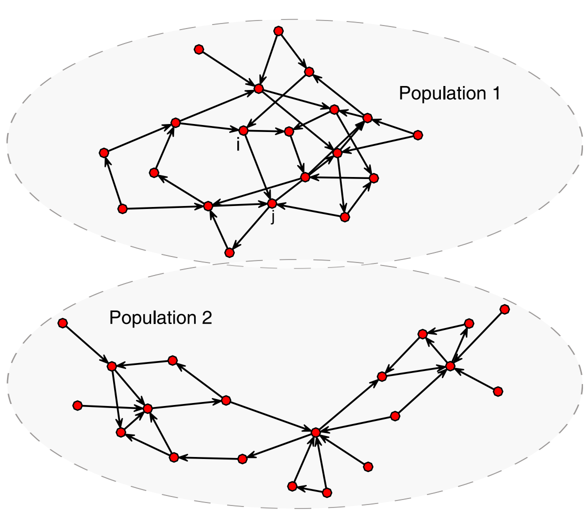

Fig. 5. A visualization of the agents and the trust matrix using Fruchterman and Reingold's algorithm at time step 150 in a single replication of the model. The circles represent each of the 40 agents in the simulation. The arrows represent the trust placed by each agent in other agents. Only trust values equal to or larger than 0.01 are visualized. For example, agent i in population 1 trusts agent j in population 1 and agent i is trusted by two other agents from the same population. Notice that, at this stage of the simulation, two subnetworks of agents have formed that trust each other only. As indicated in the figure, these groups correspond to Population 1 and Population 2. This figure illustrates the fact that matrix T can be represented as a directed graph.

$$R = 100 \cdot \sqrt{\frac{\sum_{ij}(f(S_{ij}) - S_{ij})^2}{\sum_{ij} S_{ij}^2}}. \tag{10}$$

Thus, agents are able, at least partially, to use distance as a means of 'storing' the trust they have in others. They could restrict their attention to those agents that are closest to them. This reduces the need to recognize all agents individually.

4. Conclusion

The simulations presented in this paper show that a strategy based on a limited (and readily available) knowledge of the environment is viable when incorporating a learning mechanism that requires multiple learning trials. The presented strategy remediates the problems that arise when the *Copy-If-Better* strategy is presented with untrustworthy demonstrators while also being more parsimonious. The *Copy-if-similar* strategy provides an alternative to the *Copy-If-Better* strategy for animals to choose whom to copy. Simulation 3 showed one possibility for implementing the strategy while reducing the need for detailed perception of others.

In summary, the *Copy-if-similar* strategy might be considered as a more robust variant of the *Copy-If-Better* strategy.

The parsimonious approach of not relying on the observation of the outcomes of the action of others, also makes the *Copy-if-similar* more robust with respect to time constraints. Under the *Copy-If-Better* strategy, animals need to observe pay-offs that might be delayed. In this case, animals must be present when the pay-off finally arrives in order to observe it. Additionally, they must be able

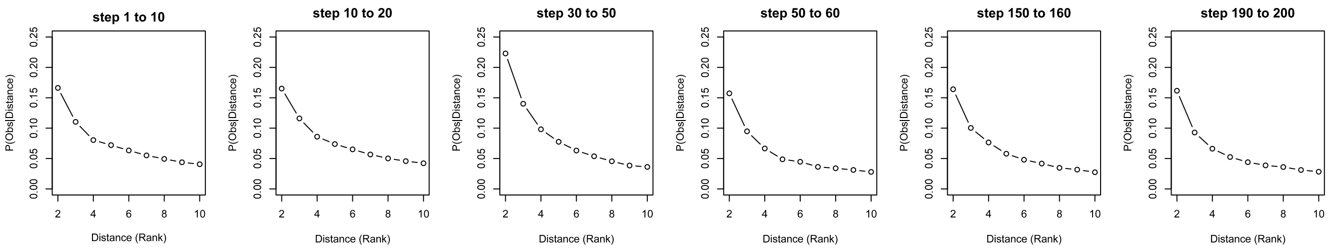

Fig. 6. The chance of an agent i to observe and agent j in function of the distance between them as fitted by Multi Dimensional Scaling (in rank order). The results are split according to different time slices of the simulation. Results averaged across 50 replications.

to relate it to the (right) previous actions of the demonstrators. This might be assumed to be harder as more time passes between action and pay-off. By only taking into account the actions of others, the *Copy-if-similar* avoids these timing issues.

Importantly, in the spatial analysis presented in simulation 3, fitting a perfect spatial model to the matrix T proved impossible. Mathematically, this follows directly from the properties of adjacency matrices. Therefore, while the spatial aggregation of animals might support the current implementation of the presented strategy, it can not do so completely. Some reliance on memory and perception (individual recognition) will be necessary to complement this.

Two broader conclusions can be drawn from the simulations.

First, computational studies of social learning should model the temporal dynamics of learning realistically. Schlag (1998) choose explicitly to model agents without internal temporal dynamics. However, as far as learning is concerned, incremental changes over time in the internal states of agents seem to occur often. This is demonstrated by the learning curves for a wide range of learning tasks. As shown in the simulations, when allowing for incremental learning, opportunities arise that might be exploited by animals that go unnoticed when modeling learning as a one-shot process.

Second, as noted in the introduction, authors have typically been ignoring the underlying implementations when evaluating social learning strategies by focusing on the information transmission only (Voelkl and No, 2008; Schlag, 1998). However, when evaluating the plausibility of these strategies, it is necessary to evaluate the possible implementations of the strategies (Laland, 2004). Strategies requiring less resources can be assumed to be more likely to evolve and to be more common in nature. For this reason, we have included an analysis showing how the cognitive and perceptual load of the presented strategy can be partially offloaded to the physical aggregation of animals. This shows that the complexity of learning strategies should be considered in terms of implementations rather than in terms of computational complexity. The physical world – here the location of the agents in space – might allow for simpler implementations of an algorithm (See Veelaert and Peremans, 1999; Lee, 1994, for analogous examples in perception and action control).

While the presented strategy was proposed purely on theoretical grounds, some predictions about which species might use this type of strategy are possible. In particular, animals capable of learning complex tasks requiring a long learning process and living in stable groups that allow for multiple encounters between the same animals are possible candidates. Primates might be the group of animals in which the occurrence of this strategy is most likely (for example, see the data reported by Ottoni et al., 2005). Indeed, some evidence for the existence of the usage of the *Copy-if-similar* strategy exists in humans. 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 objects known to them (e.g. naming a ball as a shoe). Language is probably the best example of a skill that is learned socially through repeated observation. Therefore, this is exactly the sort of social learning that could benefit from the strategy presented in this paper. It remains to be seen whether other animals exploit the theoretical opportunities pointed out in this paper.

References

Bonnie, K.E., Earley, R.L., 2007. Expanding the scope for social information use. Animal Behaviour 74 (2), 171–181.
Bonnie, K.E., Horner, V., Whiten, A., 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., Richardson, P.J., 1988. An evolutionary model of social learning: the effect of spatial and temporal variation. In: Zentall, R.R., Galef, B.J. (Eds.), Social Learning: Psychological and Biological Perspectives. Erlbaum, Hillsdale, NJ, pp. 29–48.
Butts, C.T., Handcock, M.S., Hunter, D.R., 2008. Network: Classes for relational data. r package version 1.3.
Coolen, I., Dangles, O., Casas, J., 2005. Social learning in noncolonial insects? Current Biolology 15 (21), 1931–1935.
Coussi-Korbell, S., Fragaszy, D.M., 1995. On the relationship between social dynamics and social learning. Animal Behaviour 50 (6), 1441–1453.
Fiorito, G., 2001. Socially guided behaviour in non-insect invertebrates. Animal Cognition 4 (2), 69–79.
Fruchterman, T., Reingold, E., 1991. Graph drawing by force-directed placement. Software - Practice and Experience 21 (11), 1129–1164.
Galef, B., Laland, K., 2005. Social learning in animals: Empirical studies and theoretical models. Bioscience 55 (6), 489–499.
Giraldeau, L.-A., Valone, T.J., 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.
Ihara, Y., 2008. Spread of costly prestige-seeking behavior by social learning. Theoretical Population Biology 73, 148–157.
Kendal, R., Coolen, I., van Bergen, Y., Laland, K., 2005. Trade-offs in the adaptive use of social and asocial learning. Advances in the Study of Behavior 333–379.
Koenig, M.A., Harris, P.L., 2005. The role of social cognition in early trust. Trends in Cognitive Sciences 9 (10), 457–459.
Kruskal, 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29, 1–27.
Laland, K., 2004. Social learning strategies. Learning & Behavior 32 (1), 4–14.
Laland, K., Coolen, I., Kendal, R., 2005. Why not use public information?. Science 308, 354–355.
Leadbeater, E., Chittka, L., 2007. Social learning in insects–from miniature brains to consensus building. Current Biology 17 (16), R703–R713.
Leadbeater, E., Raine, N.E., Chittka, L., 2006. Social learning: Ants and the meaning of teaching. Current Biology 16 (9), R323–R325.
Lee, D., 1994. In: Davies, M., Green, P. (Eds.), Perception and Motor Control in Birds. Springer-Verlag.
Noble, J., Franks, D.W., 2002. Social learning mechanisms compared in a simple environment. In: Proceedings of the Eighth International Conference on Artificial Life.
Noble, J., 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.
Ottoni, E.B., de Resende, B.D., Izar, P., 2005. Watching the best nutcrackers: What capuchin monkeys (cebus apella) know about others' tool-using skills. Animal Cognition 8 (4), 215–219.
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., Barto, A.G., 1998. Reinforcement Learning: An Introduction. MIT Press, Cambridge.
Tomasello, M., 1999. The Cultural Origins of Human Cognition. Harvard University Press, Harvard.
Vanderelst, D., Ahn, R., Barakova, E., 2008. Simulated Trust: Towards Robust Social Learning. MIT Press, Cambridge, MA.
Veelaert, P., Peremans, H., 1999. Flexibility maps: A formalisation of navigation behaviours Source. Robotics and Autonomous Systems 27 (3), 151–169. doi:10.1016/S0921-8890(98)00071-2.

Venables, W.N., Ripley, B.D., 2002. Modern applied statistics with *s*, Fourth ed. Springer, New York, ISBN: 0-387-95457-0.
Voelkl, B., No, R., 2008. The influence of social structure on the propagation of social information in artificial primate groups: A graph-based simulation approach. Journal of Theoretical Biology 252 (1), 77–86.
Whiten, A., Spiteri, A., Horner, V., Bonnie, K.E., Lambeth, S.P., Schapiro, S.J., 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.