

Social Learning Mechanisms Compared in a Simple Environment

Jason Noble and Daniel W. Franks

School of Computing, University of Leeds

[jasonn|dwfranks]@comp.leeds.ac.uk

Abstract

Social learning can be adaptive, but little is known about the underlying mechanisms. Many researchers have focused on imitation but this may have led to simpler mechanisms being underestimated. We demonstrate in simulation that imitative learning is not always the best strategy for a group-living animal, and that the effectiveness of any such strategy will depend on details of the environment and the animal's lifestyle. We show that observations of behavioural convergence or "traditions" might suggest effective social learning, but are meaningless considered alone.

GETTING by in the world is all about doing the right thing at the right time: eat the fallen fruit, avoid the poisonous mushroom, run from the approaching predator. But how does an animal produce the best response, or at least a satisfactory response, in each of the many possible situations it may face? We can distinguish two main sources of information used in this constant action-selection task. First, there is genetically inherited information, such as instinctive drives or tendencies. You run from the approaching predator because you were born with a brain that was wired up that way. Second, there is information based on experience, i.e., the results of learning. You avoid the mushroom because you've learned that it will make you sick.

Learning sometimes involves an individual finding out through trial and error about what to do in different situations. However, in a social species, there is often valuable information to be gleaned from the behaviour of others: for example, why risk poisoning by sampling potential foodstuffs at random when you could simply eat the same things your conspecifics are eating? In this paper we are concerned with this idea of *social learning*, and will compare the performance of different social learning mechanisms in a simulated environment.

Researchers in animal behaviour are increasingly in agreement that social learning is an important element in the overall adaptive strategy of many species (see Heyes & Galef, 1996, for a general review, or Box & Gibson, 1999, for case studies on social learning in mammals). From lion cubs learning about how to bring down prey, to rats learning through smell about which foods are safe to eat, it is clear that animals use the behaviour of their more experienced peers as a source of valuable information. However, in reviewing the literature on social learning, we often find that although investigators believe social learning to be occurring in a particular species, there is very little known about the specific mechanisms that underlie it. For example, in Box and Gibson's edited collection, many contributors discuss the ecology of their chosen

species in order to illustrate the likely roles that social learning may play, but can only speculate about exactly *how* the animals in question actually learn through the observation of another's behaviour.

Another curious fact about the social learning literature is the emphasis on imitation as a potential mechanism. Imitation — or “imitative learning” (Tomasello, 1996, p. 324) — is a cognitively complex process that requires not only perceiving and reproducing the bodily movements of another, but understanding the changes in the environment caused by the other's behaviour, and finally being able to grasp the “intentional relations” between these, i.e., knowing how and why the behaviour is supposed to bring about the goal. This emphasis on imitation is particularly strong in the case of work on primates (see, e.g., Whiten, 1998). A focus on imitation is not surprising: it is likely to be implicated in the explosive cultural evolution of our own species, and so we have a natural curiosity about the extent to which other animals in general, and our primate cousins in particular, might share this ability. Nevertheless, looking at imitation alone may blind us to the potential of much simpler mechanisms through which the behaviour of one animal can influence the learning experience of another.

Although we often do not understand *how* social learning occurs, there is a body of work on *why* it might occur, i.e., on the conditions under which it would offer a selective advantage. Models of cultural transmission (Boyd & Richerson, 1985) and related phenomena such as “highly horizontal transmission” (Laland, Richerson, & Boyd, 1996) help to delineate the conditions under which it will be advantageous for individuals to learn from others rather than finding things out for themselves. However, these models tend to result in rather gen-

eral conclusions. For example, Laland et al. (1996) summarize work on cultural transmission by relating it to rates of change in the environment: in static environments, genetic transmission of behaviour patterns can do the job, and in rapidly changing environments, only individual learning can keep up, whereas at intermediate rates of change, social learning will offer an advantage. Results like these are useful, but to get a complete picture of social learning we also need to understand the mechanisms involved.

Individual-based simulations, as used in artificial life research, represent an excellent way of instantiating a hypothesized social learning mechanism and thereby testing its plausibility. ALife models of this type have certainly been explored in related domains such as communication, dominance and territorial behaviour, and flocking or schooling. There are also a significant number of robotics researchers who share the primatologists' interest in imitation and see it as a way of improving human-robot or robot-robot communication (for a review of this work see Schaal, 1999).

However, as yet relatively few ALife models have looked at social learning using mechanisms simpler than imitation. One example is the work of Toquenaga, Kajitani, and Hoshino (1995) who constructed a simulation of foraging and nesting behaviour in egrets. They used their model to demonstrate that learning socially about the location of food, through simply approaching other birds, is more likely to evolve when food resources are patchy rather than evenly distributed. Toquenaga et al. suggest that flock foraging and colonial roosting may be based on this principle. In the same vein, Noble, Todd, and Tuci (2001) considered the selective pressures affecting social learning about food

in Norway rats. They found that a surprisingly simple strategy, which involved ignoring signs of food poisoning in others, was favoured as long as toxins were very dangerous. This finding was used to explain apparently paradoxical aspects of rat behaviour.

Noble and Todd (in press) have previously argued that a focus on imitation has led to the power of simpler social learning processes being underestimated; they describe a number of candidate simple mechanisms, and encourage the use of simulation models as a way of demonstrating the potential of such mechanisms. The purpose of the current paper is to follow up on this work by instantiating some of the mechanisms discussed by Noble and Todd, in a simple environment that reflects the learning challenges faced by a hypothetical animal. The aim is to show that simple mechanisms can perform well under the right environmental circumstances, and hopefully to demonstrate that imitative learning has received a disproportionate amount of attention. We must stress that even though our simulated agents are inspired by several different primates, our goal is not to model learning in any particular species, but to set up a specific, controlled environment in which the performance of different social learning strategies can be compared.

In looking at the data that is output by our simulation, we want to adopt the perspective of an ethologist who is observing behaviour in the field and trying to come up with conclusions about the underlying mechanisms. Noble and Todd (in press) make the point, following Braitenberg (1984), that human observers of animal and robot behaviour have a propensity to invoke mechanisms that are more complex than those strictly needed to explain that behaviour. For example, the presence of a distinctive behaviour pattern across all

members of a group could be taken as evidence for sophisticated social learning, but might result simply from similar individual learning experiences. We hope to show that much of the data a field observer might plausibly collect about a population could easily lead them astray in theorizing about what the animals were really doing.

The simulation

The environment consists of a 10×10 toroidal grid world, in which each square is either empty (20%) or contains a resource (80%). In addition, a grid square may contain a tool that agents are free to pick up and take away; agents may also drop tools. Agents occupy one square at a time, and more than one agent can occupy the same square. At each time step an agent chooses one of 12 distinct actions, and performing an action may lead to positive or negative payoffs depending on the context. An agent's choice of action is based on feeding its perceptual state into a reinforcement learning algorithm and choosing the action with the highest expected payoff (with a 20% chance of choosing a random action, in order to encourage exploration). Agents can perceive their local resource and the available tools — one on the ground and one in their grasp — as well as the presence of other agents (see Figure 1). They can also perceive the resource type in a randomly selected square drawn from the eight-square neighbourhood around their current location. If they choose to move, they will travel to this square.

The environment is deliberately simple, even playful: it is certainly not meant to be realistic. The guiding principles were, first, not to include an excessive number of distinct states, as this would make things difficult for the agents' reinforcement learning algorithm, and second, to ensure that

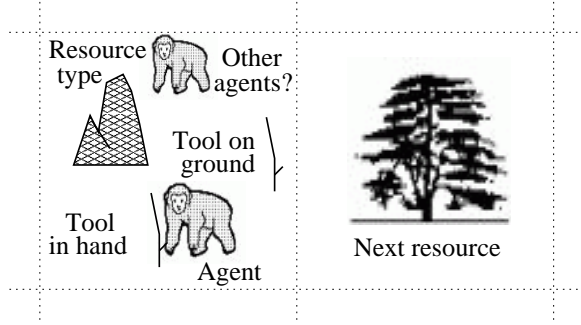


Figure 1: Aspects of the environment that can be perceived by each agent: the focal agent is in the left-hand grid square, and can perceive its local resource type — in this case a termite nest — the fact that it is grasping a stick, the presence of another stick on the ground, the presence of a second agent, and the proximity of a tree in the next grid square. If the agent chose the *travel* action, it would move to the tree.

some potential rewards were easily discovered and that others were harder to find (although possibly made easier with the right kind of social learning).

Generally, different resources are associated with distinct payoffs, and require the right action or sequence of actions to obtain them. The different resource types are listed below. Actions and tools are shown in italics, means and standard deviations for payoffs are shown as μ and σ respectively, P denotes the prevalence of that resource in the environment, and n refers to the number of times the resource can be used before it is exhausted. An exhausted resource reverts to empty, and a replacement appears in a randomly chosen empty square elsewhere in the environment.

Tree: being in the shade of a tree is pleasant, with a small constant reward ($\mu = 1$, $\sigma = 0$) regardless of the chosen action. $P = 10\%$, $n = \infty$.

Berries: an easy resource to figure out, berries are assumed to be automatically eaten ($\mu = 10$,

$\sigma = 5$) regardless of action. If the agent *shakes* the berry bush, twice the payoff is obtained ($\mu = 20$, $\sigma = 5$). $P = 10\%$, $n = 20$.

Nettles: normally cause a nasty sting ($\mu = -10$, $\sigma = 5$) regardless of action but if an agent *folds* the nettle leaves, they can be safely eaten ($\mu = 30$, $\sigma = 5$). $P = 10\%$, $n = 20$.

Termite nest: the best way to get a meal of termites is to *poke* the nest with a *peeled stick* ($\mu = 100$, $\sigma = 20$) — the termites swarm on the stick and the agent can lick it clean. Also a reward for *poking* with a *stick* ($\mu = 50$, $\sigma = 10$) or without any tool ($\mu = 20$, $\sigma = 5$). *Bashing* the nest with a *stone* or a *big stone* will get some termites ($\mu = 35$, $\sigma = 10$). $P = 10\%$, $n = 10$.

Coconuts: need to be *bashed* to have a chance of opening them and getting the milk ($\mu = 150$, $\sigma = 20$). $P = 10\%$, $n = 5$. Chance of successfully opening a coconut depends on whether *stones* or *big stones* are available in hand and on the ground.

	On ground		
	Nothing	Stone	Big stone
Holding			
None	0.01	0.01	0.01
Stone	0.2	0.6	1.0
Big stone	0.4	0.6	0.8

Monkeys: can be caught and eaten ($\mu = 200$, $\sigma = 20$) if an agent *chases* them, but only if a second agent is present to block their escape route. Alternatively an agent can *throw* a *stone* for a 10% chance of success, or *throw* a *stick*, a *peeled stick*, or a *big stone* for a 2% chance of success. $P = 10\%$, $n = 2$.

Thorns: are painful ($\mu = -20$, $\sigma = 5$) regardless of action. $P = 10\%$, $n = \infty$.

Beehive: there is a 20% chance that the bees will fly out and sting the agent ($\mu = -200$, $\sigma = 50$) regardless of action. $P = 5\%$, $n = \infty$. If the agent *shakes*, *pokes*, or *bashes* the hive the chance of stinging is 100%. The hardest trick to learn is getting honey from the hive ($\mu = 500$, $\sigma = 50$, $n = 1$): the agent must either *bash* the nest with a *stone* or a *big stone* for a 50% chance of success (and a 100% chance of getting stung) or smoke the bees out by starting a fire, i.e., *rubbing* two *sticks* of either type together (one on the ground and one in the hand).

Wasp nest: the same negative properties as the beehive but with no chance of reward. $P = 5\%$, $n = \infty$.

An important action not covered above is *scratching*, which leads to a modest reward ($\mu = 1$, $\sigma = 0$; $\mu = 2$ if the agent is carrying either kind of *stick*) regardless of the local resource. Scratching provides a kind of local optimum, in that it is easy for an agent to get locked into pursuing the modest rewards of this activity, perhaps under a tree, rather than moving around looking for more profitable resources.

Other actions include *peeling*, which converts a *stick* in the agent's hand into a *peeled stick*; *swapping*, which exchanges the tool on the ground, if any, for the one in the agent's hand, if any; and *travelling*, which moves the agent to the nearby resource they have most recently seen. Note that when an agent attempts an irrelevant action, such as bashing a tree, the payoff is always zero.

Tools are distributed generously, with 40 sticks, 20 stones, and 20 large stones present in the environment. Peeled sticks only appear if an agent peels a stick. Stones do not wear out, but sticks and peeled sticks break after 1000 or 500 time steps respectively; a replacement stick appears at a random location.

Reinforcement learning agents

The agent population is 25, but initially agents enter the world one at a time. Every 400 timesteps a new agent is born. When 25 agents are present, the oldest agent is always killed off to make room for the youngest; an agent's lifespan is therefore 10,000 timesteps. The middle-aged agent born 12 places ahead of a newborn is nominally regarded as its parent, and the new agent appears in the same square. The simulation is run for 100,000 time steps, or 10

agent lifetimes. Even though the simulation has "generations" of a sort, there is no genetic aspect to agent behaviour: each individual comes into the world as a blank slate and must learn for itself how to maximize payoffs.

Each agent makes its decisions on the basis of a history of reinforcement: it has learned to associate some states and actions with a reward and others with punishment. The different social learning mechanisms implemented all assume this underlying capacity for reinforcement learning. The specific learning algorithm used was Q-learning (Watkins & Dayan, 1992) which allows an agent to take into account not just immediate but also delayed payoffs. This would be important, for example, in realizing that a higher long-term payoff could be gained by picking up and peeling a stick rather than just poking the termite nest immediately. The learning rate was 0.1, the temporal discounting factor was 0.5, and the proportion of random actions was 0.2.

Q-learning in its simplest form does not include any generalization: if an agent learns that it is good to shake a berry bush when there is a stone on the ground, it will not automatically transfer that knowledge to the case where there is no stone on the ground. Given 60,000 state-action pairs¹ and a lifetime of only 10,000 timesteps it is clear that an agent could not exhaustively sample its space of perceived situations. We therefore included some basic generalization: after updating a state-action pair through Q-learning, all other state-action pairs that differed only in terms of the tools available, the presence of others, and the resource seen in the distance (exception: the resource being left behind for the travel ac-

¹10 resources \times 5 tool-on-ground options \times 5 tool-in-hand options \times 2 others-present options \times 10 next-resource options \times 12 actions = 60,000.

tion) also had their expected returns updated by a small amount in the same direction.

Implementing social learning

To find baseline levels for payoffs in this environment, we looked at the performance of agents with completely random behaviour, and at that of agents capable only of individual learning. We then implemented three of the simple mechanisms described by Noble and Todd (in press): following or stimulus enhancement, contagious behaviour, and emulation. In addition, we implemented a version of imitative learning. Finally, we looked at the performance obtained when following was combined with each of the other mechanisms.

Following: (otherwise known as stimulus enhancement) is implemented simply by having each agent occupy the same square as its parent for the first 25% of its lifetime. Beyond that, agents learn for themselves.

Contagious behaviour: if another agent is present, there is a 10% probability that the first agent will do whatever the second just did. An instinctive behavioural contagion is assumed, as when one person's yawning stimulates another to yawn; actions are equally likely to be copied regardless of their payoff.

Emulation: if another agent is present and obtains a non-zero payoff, the first agent adjusts its estimated return for *all* state-action pairs related to its current state. The idea is that the agent has become aware of an affordance in the local environment, but has no idea about what specific action is needed to exploit it.

Imitation: if another agent is present and obtains a non-zero payoff, the first agent is able to take the perspective of the second. The first agent notes the perceptual state of the other, as well as the action the other chose, and updates its own relevant state-action pair accordingly. The advance on emulation is that the first agent not only knows which action achieved the reward, but also can perceive the other agent's state as opposed to its own (e.g., it might be holding a stick).

Data collection

How can we collect data from this simulation that parallels the data ethologists would collect when observing a real species? We assumed that observers watched the population for the final 1000 timesteps; i.e., they had a snapshot of the group's behaviour rather than being able to observe the rise and fall of socially transmitted behaviours in the long term. We then looked at the average payoff per agent during these 1000 timesteps. We can imagine that the ethologists are observing distinct groups in similar environments and comparing the average energy intake across groups. It is natural to assume that a population equipped with a richer social learning ability will be able to extract more energy from its environment.

We also wanted to measure behavioural convergence, i.e., the degree to which the various members of the population tend to choose the same action when faced with the same situation. This was done by compiling a list of 1000 states perceived by members of the population (the state of a randomly chosen individual was noted at each timestep). Then all 25 members of the population were faced with each of the 1000 states in turn, and the proportion of individuals who chose the modal response for each state was recorded and averaged. We can imagine that the ethologists have an outstanding ability to set up naturalistic experiments in the field, and are able to engineer things so that each member of the population can be placed into the same series of states and allowed to choose an action.

We looked at behavioural convergence because uniformity of behaviour within a population, especially when combined with behavioural variation *across* populations, is usually taken as a sign that social learning is going on. For example, if one group is

observed to hunt monkeys by chasing them down, while another group tends to throw stones at them, then it seems to follow that each animal is not arriving at its own solution to the problem, but rather is being influenced by its fellow group members.

Finally, we looked at simple counts of the occurrence of various behaviours: the performance of more complex and hard-to-learn behaviours would probably be taken as a sign of sophisticated social learning. So, for example, if one species has mastered the trick of peeling sticks before poking them into termite nests, and another related species has not, one might assume that the former species has a better social learning faculty.

Results and discussion

All of the learning mechanisms, both individual and social, ensured that the agents did far better than they would by choosing a random action. Figure 2 shows that the mean payoff per agent per timestep was at least 6 units; the mean payoff for random behaviour was -8.997 units. Behaviours such as scratching, eating berries, folding nettle leaves, and poking termite nests (although not often with tools) were all quite common and allowed the agents to extract a good return from the environment.

Omitting for a moment the cases in which following is combined with another mechanism, we have the surprising result that the best mean payoff is for emulation, not imitative learning. Why does this occur? Emulation means that an agent increases its estimated return for *all* state-action pairs associated with any state in which it witnesses another agent receiving a positive payoff. Effectively this means that emulation encourages exploration around beneficial resources, and results in more inquisitive agents who are persistent enough

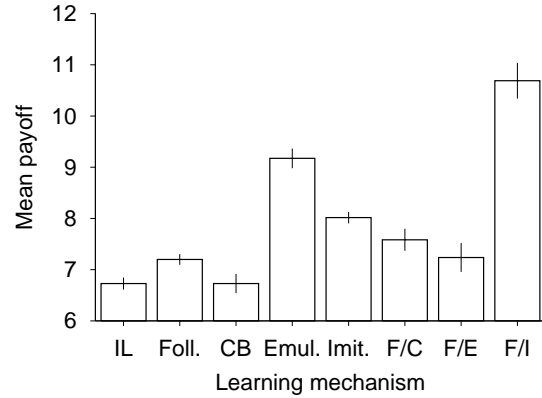


Figure 2: Mean payoff per agent per timestep across the different learning conditions: individual learning, following, contagious behaviour, emulation, imitation, following with contagious behaviour, following with emulation, and following with imitation. Each data point is averaged across 10 runs, and is shown ± 1 standard error.

to crack a tricky resource like nettle leaves. Imitators, on the other hand, are quite conservative. Imitative learning is sophisticated enough to allow an agent to pick up any clever trick it might witness, but imitators appear to be caught in a vicious circle in which the only role models they observe are other conservative imitators (see the behavioural profiles in Figure 4).

Figure 2 also shows that the simple matter of following your parent around for the first quarter of your life is enough to increase your mean payoff significantly. This is more impressive than it seems, as the imposition of following behaviour threatens to *decrease* payoffs by increasing feeding competition: a resource will run out twice as fast if it is being consumed by both parent and child. Contagious behaviour, on the other hand, did not do any better than individual learning. It may be that occasional instinctive copying is unhelpful because agents already try random actions 20% of the time and will thus reach the

same result in the end.

When following behaviour is combined with other mechanisms, the story becomes more complex. It is obviously not the case that following and the other three mechanisms mesh together in a simple additive fashion: the efficiency of imitation and contagious behaviour is increased, but the efficiency of emulation is decreased. Clearly it is important to imitate or emulate the right people: an experienced parent in one case, and a random cross-section of society in the other.

Behavioural convergence (Figure 3) was 0.55 even in the individual learning case. The convergence score for random behaviour was inevitably much lower at 0.18, and thus we are reminded that similar behaviour may be brought about by similar learning experiences rather than any form of social learning. Convergence was highest in the two imitation conditions, which makes sense. However, the low convergence for emulation belies the fact that these agents were doing well in terms of mean payoff, and convergence rates are the same across the two imitation conditions, despite the fact that it is only in the following / imitation condition that high pay-offs are achieved. An observation of high or low behavioural convergence, which might lead one to suspect the presence or absence of effective social learning, is meaningless considered alone.

Figure 4 shows frequency counts for several behaviours across the different conditions. It shows that different learning mechanisms lead to quite different behavioural profiles, e.g., emulators spend a lot of time folding nettle leaves, whereas imitators have mastered the art of shaking berry bushes. It is not simply the case that the more “advanced” the social learning mechanism, the more behaviours are added to the repertoire.

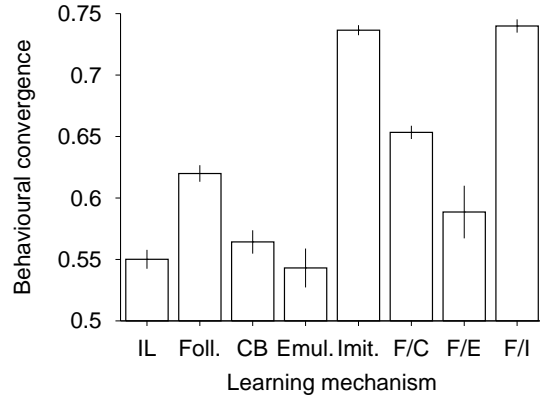


Figure 3: Behavioural convergence across the different learning conditions. Each data point is averaged across 10 runs, and is shown ± 1 standard error.

Another surprising finding is that some behaviours that would seem to suit imitative learning, such as learning to peel a stick and then poke it into a termite nest, are in fact more often performed by agents with simpler mechanisms (Figure 4, bottom). Agents who follow their parents are the ones who do this most often. Even so, it seems that this behaviour was largely too hard: note the low frequency counts overall. Imitators could certainly have picked up this trick, but again it seems they did not have access to the right role models.

Conclusions

We hope to have demonstrated that imitative learning is not always going to be the best strategy for a group-living animal, and that the effectiveness of any social learning strategy will depend on details of the environment and additional facts about the animal’s lifestyle. We readily admit that the impressive performance of emulation in our model, for example, is not likely to be a general result, but is tied to the particulars of the simulated environment and the

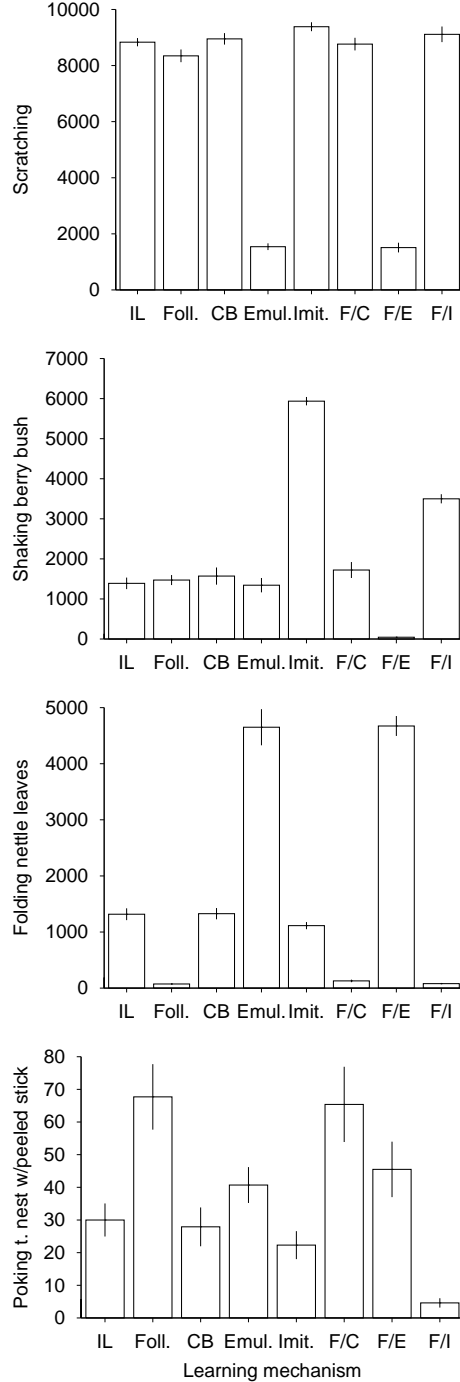


Figure 4: Frequencies with which various behaviours were observed across the different learning conditions: from top to bottom, scratching, shaking berry bushes, folding nettle leaves, and poking termite nests with peeled sticks. Each data point is averaged across 10 runs, and is shown ± 1 standard error.

underlying reinforcement learning system. But that is exactly the point.

The argument which inspired this paper (Noble and Todd, in press) had two aspects: first, simpler social learning mechanisms are inherently more likely to evolve; one reason for this is that complex mechanisms such as imitation will involve significant costs. It follows that we should look for simple mechanisms in nature. Second, in cases of doubt about which mechanism is at work, from an epistemological point of view it is good practice to start by proposing simpler mechanisms. The current simulation results show that an even stronger argument is possible: even if imitation was available via a cost-free, one-step mutation, it still might not be the best learning mechanism for a given species.

A focus on imitation as the only kind of social learning worth having is misplaced. Researchers interested in imitation as a potential tool for instructing robots or software agents, for example, might do well to consider giving their creations more spartan mechanisms, which have the additional benefit of being far easier to program. If the agent is equipped with a reinforcement learning system, it would be a good start to simply have it follow a more experienced agent (e.g., a person).

We believe our model also shows that ethologists comparing how well two different groups of animals get on in similar environments, or looking at data on “traditions” or “proto-cultural behaviour” (i.e., behavioural convergence) may find that data to be quite misleading as to the underlying social learning mechanisms at work. It is not the case, for example, that low behavioural convergence necessarily means an absence of social learning. One has to be careful in looking at exactly how behaviours are transmitted from one individual to another. Theories about the adap-

tive value of social learning, such as Laland et al. (1996), are useful, but ideally we want to move beyond such theories and look at hypotheses about the utility of a particular mechanism in a particular environment.

We deliberately omitted an evolutionary dimension from the current model as we wanted to look at the effects of learning alone. However, future work could incorporate an evolutionary aspect, in which different social learning mechanisms were selected for over generational time. The mean payoff results certainly suggest that this would not be a straightforward story of hill-climbing with imitation at the peak. Let us assume that a single-step mutation F controls following behaviour, and a two-step mutation leads from individual learning to emulation (S_1) and thence to imitative learning (S_2). This latter assumption is reasonable as imitative learning can be seen as a “sharpened” version of emulation. For an evolving population of individual learners, the final strategy will be emulation if S_1 occurs first, as F and S_2 will both be associated with a loss in fitness. On the other hand, if F occurs first, both S_1 and then S_2 will mean increases in fitness, and the population will end up as follower / imitators.

It would be useful to move away from our simplistic lookup-table implementation of Q-learning and use a neural network as a more natural way of generalizing across perceptual states. We are also interested in investigating whether infinite negative pay-offs, i.e., actions that result in the death of the agent, push an evolving population towards imitative strategies. Finally, we would like to use the same basic technique described here to produce a high-fidelity model of social learning in a particular species, e.g., chimpanzees, and thereby comment on which mechanisms are plausibly being used by that species, much as

Noble et al. (2001) do for Norway rats.

References

- Box, H. O., & Gibson, K. R. (Eds.). (1999). *Mammalian Social Learning: Comparative and Ecological Perspectives*, Vol. 72 of *Symposia of the Zoological Society of London*. Cambridge University Press, Cambridge.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. The University of Chicago Press, Chicago.
- Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*. MIT Press, Cambridge, MA.
- Heyes, C. M., & Galef, Jr., B. G. (Eds.). (1996). *Social Learning in Animals: The Roots of Culture*. Academic Press, San Diego, CA.
- Laland, K. N., Richerson, P. J., & Boyd, R. (1996). Developing a theory of animal social learning. In Heyes & Galef (1996), pp. 129–154.
- Noble, J., & Todd, P. M. (2002). Imitation or something simpler? Modelling simple mechanisms for social information processing. In Dautenhahn, K., & Nehaniv, C. (Eds.), *Imitation in Animals and Artifacts*, pp. 423–439. MIT Press.
- Noble, J., Todd, P. M., & Tuci, E. (2001). Explaining social learning of food preferences without aversions: An evolutionary simulation model of Norway rats. *Proceedings of the Royal Society of London: Biological Sciences*, 268(1463), 141–149.
- Schaal, S. (1999). Is imitation learning the route to humanoid robots?. *Trends in Cognitive Sciences*, 3(6), 233–242.
- Tomasello, M. (1996). Do apes ape?. In Heyes & Galef (1996), pp. 319–346.
- Toquenaga, Y., Kajitani, I., & Hoshino, T. (1995). Egrets of a feather flock together. *Artificial Life*, 1(4), 391–411.
- Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8, 279–292.
- Whiten, A. (1998). Imitation of the sequential structure of actions by chimpanzees *Pan troglodytes*. *Journal of Comparative Psychology*, 112(3), 270–281.