

Chapter Twelve

Individual, Social and Evolutionary Adaptation in Collective Systems

Evert Haasdijk, A.E. Eiben, Alan F.T. Winfield

12.1 Introduction

This chapter focusses on adaptivity as a pivotal enabler of future robotic systems. It is *the* fundamental premise of our vision that future robots will have to be capable of autonomous adaptation, that is, able to change their control systems without human intervention. This vision has also been articulated by Nelson *et al.*, cf. [Nelson *et al.* (2009)], who phrased it as follows:

Advanced autonomous robots may someday be required to negotiate environments and situations that their designers had not anticipated. The future designers of these robots may not have adequate expertise to provide appropriate control algorithms in the case that an unforeseen situation is encountered in a remote environment in which a robot cannot be accessed. It is not always practical or even possible to define every aspect of an autonomous robot's environment, or to give a tractable dynamical systems-level description of the task the robot is to perform. The robot must have the ability to learn control without human supervision.

To define **adaptation** –“learning control without human supervision”– clearly, consider a robot's controller as a process that maps inputs, read from the robot's sensors and internal states, to outputs, typically actuator and state settings. Adaptation is then defined as any changes to this mapping process, including the setting of its parameters.

According to this definition, changing the output threshold on some artificial neural net controller constitutes adaptation because the mapping from in- to outputs changes, but varying outputs due to some internal state does not, because the mapping remains the same, even though behaviour changes. Adaptation is necessarily on-line and without human intervention: the robot controller changes

on-the-fly, as it goes about its tasks. We can distinguish two stages in the robot life-cycle: design time and operational time, separated by deployment. In these terms, adaptivity amounts to changing robot controllers autonomously during operational time. There are various optimisation and design techniques based on adaptive systems, e.g., evolutionary algorithms, particle swarm optimisation, neural networks, etc., that can outperform traditional methods. Such techniques can be used, and often are to great effect, during design time to find (near-)optimal robot controllers. However, these adaptive techniques fall outside of the scope of this chapter if the controllers remain static after deployment.

The vision that underlies this chapter is that adaptivity is a necessary feature in collective robotic systems to cope with a number of fundamental challenges:

- (1) **Unforeseen environment** The environment where the robots operate may not be fully known during the design process. Therefore, the robot controllers at the time of deployment are only approximate solutions that need to be adapted to the real requirements during operational time.
- (2) **Changing environment** The environment may change to such an extent that the given skill set of the robots is not adequate anymore. In a robot collective this environment might include the robots' social environment. Hence, controllers must adapt to the new situation.
- (3) **Reality gap** Even if the environment is known in advance and is not changing during operational time, it is very likely that the design process is based on approximations and simulations of the real operational conditions. Hence, the robot controllers have to be fine-tuned after deployment.

In this chapter we elaborate on the notion of adaptation and place adaptive systems into one conceptual framework, called *Population-based Adaptive Systems* (PAS). The notion of PAS serves as the unifying concept and the name PAS as an umbrella term. Within this framework we further distinguish different types of adaptation. One of the fundamental distinctions we make is based on differentiating learning and evolution. In turn, this is based on distinguishing phenotypes and genotypes regarding robot controllers [Eiben *et al.* (2010)]. Simply put, this distinction means that:

- We perceive the controllers with all their structural and procedural complexity as phenotypes.
- We introduce a (typically structurally simpler) representation of the controllers as genotypes.
- We define a mapping from genotypes to phenotypes, that might be a simple mapping or a highly complex transformation.

For example, a robot controller may consist of two artificial neural nets (ANNs) and a decision tree, where the decision tree specifies which ANN will be invoked to produce the robot's response in a given situation. This decision tree can be as simple as calling ANN-1 when the environment is lit and calling ANN-2 when the environment is dark. This complex controller, i.e., phenotype consisting of a decision tree and two ANNs, could be represented by a simple genotype of two

vectors, showing the weights of the hidden layer in ANN-1, respectively ANN-2. A technical distinction between learning and evolution is now straightforward if we postulate that learning acts at the phenotypic level, while evolution only affects the genotypes.

This chapter is structured as follows. Section 12.2 establishes a framework that identifies three main forms of adaptation (evolution, individual learning, and social learning) in the context of population-based adaptive systems, ranging from artificial life systems to robot swarms. Section 12.3 presents a case study carried out in a system where individual learning and evolution are combined in such a way that they can directly influence each other, rather than acting independently on the agent/robot population. We demonstrate that in such a system learning – that optimises for the benefit of the individual – can effectively kill the population by ignoring the group level benefits of reproduction. In Sec. 12.4 we investigate social learning as a mechanism to disseminate ‘knowledge nuggets’ –bits of adapted controller– in a population of agents/robots. Thus we show how the results of individual learning efforts (that would normally disappear if the individual dies) can be kept. In other words, here we demonstrate how social learning can facilitate the emergence of a knowledge reservoir in a population. While the experiments reported in these sections were conducted in an artificial life setting, the conclusions are just as pertinent to robot swarms that implement combinations of these forms of adaptation. Section 12.5 presents a case study of robots implementing social learning through embodied imitation. Here we show that (real) embodiment gives rise to variation in socially learned behaviours (which we refer to as memes). In a robot collective we have the three evolutionary operators: variation, selection (because robots choose which memes to enact) and inheritance (because we see inherited characteristics in n-th generation memes). This, we suggest, demonstrates the possibility of population-level social learning in collective robotic systems through memetic evolution. Finally, Sec. 12.6 concludes the chapter.

12.2 Population-based Adaptive Systems

We coin the phrase “*Population-based Adaptive Systems*” (PAS) to label systems such as robot swarms or artificial life systems that have adaptive behaviour at agent and/or population level. Such systems can be characterised by two essential features:

- A group of basic units (agents or robots) that can perform actions, e.g., computation, communication, interaction, etc. By acting, these units exhibit behaviour – individual behaviour at unit level, as well as collective behaviour at the group level.
- The ability to adapt at individual and/or group level. If the exhibited behaviour is generated through behavioural rules¹ inside the units, then adaptation implies that these rules change. For instance, a change can take place

¹We do not necessarily mean a set of IF-THEN rules, but any representation, including, for instance, neural nets, decision trees, etc.

298 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

inside an existing unit by replacing an existing rule by a new one, or a change can take place on population level by creating a new unit with a new set of rules.

There is a large variety of PASs with quite different examples. For instance, a peer-to-peer computer system where each node (peer) is able to improve its workings through experience, a genetic algorithm seeking an optimal solution to the travelling salesman problem, a group of learning robots collectively gathering red rocks on Mars, or a simulation of socio-economic processes by means of adaptive agent society. Such systems have received increasing interest over recent years with an increasing number of related papers. However, the lack of a common underlying framework of terminology means that the presentation of related problems and solutions shows a large (application dependent) variation. This forms an obstacle for identifying similar concepts, problems, solutions, etc. over various publications and implies the risk that individual researchers reinvent the wheel. A common conceptual framework describing a large class of PASs forms a helpful stepping stone towards further developments in the area.

We introduce the notion of Population-based Adaptive Systems and identify related concepts and research issues in this section. We focus our study on a class of PAS where adaptation occurs through three fundamental adaptation mechanisms: evolution, individual learning and social learning.

In the remainder of this section, we present a conceptual framework that captures a wide class of adaptive systems and identify research issues of general relevance in PAS.

12.2.1 *Three Tiers of Adaptation*

We use an agent-based metaphor, where the group of basic units is perceived as a population of agents (be they software agents or robots) whose behaviour is controlled by themselves — subject to environmental constraints, of course. That is, we assume that each agent has a controller that takes observations regarding the environment and the agent's internal state as input and generates actions to be executed by the agent as output. Furthermore, we assume that two levels of change can occur:

- (1) Changes at agent level: the controllers of the agents can change;
- (2) Changes at population level: it is possible to delete existing and to add new agents. In common parlance, this amounts to birth and death in the system.

As mentioned above, we see adaptation as change of controllers in a population of agents and distinguish three fundamentally different adaptation mechanisms. Denoting the set of all possible controllers by C , we can perceive adaptation mechanisms in PASs as search algorithms traversing the space C in a volume oriented manner – maintaining a population of controllers $P = \{c_1, \dots, c_n\} \subset C$ simultaneously. Adaptation or learning then amounts to taking search steps, moving from the presently given set P of controllers to a new set P' and we distinguish adaptation on agent level (cf. property 1) and adaptation on population level (cf. property

2). We will call these *lifetime learning* and *evolution*, respectively. Furthermore, we make an additional distinction between two types of lifetime learning. In *individual learning*, an agent adapts its controller through a purely internal procedure, not through some oracle or other agents. If agents adapt their controllers by communicating controller information to each other and incorporating (good) pieces of knowledge from each other, we speak of *social learning*. Figure 12.1. illustrates this taxonomy and the corresponding terminology.

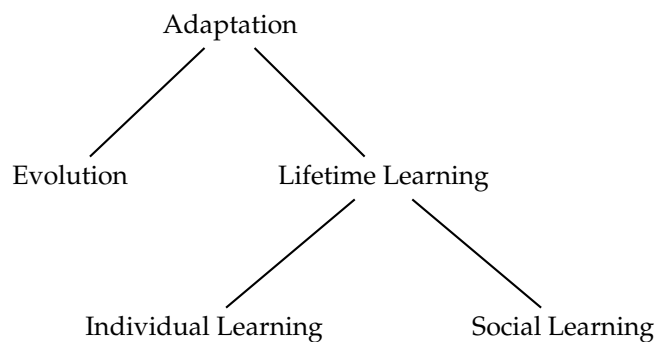


Figure 12.1. Taxonomy of adaptation mechanisms in PASs

To delineate this framework, consider a few examples. (1) A genetic algorithm solving the Travelling Salesman Problem has birth and death, but the agents (individuals, candidate solutions) do not have a controller because they are not supposed to *do* anything other than producing offspring. Reproduction, moreover, is not actively controlled by the individuals themselves. Rather, they undergo it, arranged by an “oracle”—the outer loop of the evolutionary algorithm procedure. Thus, in this example we have no controllers and changes occur only at population level. (2) In embodied evolution [Ficici *et al.* (1999); Watson *et al.* (2002)], the robots broadcast (possibly mutated) genes at a rate proportional to their fitness (measured as the number of batteries collected). Robots also resist “infection” at a similar rate. A good individual, collecting many batteries by virtue of its superior controller will infect many others before being replaced (i.e., infected) itself. If we see infection as death and immediate replacement, the robots in such a system do not adapt individually and changes occur at population level. (3) As a third example consider a single Web-agent serving a single user by selecting news items every morning using some given set of rules that are continuously improved through reinforcement learning. Here, the agent does have a controller (the rule set) that can change, but the population is a singleton and there is no death – no changes at population level. (4) Finally, consider the AEGIS artificial life system [Buresch *et al.* (2005); Eiben *et al.* (1999)], where a population of agents exists in an artificial habitat. The agents can move, eat, mate, fight, etc. as determined by their controllers and they undergo adaptation of their body characteristics (by evolution from generation to generation) and their controllers (by evolution from generation to generation or by learning during lifetime). In this system, we have controllers

300 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

and changes occur at both individual and population levels.

As an example of a PAS with adaptation through evolution as well as individual and social learning, we consider the *NEW TIES* system², which we will describe before we elaborate on the three adaptation mechanisms (Sec. 12.2.3) and their interactions (Sec. 12.2.4) and research challenges these raise. Note, that *NEW TIES* serves as an example only and that, although we describe many design choices that were made for this particular system, the interactions between adaptation mechanisms that we describe are not specific to this example implementation and mostly do not depend on the design choices described.

12.2.2 The Environment and the Agents

The *NEW TIES* system provides a simulation platform in which a cultural society develops through evolution, individual learning and social learning of autonomous agents [Gilbert *et al.* (2006)]. The artificial, virtually embodied agents that make up this artificial society live in a grid world containing objects such as food sources (plants), tokens, places and building bricks.

In this world, time passes in discrete steps. Every time-step, the agents receive stimuli regarding objects (including agents) that they see or carry, messages from other agents that they hear and their internal state (e.g., their own energy level). The agents process these stimuli to select actions such as move or turn, pick up or put down objects, eat, communicate or interact otherwise with other agents (e.g., mating, or giving or taking objects to/from other agents). To process these inputs and arrive at a decision about which action to take, the agents use their individual controllers.

The project models agents anthropomorphically, thereby imposing strict autonomy, (virtual) embodiment and situatedness. This limits our options when designing agent interactions (e.g., agents cannot communicate unless they are within each other's vicinity), perception (e.g., they cannot see inside each other's heads) and learning mechanisms (e.g., no supervised learning).

Agents have to husband their energy: performing the selected action, even if that amounts to inactively surviving a time-step, costs energy. Should an agent run out of energy, it dies. To gain energy, an agent must eat food (plants). Other than that, agents die when they reach a certain maximum age. There is no other selection mechanism: as long as an agent lives, it can act—and therefore, engage in mating or social learning. To gain energy, an agent must eat food (plants). The laws of nature governing the environment determine the preconditions and the results of actions, e.g., specifying the amount of energy a plant yields when eaten and the costs of movement, the maximum lifetime for agents, or a minimum age and energy level at which agents can mate. Agents decide on their actions using a controller. In other words, the controller is the decision making unit inside an agent that maps inputs, i.e., perceptions of the agent regarding the world and its own internal state, to outputs, i.e., the agent's action.

²New and Emerging World models Through Individual, Evolutionary and Social learning (*NEW TIES*), EU FP6 Project, <http://www.new-ties.org>

12.2.2.1 Decision making and agent controllers

At every time-step, the agent processes the incoming information and describes the situation it finds itself in in terms of concepts. Then, based on this description, the agents decides on an action to perform.

Categorisation and conceptualisation To reduce the dimensionality of the observation space (the raw data where attributes are the elementary attributes of all possible entities in the world), a process of categorisation and conceptualisation map it onto another space, where the attributes are the so called concepts. Raw data is aggregated in two steps. First, it is aggregated to form categories that are then further aggregated to concepts. The incoming information is processed by categorising the raw data-bundle of features. Each feature concerning objects in the world, like color or shape, can be regarded as an axis in the features space; a category is defined by a range of possible values within the whole range of a feature. For example, for the feature colour everything between 1, ..., 10 could be considered green. Concepts are (multi-dimensional) entities composed from (one-dimensional) categories. For instance, plants could be the green and triangular objects while agents could be pink and circular. Concepts are stored in an agent's ontology and are used to provide a characterisation of a given situation on a higher level than the original raw data.

Decision making The agents' controllers are implemented as special kinds of decision trees, *decision Q-trees* (DQTs). The 'Q' refers to the fact that they can be adapted through Q-learning [Sutton and Barto (1998a)], the NEW TIES implementation of individual learning. With crossover and mutation operators inspired by those used in genetic programming [Koza (1992)], these trees can also be adapted through evolution when two agents mate to create offspring.

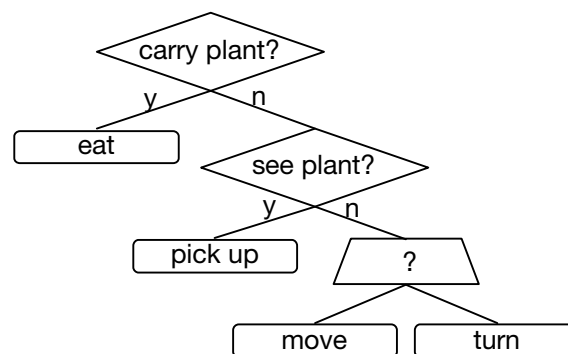


Figure 12.2. A simplified example of a decision Q-tree (DQT).

DQTs consist of test, bias and action nodes (Fig. 12.2.; depicted as lozenges, trapezoids and rounded rectangles, respectively).

A **test** node evaluates a Boolean query based on concepts known to the agent, e.g., "Is there a plant ahead?" or "Is there an agent nearby?", and depending on

302 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

the answer (Yes or No) the tree is further traversed through either of the two child nodes. Thus, a full path from the root to a leaf (an action to be performed) node forms a conjunction of statements that together provide a partial situation description in terms of the agents' concepts.

To traverse a bias node, the agent probabilistically selects one of multiple branches for further traversal – each of these branches has a bias that determines the likelihood of it being selected. These biases are determined genetically through evolution and onto-genetically through individual and social learning.

The leaves of the DQT are action nodes that select an action. Action nodes, like bias nodes, are probabilistic: the actual action is stochastically chosen according to a weight distribution over all possible actions. The available actions are simple actions – such as move, turn-left or turn-right –, unary – such as eat(x) or hit(y) –, and binary actions such as give(a,o). The arguments for the higher arity actions are implied by the tests that were traversed to select an action –e.g., testing for visible agents implicitly selects all agents in sight– and can be any object, but if, e.g., an agent attempts to eat a non-food item, this action will fail in the world.

12.2.3 *Adaptation Mechanisms*

As outlined in Sec. 12.2, we envision adaptation as the change of controllers in a population of agents. In *NEW TIES*, this amounts to changing DQTs. In this subsection we discuss how the general trinity of adaptation is instantiated in *NEW TIES*. To begin with, we note that all three adaptation mechanisms work in the same search space – that of all possible DQTs.

12.2.3.1 *Evolution*

NEW TIES deliberately adopts a non-Lamarckian notion of evolution[Lamarck (1809)], so inheritable material cannot change during an agent's lifetime. This means that an agent created with a controller c seeds its descendants by exactly this controller c , regardless of any changes brought by lifetime learning.

The two pillars of evolution are selection and variation; variation is realised by straightforward tree-crossover and tree-mutation operators, much as in genetic programming. Viewing adaptation as search through the space of controllers, one elementary search step in this context amounts to combining two existing controllers c_1 and c_2 into a new one c_3 .

It is an essential aspect of this system that selection is not based on some objective function to be maximised [Menczer and Belew (1996); Mitchell and Forrest (1994)]. Survivor selection is strongly environmental: agents die if they run out of energy or reach the maximum age. As for parent selection, an agent can decide any time to mate (subject to some constraints). If the controller chooses to mate, the agent selects itself as a would-be parent. To procreate, it needs to find and “convince” another agent: it sends a special message, a mate proposal. Only if the other agent accepts this mate proposal do the two agents become actual parents and produce a child. To make the child viable, each parent donates a portion of its current energy, consequently incurring a cost.

The *NEW TIES* evolutionary system differs from usual evolutionary algorithms in a number of essential aspects.

- (1) Fitness is not an a priori utility measure that determines the number of offspring. One could say there is no notion of fitness at all, or rather, that in *NEW TIES* fitness is a secondary, observable measure determined by the number of offspring rather than vice versa – a truly Darwinian notion.
- (2) Reproduction is not orchestrated by some central authority. Individuals autonomously and asynchronously decide to mate.
- (3) Reproduction is detached from survivor selection. Newly produced individuals can be added to the population without removing old ones. Likewise, an individual can die without being replaced by a new one. As a side-effect, there is no clear definition of a generation here.

These properties have two prominent consequences. Firstly, in the absence of an explicit objective function the selection probabilities (that embody the system bias for quality) must be based on indirect quality indicators. In general, the age and the energy level of agents can be used here: an agent that survives for a long period and/or has accumulated much energy must be well adapted, hence worthy of being reproduced. In this respect, PAS of this kind are closer to natural selection than, for instance, Genetic Algorithms where selection probabilities are calculated from an objective function.

The second effect is that points 2 and 3 imply a kind of reproduction – “*natural reproduction*” – where the population size inherently changes over time. Users of such systems face a tough challenge concerning the calibration of the system to avoid unlimited population growth (explosion) or complete extinction (implosion). In a particular system, such as *NEW TIES* or *AEGIS*, ad hoc solutions can work, based on balancing energy supply (number of plants, energy of plants, reproduction rate of plants) and energy consumption (costs of actions). From a general evolutionary point of view, population size can be controlled by tuning the selection mechanisms. For instance, the parameters specifying the minimum age or energy required for mating. At the moment, there are no general guidelines or design heuristics available to cope with this problem.

12.2.3.2 Individual Learning

A newborn agent, and with it individual learning, starts with the controller that is provided by (one of) its parents. The most appropriate learning type for individual learning is reward based: supervised learning is difficult, because agents can be in an environment of which the most optimal (set of) action(s) is unknown. Unsupervised learning is inappropriate, because information present in the environment is wasted if not used as feedback for learning.

NEW TIES implements reward based individual learning as reinforcement learning [Kaelbling *et al.* (1996); Sutton and Barto (1998a)]. Reinforcement learning can change the DQT by policy change. An agent’s *policy* is –in the context of reinforcement learning– represented by its DQ tree. Any path in the DQT leading to an action is a result of the policy. Policies can be altered by changing the values of the

304 | Individual, Social and Evolutionary Adaptation in Collective Systems

edges that change the likelihood of taking a specific path. *NEW TIES* uses SARSA, one of the variants of Q-learning[Sutton and Barto (1998a)].

In *NEW TIES*, the reward is usually based on energy, but other types of ‘currency’, e.g. something based on emotions or some mix of simpler currencies, are possible. The currency must in any case be accessible to the agent, or the agent would not be able to use it for computing rewards. Such a mixture is probably needed for the problem described in Sec. 12.2.4.1, where agents would unlearn to reproduce if reward is only based on energy – this is investigated in detail in Sec. 12.3.

An important challenge for reinforcement learning is that the state-space created by the perceptual input is huge. To illustrate, the state-space for the visual field is $\#typeOfObjects \times \#gridcells \times VisualField$. Given that *NEW TIES* has at least 3 types of objects and that the visual field is 50 grid cells, it is obvious that the state space is very large, probably intractably so. Moreover, the state space is further extended by non-visual perceptual input of auditory, internal and reproductive stimuli. To cope with the size of this state-space, it is not partitioned by the input stimuli, but by the tests in the test nodes of the DQT. The tests in a test-node test for certain concepts, for instance *green agent*. This divides the state space in agents that are green and all other coloured objects. The test-node uses the input, only testing for particular aspects of the environment.

12.2.3.3 Social Learning

Many studies have focussed on social learning with approaches including imitation (as described in Sec. 12.5), copying behaviour as well as using socially provided corrective feedback [Dautenhahn and Nehaniv (2002); Acerbi *et al.* (2008)]. In *NEW TIES*, by contrast, agents communicate explicitly and social learning entails an agent modifying its controller by incorporating a piece of knowledge it receives from another agent. Social learning requires at least two agents a_1 and a_2 with controllers c_1 and c_2 ; one search step amounts to changing c_1 into c'_1 (assuming that a_1 learns from a_2), where c'_1 is some combination of c_1 and c_2 .³

Agents communicating in this manner implies a multi-faceted set of features and parameters that govern issues such as (social) networks of knowledge exchange, levels of trust and relative merit of knowledge, etc. In general, they concern:

- when and with whom to exchange knowledge;
- the selection of knowledge to send or elicit;
- when and how to accept offered knowledge.

Obviously, a general consideration when designing these features is including a bias for quality. In other words, at least some of the choices involved in importing a “knowledge nugget” from another agent must favor learning from a better agent. Similar to introducing a bias for quality in evolution (cf. Sec. 12.2.3.1), the age and the energy level of agents can be used as quality indicators here. Apart from any

³Remember the non-Lamarckian nature of *NEW TIES*’ reproduction: these controller changes do not affect the genetic material (which in effect is a copy of the initial controller with which an agent is created).

specific quality-driven social learning scenario, there is always qualitative pressure as described in Sec. 12.2.3.1: agents with poor controllers die sooner and therefore cannot participate in social learning exchanges (“teach”) as often as agents with good controllers.

Note, that communication introduces a “social dimension”; an overlay network in technical terms. The properties of this network depend on the given implementation, but in general, the network changes over time. In *NEW TIES*, this is realised by a protocol similar to gossiping in peer-to-peer systems. Every agent maintains a (fixed length) list of acquaintances – agents it has seen and talked to before. This list is updated with new observations (encounters with other agents) using a FIFO policy. The construction and maintenance of this social network can also be influenced by quality indicators of peers.

A knowledge nugget in our system is represented by a sub-DQT (extracted from the sender’s controller). In the current implementation, this sub-DQT is included in the tree of the receiving agent by inserting –at some appropriate location in the DQT– a bias node that has two children: the foreign sub-DQT and the already existing native sub-DQT. These alternatives are weighted by newly defined biases based on the ratio between the sender and recipient’s age and energy levels.

Section 12.4.2 provides a more detailed description of the social learning mechanism in *NEW TIES*.

12.2.4 Relationships Between Adaptation Mechanisms

To position evolution, individual learning, and social learning it is helpful to consider them from the knowledge transfer perspective, where knowledge is seen as (good) pieces in the agent controllers. From this point of view, knowledge is transferred vertically by evolution, down along the line of descendants. (Recall the note from Sec. 12.2.3.1 that we do not have a clear notion of generations here, because agents residing on different levels of the family tree can live at the same time in the same population.) On the other hand, individual learning is a sink: in the absence of social learning, individually accumulated knowledge simply disappears when the agent carrying it dies. Social learning can alleviate this, since it amounts to horizontal knowledge transfer, passing knowledge nuggets within the current population. In this respect, social learning makes the population into a knowledge reservoir, reducing (at least potentially) the risk that knowledge must be rediscovered over and over again.

12.2.4.1 Evolutionary and Lifetime Learning

A marked distinction between evolution and lifetime learning is that evolutionary operators do not change the controllers of agents during their lifetime, while lifetime learning operators obviously do. If evolution were the only adaptation mechanism, agents would die with the controller they were born with. Hence, evolution does not take place on an individual, but strictly on a population level. From this perspective, the death of an agent represents a contribution to the evolution pro-

cess, because the population adapts with each death.⁴ This is particularly not the case for individual learning, where the death of an agent terminally destroys the results of the learning process.

In our example, evolution also differs from lifetime learning in the entity that initiates a learning step: individual learning and social learning steps are initiated without the influence of the agent's controller – by an oracle, or subconsciously, if you will. This is not the case for evolution search steps, because the agent has to decide itself to reproduce by sending or accepting a mate proposal. As a compelling consequence, agents can unlearn reproduction through lifetime learning because the individual reward for mating is negative: it costs energy without any mitigating personal benefit. To counteract such tendencies, one can introduce some specific reward for mating (orgasm), make mating a subconscious process or take population-level benefits into account in lifetime learning.⁵ Section 12.3 investigates this consequence in detail.

Memetic algorithm research has pointed out positive interactions between evolution and lifetime learning, by showing that combinations of evolution and individual learning are particularly beneficial[Krasnogor (2002)]. An interesting and promising interaction between evolution and lifetime learning is described in [Best (1999)]. This study finds that lifetime learning decreases the need for evolution to get it spot-on: the chance of finding the optimal solution is much greater with lifetime learning and evolution combined.

12.2.4.2 *Individual and Social Learning*

As noted above, the non-Lamarckian nature of evolution in NEW TIES entails that knowledge that an agent acquires through individual learning cannot affect inheritable material, and therefore is lost when that agent dies. By proliferating knowledge over the population of agents, social learning preserves such knowledge pieces that would otherwise disappear. Thus, social learning turns the population into a reservoir of (individually acquired) knowledge.

A system that combines individual learning and social learning can be thought of as having division of labour: individual learning generates novel knowledge nuggets and social learning disseminates these. Social learning can also be seen as an accelerator making the system more efficient. Think, for instance, of agent a_1 learning x , agent a_2 learning y and a_1 and a_2 telling x and y to each other, rather than having to learn these knowledge pieces themselves. In general, efficiency improves if the costs of, respectively, time needed for and learning through communication, are lower for the agents than the costs/time of acquiring knowledge individually – an assumption that holds in a great many systems. As a net effect, combining social learning and individual learning allows agents to possess knowledge regarding situations they never encountered themselves, acquired at greater speed and at lower costs. Such constellations have been shown to outperform either adaptation mechanism by itself, e.g. in [Bull *et al.* (2007b)].

⁴Supposedly changing for the better, cf. survival of the fittest.

⁵Taking a learning step in both individual learning and social learning could also be made into a conscious action, in which case similar considerations would apply.

Section 12.4 investigates this interaction in detail.

12.2.4.3 Individual and Social Learning as Evolution

Recall from Sec. 12.2.3.3 that knowledge nuggets are sub-DQTs. Incorporating such sub-DQTs into an agent's controller amounts to an operation similar to crossover in GP. Similarly, one can see an analogy between a learning step in individual learning and a GP mutation operator: both turn some controller c into c' . From this perspective it is quite natural to see the combination of individual learning and social learning as an evolutionary process. Similar observations can be found in e.g. [Bull *et al.* (2007b)], [Smith *et al.* (2000)] and [Richerson and Boyd (2005)].

The selection components for this evolutionary system consist of the mechanisms regulating when two agents engage in sending/receiving knowledge pieces (parent selection)⁶ and the policies to accept and incorporate a received piece of knowledge (survivor selection).

It should be noted that this constitutes an evolutionary process quite different from the one described in Sec. 12.2.3.1. The most visible difference lies in the replacement strategies: in the lifetime learning-based evolutionary process, reproduction and survivor selection *are* coupled: a new controller, whether made by mutation or crossover, immediately replaces an existing one: its parent and the population size remains unaffected. Another difference is that here, a new controller is created by either crossover (social learning step) *or* mutation (individual learning step), while in evolution this happens by crossover *and* mutation (which occurs sequentially in the reproduction procedure). Furthermore, we should note that here we *do* have an explicit fitness measure, used in at least some parts of the system. For the parent selection component this is not necessarily the case; an agent can perform a mutation (do an individual learning step) regardless of the quality of its present controller c – making c the parent of the new c' – and the same holds for an agent a_1 deciding to talk to a_2 – selecting their controllers c_1 and c_2 as would-be parents. We can distinguish two cases of survivor selection: in the case of mutation (an individual learning step), survivor selection does not involve fitness either: the old c (the controller being improved by individual learning) is simply deselected and replaced by c' (the improved controller). However, if a new controller is created by crossover (an social learning step), a utility function is used to determine the relative merit of the received knowledge when integrating it with the already known c_1 to create the new c'_1 . This utility is related to the relative ages and energy levels of the two agents involved.

Section 12.5 examines social learning as evolution in a robot collective. In that case study the property of the population that evolves is the set of socially propagated (by imitation) behaviours, which we refer to as memes (equivalent to knowledge nuggets); thus we have a PAS of memes. Memes are replicated (by robot-robot embodied imitation), mutate (because of imperfect fidelity imitation), are selected (because robots choose which memes to enact) and have inherited characteristics across several generations of memes. In this population the overlay network (Sec. 12.2.3.3 is provided by the physical position and proximity of robots,

⁶Combined with Darwinian survival of the fittest as described in 12.2.3.1

which of course changes as the robots move. Thus here we see an evolutionary process at the behavioural or meme-level, that we can label memetic evolution⁷.

Considering individual learning and social learning in this light raises two prominent research questions. First, how does existing evolutionary computing knowledge, e.g., regarding variation, selection and their balance, translate into these contexts? Second, how do the two evolutionary processes, genetic evolution on the one hand, social learning and individual learning on the other, interact in one system?

12.2.5 Discussion

Most of the technical details we introduced are merely illustrative in the sense that they do not restrict the generality of our discussions. Using trees to represent agent controllers is one such detail. Our line of thought about variation operators in evolution and merge operators in social learning can be repeated for other data structures as well. A similar argument holds for the categorisation and conceptualisation mechanism to pre-process sensory input of the agents; the general point here is dimensionality reduction. This is critical when using reinforcement learning algorithms, because they scale very badly with the size of the state-space, but this aspect is likely to occur in many systems.

The main contribution of this chapter is the definition of a system where three different adaptation mechanisms – genetic evolution, individual learning, and social learning – can work simultaneously, yet clearly distinctly. The separation of the learning mechanisms is based on a differentiation between inheritable and learnable agent characteristics.⁸ Designating agent characteristics as inheritable or learnable is one of the major design decision when implementing PASs. Inheritable properties can undergo evolution through appropriate variation operators and environmental selection, learnable properties can undergo lifetime learning through individual and social learning. By the clear separation between evolution, individual learning, and social learning, particular adaptation mechanisms can be switched on and off independently, thus allowing research on their effects separately or in various combinations. This allows us to gain insight in their mutual effects on each other and on the adapting population. Research in this area offers great benefits by the high potential of “fully powered” adaptive systems. In general:

- Social learning can act as an accelerator for individual learning in each agent and can preserve the individually discovered knowledge nuggets for the population that would otherwise be lost after the death of the individual that learned them;
- The combination of individual learning and social learning can be seen as an evolutionary system, creating an opportunity to use existing knowledge in evolutionary computing when designing such combined systems.

⁷Memetic evolution is distinct from memetic algorithms [Krasnogor (2002)] since, in memetic evolution only the memes (which we can think of as extended phenotypes) evolve, the genotypes do not.

⁸In the system we described, these are the same, but our considerations are still valid if this is not the case.

The specific choices concerning evolution in *NEW TIES* are reflected in our treatment of evolution. In particular, we focus on systems with natural reproduction, cf. Sec. 12.2.3.1. In many applications, e.g., ALife, social simulations, peer-to-peer systems, this is the obvious choice of reproduction scenario, so we can safely state that the subset of PAS with natural reproduction is large and interesting. Considering such systems we observed that:

- In an evolutionary process relying on natural reproduction, population size is inherently volatile. This creates a tough challenge for designers and users of such systems: to design (selection) mechanisms that prevent explosion and implosion of the population;
- While, in general, combining lifetime learning and evolution is a powerful combination (cf. memetic algorithms), in PAS with natural reproduction lifetime learning can counteract evolution by unlearning mating.

In many instances of PAS, one is mainly interested in emergent phenomena, particularly in emerging behaviour and emerging structures, such as the controllers of the agents (world models) or the social network generated by social learning, or the emergence of ‘traditions’ in the socially learned behaviours across the population. It is characteristic that the experimenters can influence system properties only indirectly, via the adaptation mechanisms. Given some demanding world where agents only survive if they adapt to the particular challenges of that world, the experimenter’s task is to engineer an appropriate mix of the adaptation mechanisms so that these mechanisms will generate the desired emergent behaviours and structures. It is this aspect that makes understanding the trichotomy of evolution, individual learning, and social learning crucial to apply them successfully in any PAS, be it *NEW TIES* or a robot swarm.

12.3 Learning Benefits Evolution

This section considers the interplay between two of the three levels of adaptation introduced in Sec. 12.2, namely evolution and individual learning. Combinations of evolution and learning have been investigated before [Belew and Mitchell (1996)], cf. the hundred years of the Baldwin effect [Turney *et al.* (1996)]. Prominent related work can be found within memetic algorithms, or hybrid evolutionary algorithms [Moscato (1999); Krasnogor (2002)], evolutionary robotics [Nolfi and Floreano (1999); Ijspeert *et al.* (1998)] and ALife [Todd and Miller (1990); Belew *et al.* (1990); Munroe and Cangelosi (2002); Curran and O’Riordan (2006); Buresch *et al.* (2005)].

As described in Sec. 12.2.3.1, the combination of features in *NEW TIES* implies that the population size can change, even to extinction. This property is typically absent in related work⁹, even some work that claims to model natural systems [Ruppin (2002)], although it is evident that in nature populations can and do die out. Past research has focussed on the costs and benefits of learning in evolu-

⁹Research on predator-prey phenomena is not usually concerned with combinations of evolution and learning.

310 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

tion [Hinton and Nowlan (1987); Mayley (1996); Munroe and Cangelosi (2002); Nolfi and Floreano (1999)] and on identifying factors that influence this relationship [Mayley (1996); Nolfi and Floreano (1999)]. This section continues research in this direction, but specifically in the context of a changing population size.

Remember that in *NEW TIES*, the evolutionary mechanism is under the control of the agents, because it is the agents themselves who decide if and when to create offspring. This means that the development of agent controllers (through evolution and/or learning) can lead to intensively reproducing agents or just the opposite: the evolutionary mechanism itself is subject to changes over time.

Evolution and individual learning act in a common search space: that of the set of all possible agent controllers. Hence, an agent can be born with controller C , created by some evolutionary operators applied to its parents' controllers, and can change C into C' , C'' , etc., during its lifetime through individual learning. Evolution is non-Lamarckian: when this agent reproduces, only its original controller C is used for creating a child, any individually learned modifications in C' etc. are disregarded as inheritable material. As noted in Sec. 12.2.3.2, individual learning is implemented as reinforcement learning. In essence, reinforcement learning changes the controller by regulating agent preferences for actions based on a reward system. It is important to note that reinforcement learning can strengthen or weaken preferences for any agent action, including the mating action required for offspring creation. Thus, it is possible that individual learning unlearns reproduction and effectively counteracts evolution.

The questions we have to ask ourselves, then, are these:

- (1) What is the effect of adding individual learning through reinforcement learning?
 - On the viability of the population?
 - On the performance of the population?
 - On the evolutionary engine?
- (2) How does this depend on the rewards used by reinforcement learning? In particular:
 - When rewards are energy-based.
 - When rewards are hard-wired by the user.

12.3.1 *The Experiments*

As noted above, the system is not meant to set the agents any specific task other than to win the struggle for life. The environment can, of course, be set up to challenge the agents in specific ways. The agents then have to deal with these challenges in order to survive and prosper. In other words, an experimental set-up in *NEW TIES* represents a particular challenge or learning task that agents must solve through adaptation.

In the experiments we describe here, the environment is set up so that agents can only survive if they successfully tackle the well-known poisonous food problem [Cangelosi and Parisi (1998); Nolfi and Parisi (1995); Todd and Miller (1990)].

The agents find themselves in an environment where there are two types of plants, both of which can be picked up and eaten. One type is nutritious and yields an energy increase, the other type is poisonous and eating them actually drains energy. Agents can choose not to, but they can distinguish between the two types of plant. They do not, however, know *a priori* that one kind – let alone which kind – is poisonous. Because agents *must* eat to replenish their energy level as mentioned above, they have to learn to disregard poisonous food if they are to survive.

In these experiments, the world is a 200×200 grid, initialised with 500 agents, 8,000 edible plants and 10,000 poisonous plants. There is a maximum to the number of agents: agents are unable to reproduce when this limit is reached, but it may be exceeded through the concurrent creation of a number of new-born agents. Agents and both types of plants are randomly distributed over the grid. We call our atomic time step a day and 365 days a year; the minimum mating age for agents 1,000 days: i.e., they cannot successfully reproduce for the first 1,000 days of their lives. The maximum age for agents is 7,300 days (7.3 times the minimum mating age), when they reach this age they die, whatever their energy level. The experiments run for 30,000 days. Initially, agents are assigned a random age between zero and one year. The initial controller of all agents is the same; in this controller some behaviours are pre-wired¹⁰, like looking for food. However, the behaviour for eating the correct type of food is not present. This can be acquired by changing the tree structure and/or tuning the biases of bias nodes and action bias nodes, although the probability that the latter succeeds is small in the tree-structure of the initial controller. Evolution (without sub-tree mutation) and individual learning are the only active adaptation mechanisms; social learning is turned off.

12.3.1.1 Measurements

To answer the research questions we must measure the viability and performance of the population and provide insights into the evolutionary engine.

To measure the viability, we use the population size. A run is considered successful if the population size did not reach zero during that run.

As a behavioural performance measure we use a function based on the ratio of the different types of food the population eats:

$$g(t) = \frac{\sum_{t-1}^t eat_h}{\sum_{t-1}^t eat_p + \sum_{t-1}^t eat_h} \quad (12.1)$$

Where $\sum_{t-1}^t eat_h$ and $\sum_{t-1}^t eat_p$ are the number of wholesome and poisonous plants eaten by the population between $t - 1$ and t .

Additionally, we measure the total and average energy of the population and the total and average age.

To measure the performance of the evolutionary system we monitor the average number of mate-agreements of the population.

¹⁰Pre-wired is not the same as hard-wired: pre-wired controllers can be modified by the adaptation mechanisms.

12.3.2 Experiment I

In the first experiment, poisonous plants drain twice the energy that an edible plant yields. We ran two sets of experiments; one where individual learning was either turned off or used only energy-based rewards and one where we introduced a specific reward for reproduction. The results are summarised in Fig. 12.3..

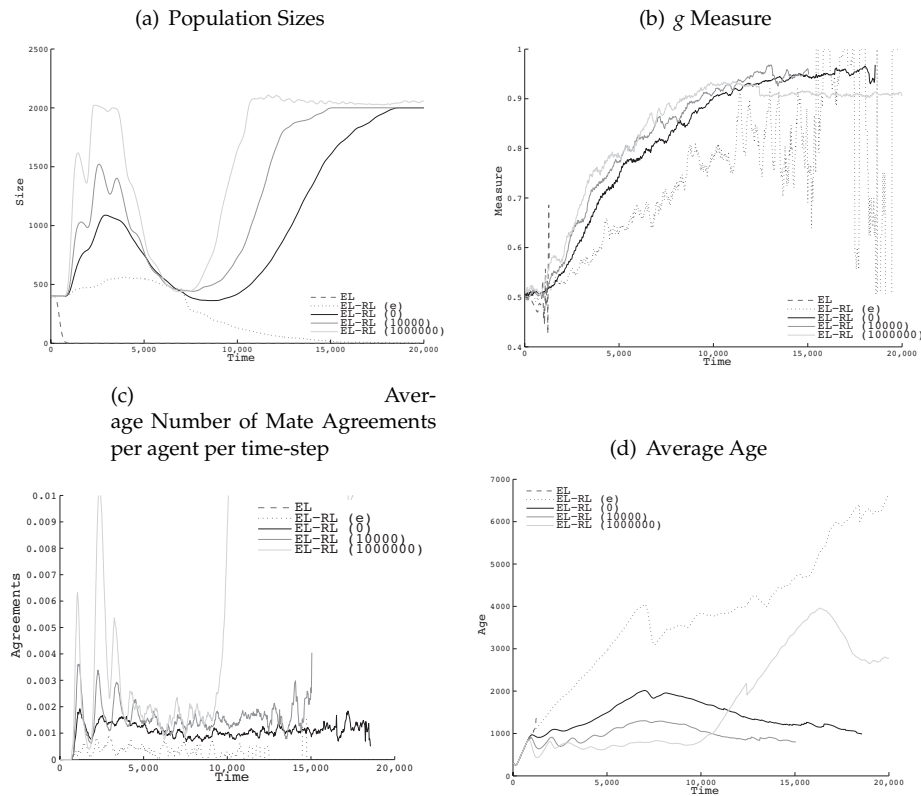


Figure 12.3. Results for experiment I; graphs compare results for evolution by itself (EL), combined with reinforcement learning based on energy (EL-RL (e)) and combined with reinforcement learning with different “orgasm” levels (EL-RL (0;10,000;1,000,000)).

12.3.2.1 Evolution only and Evolution-reinforcement learning combination with energy based rewards

Figure 12.3.(a) shows clearly that evolution only (indicated by the dashed line “EL”) survives for approximately 1,000 time steps and thus does not yield a viable population. Adding energy based reinforcement learning to evolution markedly improves viability, as can be seen in Fig. 12.3.(a) (the dotted line “EL-RL (e)”). In the long run, however, this is not a viable solution, because after 15,000 time-steps the population is as good as extinct.

So, the combination with energy based reinforcement learning is thus unable to make a population viable. This might be because reinforcement learning is un-

learning reproduction, since it costs energy and therefore produces negative rewards. The rewards for other actions, except the eat action, are also negative, but usually not as bad as reproduction, because that costs a third of the agent's energy. The "EL-RL (e)" curve in Fig. 12.3.(c) proves that reward based reinforcement learning is unlearning reproduction, because the total number of mate-agreements steadily decreases. Moreover, Fig. 12.3.(d) indicates that agents do not reproduce enough to sustain the population: agents reproduce once every 3,000 time-steps, while the average age in the population is only 1,000.

There are two reasons why agents reproduce at all in the face of the negative reward. Firstly, they have to try to reproduce at least once to learn its negative effects. Secondly, during exploration agents can still choose the mate-agreement action, even if they unlearned it. The periodic behaviour of the curve is a side-effect of the setting of the ages of the initial population and the minimal reproduction age.

12.3.2.2 *Combination of evolution and reinforcement learning with a hardwired reward*

The results in the previous subsection suggest that reproduction is unlearned or becomes so rare that the agent population is unable to sustain itself. To test the explanation that this is due to agents receiving negative rewards for reproduction, we introduce a special reward for reproduction. Its only role is to make mating actions attractive, so it can be regarded as a kind of pleasure or orgasm. We ran experiments with three levels of reward: 0, 10,000 and 1,000,000.

The most striking result is that a hardwired positive reward renders the population viable. Note, that even a reward of zero works because all other actions except eating yield a negative reward.

In terms of population performance, the results show that higher rewards for reproduction result in better performance. For instance, in the g measure graphs (Fig. 12.3.(b)), the curve for a reproduction reward of 1,000,000 increases much more steeply than for a value of 10,000, indicating that the population learns to avoid poisonous food very quickly.

The intensity of the evolutionary engine is measured by the number of mate-agreements, displayed in Fig. 12.3.(c). The general trend is that the higher the mate-reward the higher the number of mate-agreements.

Note that in all different simulations, including that of evolution alone, the g measure is similar for the first 1,000 time steps. This means that the combination of reinforcement learning and evolution is unable to learn the task during this period, implying that individual learning somehow keeps agents alive that would die in the case of evolution alone. To find out how agents were able to survive, we analysed the results by tracking the agents' actions. This analysis showed that agents often choose to do nothing (the NULL action). Agents thus learn to save their energy. This suggests a *hiding* effect: individual learning preserves agents with a non-optimal strategy [Mayley (1996)].

12.3.3 Experiment II

To test whether a hiding effect occurs as suggested above, we change the environment so that evolution alone can make the population viable. The only change from the previous experiments is that the levels of nutrition and poisonousness have been set so that a poisonous plant drains an equal amount of energy that an edible plant yields. We run experiments with evolution only and with both evolution and reinforcement learning. The average results over 10 runs are shown in Fig. 12.4..

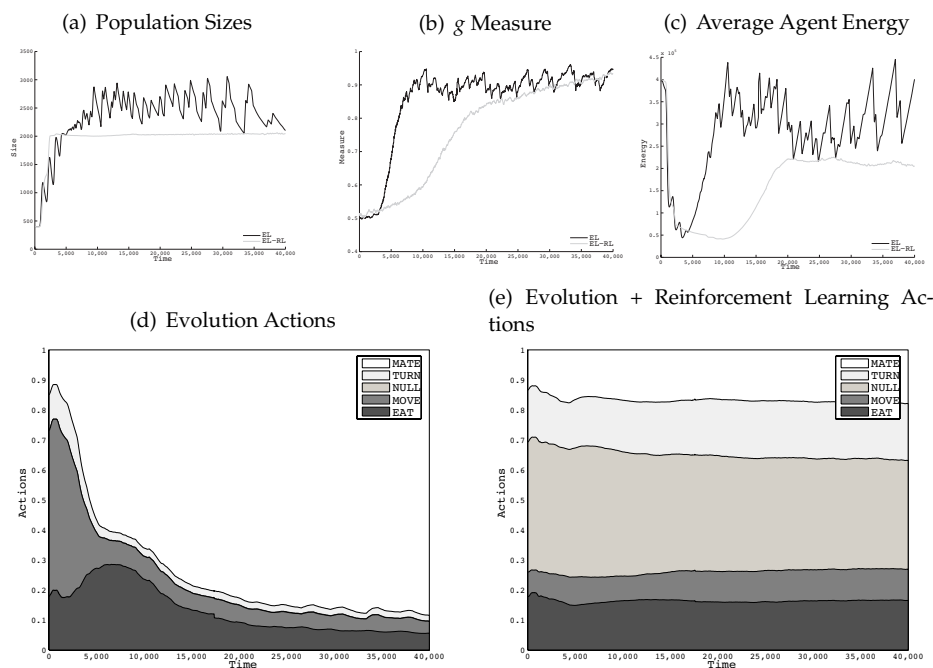


Figure 12.4. Results for experiment II; graphs (a), (b) and (c) compare results for evolution by itself (EL) and combined with reinforcement learning (EL+RL).

Both the combination of evolution and reinforcement learning and evolution by itself yield viable populations in this set-up. The soft cap on population size causes some boundary effects such as the fluctuating population size and the decrease of energy levels at some point.

The main result, however, is that there clearly is a hiding effect: the combination of evolution and reinforcement learning hides the ill-adapted nature of non-optimal agents. The population with only evolution very rapidly learns to eat only wholesome plants (Fig. 12.4.(b)) and therefore accumulates much more energy than the combination of evolution and reinforcement learning (Fig. 12.4.(c)).

With evolution only, the agents accumulate enormous reserves of energy so that they can get by without any food; this changes the evolutionary pressure from eating (and preferring edible plants) to reproduction: we see the evolution of agents that only perform actions involved with reproduction.

Figure 12.4.(e) clearly shows that the average number of mate-agreements is much lower with the combination of evolution and reinforcement learning than with only evolution; reinforcement learning apparently hinders evolution. The difference in number of mate-agreements already appears within the first 5,000 time steps, while there is no appreciable difference in g value or population size to explain the difference. One possible explanation is that the combination of evolution and reinforcement learning creates another type of agent that doesn't reproduce as often because while evolution is mainly focussed on reproduction, the agents with reinforcement learning try to balance between both eating and reproduction in order to maximise their rewards.

12.3.4 Discussion

Over the years there has been a fair amount of research into combinations of learning and evolution, in particular regarding their costs, benefits [Hinton and Nowlan (1987); Mayley (1996); Munroe and Cangelosi (2002); Nolfi and Floreano (1999)] and factors that influence this relationship [Mayley (1996); Nolfi and Floreano (1999)]. We now (re)consider these issues in a context where agents:

- decide autonomously if and when they reproduce (*natural reproduction*, implying a dynamic population size);
- can adapt their controllers to unlearn the mating action.

Our experiments show that in such systems learning can counteract evolution. To be concrete: with a straightforward reward system based on energy, reinforcement learning will cause the agents to lose interest in mating because of the high individual costs. Hereby the group benefits (maintaining the evolving population) are lost. This effect can be counteracted by introducing a specific reward for the mating action that gives positive feedback to the agents, regardless of the related energy costs. One could of course argue that this trick is known in nature, commonly called an orgasm. All in all, this indicates that we must consider the reward for reproduction as another factor that influences the effect of learning on evolution in addition to the list proposed by Mayley [Mayley (1996)].

In terms of the viability and performance of the population, our experiments show that learning can quite literally be a matter of life and death. In our first scenario, evolution by itself was not powerful enough to sustain the population. Adding reinforcement learning changed this, yielding populations that survive and prosper until the end of the simulations. Simply put: learning keeps the population alive. It can do so by creating controllers that minimise energy expenditure, a non-optimal behaviour, in the sense that such agents do not learn to eat the correct plant type. This is one of the costs of learning: learning causes a clear hiding effect because it allows non-optimal controllers to survive. By contrast, evolution by itself optimises by harshly cutting out the bad agents, but always with the risk that there is no population left. In a system allowing a changing population size this can be lethal.

Further research could show whether there is an optimal value for the reproduction reward (i.e., the extent of "pleasure" during mating). A good value would

not frustrate evolution and still make a population viable when needed. One possibility is to make this value self-adaptive by adding it to the genome, allowing evolution to tune itself.

12.4 Social Learning as Enabler of a Knowledge Reservoir

As mentioned in Sec. 12.2.3, agents in the NEW TIES PAS decide autonomically on the actions they perform by means of a controller that is inherited (for the initial population: generated) at birth. They implement evolution and reinforcement learning for individual learning. Through evolution, only the inherited controller is passed on (i.e. *non-Lamarckian* evolution [Lamarck (1809)]): agents do not inherit knowledge (modifications to the controller) that their parents may have gained through experience; they can only recombine the controllers that their parents had at birth (with some mutation added). This means that, without some additional method of spreading knowledge through the population of agents, everything an agent learns through experience (i.e., through individual learning) will be lost when that agent dies.

This is where social learning comes into play: with social learning in place, anything an agent learns during its lifetime can be taught to other agents, so that this knowledge does not necessarily die with the agent that originally discovered it. With agents exchanging knowledge pieces –bits of adapted controller– through social learning, the population as a whole effectively becomes a knowledge repository – although not a randomly accessible one for individual agents – for individually discovered adaptations. Obviously, social learning can also speed up the learning process at the population level as found in e.g., [Acerbi and Nolfi (2007); Denaro and Parisi (1996); Bull *et al.* (2007b)]

Social learning can only play this role if it can effectively disseminate individually acquired knowledge pieces. The question, then, that we seek to answer is the following:

Is social learning an efficient mechanism to spread knowledge pieces through the population, thus creating a knowledge repository for individually acquired knowledge?

In nature, social learning can be achieved through a host of mechanisms ranging from imitation to social guidance in individual learning [Acerbi and Nolfi (2007)]. Here, we consider the case where social learning consists of agents actively suggesting behavioural rules (knowledge pieces) for the consideration of other agents in a peer-to-peer fashion. The recipients of these knowledge pieces then choose whether or not to integrate them into their own set of rules. The fact that all agents participate in social learning on an equal footing implies an inherent parallelism in the spreading of knowledge pieces: all agents that have acquired a knowledge piece can simultaneously share it with other agents, who can then share it in turn, and so on.

Cultural algorithms employ belief spaces [Reynolds (1999)], which can be seen as explicit knowledge repositories that the individuals build collectively. In the re-

12.4. Social Learning as Enabler of a Knowledge Reservoir | 317

search presented in this subsection, however, knowledge repositories are formed implicitly by the population and any individual agent can use only that part of the repository it embodies. It has been shown that social learning through imitation (sometimes called ‘cultural evolution’) can be beneficial by decreasing the learning time for individuals, particularly in cases where the required behavioural rules are difficult to acquire [Acerbi and Nolfi (2007); Denaro and Parisi (1996)]. Such implementations of social learning typically focus on a limited number of ‘experienced’ individuals instructing uninitiated individuals one by one and thus do not exploit the inherently parallel ink-stain effect present in the peer-to-peer knowledge exchange that we envisage. Similarly, in ensembles of learning classifier systems, social learning – termed ‘rule-sharing’ – has been shown to boost the learning speed [Bull *et al.* (2007b)] of the ensemble. Comparing such ensembles with a population of interacting, mortal agents is tenuous, however: the constituent parts of the ensembles are not considered separately, only the performance of the ensemble’s aggregated behaviour is taken into account.

12.4.1 Energy and Agent Quality

As mentioned in Sec. 12.2 and contrary to typical evolutionary algorithm or evolutionary robotics applications [Eiben and Smith (2003); Nolfi and Floreano (2000)], the PAS we use as an example lacks a crisp optimisation criterion as well as a concrete task to be performed optimally. The agents survive whatever the environment throws at them or they do not—that’s all there is. This also entails that there is no measure of fitness in this system: the only selection mechanism is—truly Darwinian—the struggle for life in the environment: environmental selection.

To gauge their relative quality, agents can, however, be compared in terms of their perceivable attributes such as age or energy level. Crucially, such comparisons cannot be performed by some central selection mechanism – as would be the case in traditional evolutionary algorithms – but by the individual agents themselves when they *autonomously* decide to mate, engage in social learning, or otherwise interact with another agent.

12.4.2 Social Learning in Detail

Social learning is implemented in a push model, where teachers volunteer knowledge pieces that the students then may accept¹¹. Alternatively, social learning can be implemented in a pull model, where agents request knowledge from other agents. A combined model, where agents advertise that they believe that they have useful knowledge to share and other agents can then request that knowledge (similar to the plumage concept in [Smith *et al.* (2000)]) could be implemented as well. Social learning as described here uses a measure of relative quality $R(a, b)$ (described below) that compares agents a and b in terms of energy and age, but could have used, for example, a reputation-based measure just as well.

Generally, this subsection describes the implementation of social learning within NEW TIES—alternative design choices could be made and implemented at

¹¹Section 12.5 presents an example of this form of social learning by imitation.

318 | Individual, Social and Evolutionary Adaptation in Collective Systems

every level described here. As mentioned above, however, some options are infeasible because of the anthropomorphic nature of agents in these experiments. For instance, agents have to be within range ('earshot,' if you will) to be able to communicate and hence engage in social learning.

Social learning is implemented in the following sequence for every agent at every time-step:

- (1) An agent chooses to initiate sending ('teaching') probabilistically ($p = 0.2$).
- (2) If it decides to send, the agent describes the trace through its DQT that led to the current action (e.g., "I'm moving because there is no food to pick up").
- (3) Of all the agents in range, the teacher then selects the one with the lowest energy as the 'student'.
- (4) When an agent receives a knowledge piece, it stochastically chooses to integrate ($p = 0.2$) or disregard it.
- (5) When an agent s incorporates a DQT path P it received from an agent t , agent s selects the most similar path P' in its own DQT according to the following criteria:
 - (a) Percentage of matching tests
 - (b) The number of tests P but not in P'
 - (c) The number of tests in P' but not in P

If the percentage of matching elements in P is 100%, the bias for the action that P results in is multiplied with the relative quality $R(t, s)$ (see below). Otherwise, the agent engages in a kind of dialectic: it inserts a bias node at the first point of divergence between P and P' . The remainder of P' is inserted as one option at that node, a sub-tree corresponding to the non-matching entries in P is inserted as the alternative. The biases for the options are set proportionally to the relative quality $R(t, s)$. Figure 12.5. illustrates this procedure.

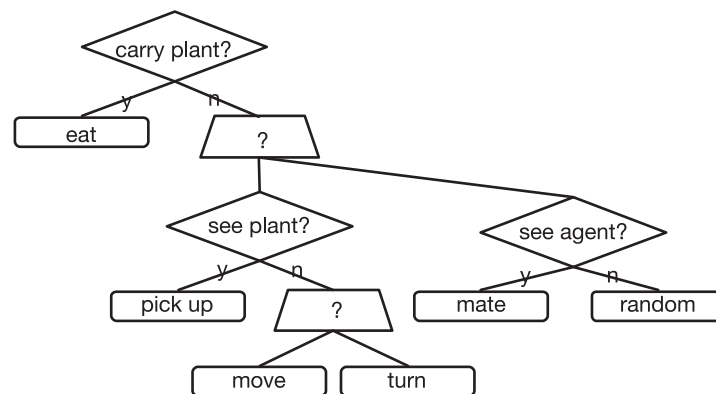


Figure 12.5. The result of integrating the path [not carry plant; see agent] \Rightarrow mate into the DQT from Fig. 12.2..

As described above, this social learning implementation requires some measure of (relative) quality for agents to be able to assess the merit of received knowledge

12.4. Social Learning as Enabler of a Knowledge Reservoir | 319

pieces when incorporating those pieces. To that end, an agent a can determine the relative quality $R(a, b)$ of another agent b from their relative ages A_a and A_b and energy levels E_a and E_b , respectively:

$$R(a, b) = 0.5 \cdot \left(\frac{A_a}{A_a + A_b} + \frac{E_a}{E_a + E_b} \right)$$

This measure ranges from 0, where agent b devastatingly outperforms agent a to 1, where the converse is true. If the agents have the same energy and are equally old, $R(a, b)$ equals 0.5. Note that this measure does not constitute an optimisation criterion as typically used in evolutionary algorithms: it does – without specifying any goal – allow for the comparison of the success of adaptation of individuals.

Social Learning as an Evolutionary Algorithm

[Smith *et al.* (2000)] already showed that an agent-based knowledge exchange mechanism similar to social learning constitutes an evolutionary algorithm. Moreover, as pointed out in [Eiben and Smith (2003)], an evolutionary algorithm requires:

- Selection as a force to push quality;
- Variation operators to create the necessary diversity and thereby create novelty.

This implementation of social learning achieves the former of these at various levels. Firstly, ill-adapted individuals tend to die relatively quickly, and hence cannot further distribute their knowledge, while well-adapted individuals tend to survive and have ample opportunities to distribute their knowledge. The second level is that of student selection mentioned above: when an agent has to choose between potential recipients of a knowledge piece, it selects the one with the lowest energy. Finally, the integration mechanism uses the relative quality $R(a, b)$ to set the bias for already known or newly received knowledge.

Variation is provided by the knowledge integration mechanism, which can be seen as a guided adaptation of crossover such as commonly used in genetic programming. Although this suffices, individual learning and social learning dovetail very nicely in this respect (as well as because of the benefit that we expect from social learning providing a knowledge repository for individual learning): individual learning then plays the part of a mutation-like variation mechanism.

The notion of social learning as an evolutionary process is further developed in Sec. 12.5.

12.4.3 Experimental Set-up

In this section, we –or rather, the agents– revisit the poisonous food challenge described in Sec. 12.3.1 where agents have to learn to avoid poisonous food and eat only healthy food.

To measure the efficacy of social learning as a mechanism for the proliferation of knowledge pieces through a population (i.e., for the population as a whole to adapt from individually learnt adaptations), we ran a series of experiments where

320 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

the population consists of two kinds of agents: knowers and students. The knowers have pre-built controllers that allow them to tackle the poisonous food problem. The students have a partially randomly constructed controller—they know how to pick and eat plants (regardless of their being poisonous or not) and how to mate, but the rest of their DQT is constructed randomly. A varying proportion of the agents with pre-built controllers can send, but not receive social learning messages ('teachers'), while students both send and receive social learning messages. The remaining knowers do not engage in social learning in any way; they are only there to ensure that the environment contains the same amount of agents eating away at the wholesome plants across the experiments, so that the results are comparable.

Another difference between students and knowers is that the former can mate to produce offspring where the latter cannot. Note, that this does not – in these particular experiments – constitute evolution: there is no variation operator because it does not entail recombination, but cloning of either parent. Therefore, there is no evolution at play to disturb our measurements. Neither kind of agent can perform individual learning in these experiments.

This set-up serves as an idealised exemplar of a population where some agents – represented by the teachers – have discovered, through individual learning or otherwise, a particularly useful bit of knowledge: to eat only wholesome plants. Note, that these teachers play quite a different role from the 'experienced individuals' employed by [Acerbi and Nolfi (2007); Denaro and Parisi (1996)]: from the students' point of view, they are no different from any other agent they encounter. We ran the experiment with varying numbers of teachers to compare the rate at which the population of students learns to differentiate between nutritious and poisonous food.

In our experiments, the agents can move in a 200×200 grid. There are initially 250 students and 100 knowers, of which 0, 1, 5 or 50 individuals are teachers. Agents can live well beyond the length of the experiments, so agents can only die of lack of energy. Each experiment was repeated 20 times. Poisonous plants drain 1.5 times the energy that wholesome plants yield, the environment is initialised with 16,000 plants of each type. Plants regrow practically immediately (within 2 time-steps), even if they've been picked, similar to food in SugarScape [Epstein and Axtell (1996)]. Thus, there is always food (and poison) available and the ratio poisonous/wholesome plants remains more or less at the initial value of 0.5.

To quantify behaviour, we use the g measure introduced in Eq. 12.1 – the ratio between wholesome and poisonous plants eaten. We also employ a structural measure that actually detects the presence of the required knowledge. There are, of course, many different strategies that allow the agents to eat only wholesome plants—e.g., "only pick up wholesome plants and eat anything you carry", or "drop any poisonous plant and eat anything you still carry". In these experiments, however, we know exactly which knowledge piece we expect to find because it is the relevant trace through the handcrafted knowers' DQT: it's [carry wholesome plant] \Rightarrow eat. This allows us to identify, during a run, those students that have incorporated this rule by inspecting their DQTs. Thus, we can

measure the incidence among the students of the appropriate knowledge piece.

Note, that the measurements we present here were taken only over the population of students.

12.4.4 Results

Figure 12.6. shows the development over time of $g(t)$ – averaged over 20 runs – for the students with 0, 1, 5 and 50 teachers. For reasons of legibility we omitted error bars; the 4 curves do differ considerably, although the standard deviation for 0 and 1 teacher is large, due to the fact that in many of these simulations, the students didn't eat at all.

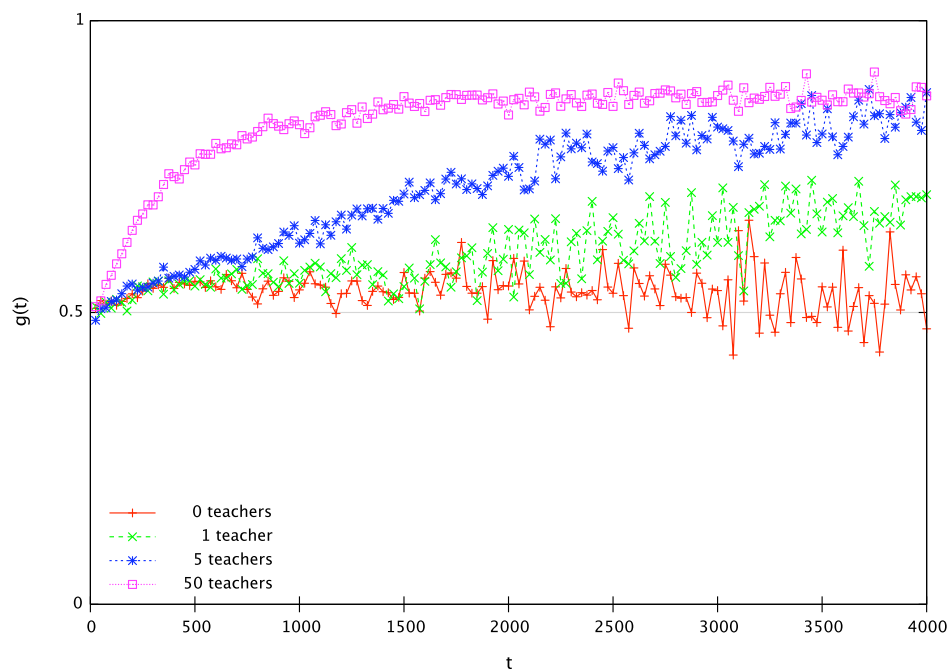


Figure 12.6. Development over time of $g(t)$ –the ratio between wholesome and poisonous plants eaten– for the student population for different numbers of teachers.

As can be seen, $g(t)$ remains level just above 0.5 for 0 teachers – there is no learning at all – the slight improvement over fully random behaviour is due to environmental selection: agents that eat too much poisonous food simply die at a faster rate than agents that do not or less so, leaving a slightly better set of surviving agents. In the case with a single teacher, the performance of the students increases substantially: even from so small a seed, a knowledge repository can grow. For 5 and 50 teachers, the population behaviour improves rapidly until $g(t)$ reaches a plateau between 0.8 and 0.9—there is no substantial difference between these experiments after that point. This seems to imply that in both cases the population of students becomes saturated – at least at a behavioural level – with the appropriate knowledge piece.

322 | Individual, Social and Evolutionary Adaptation in Collective Systems

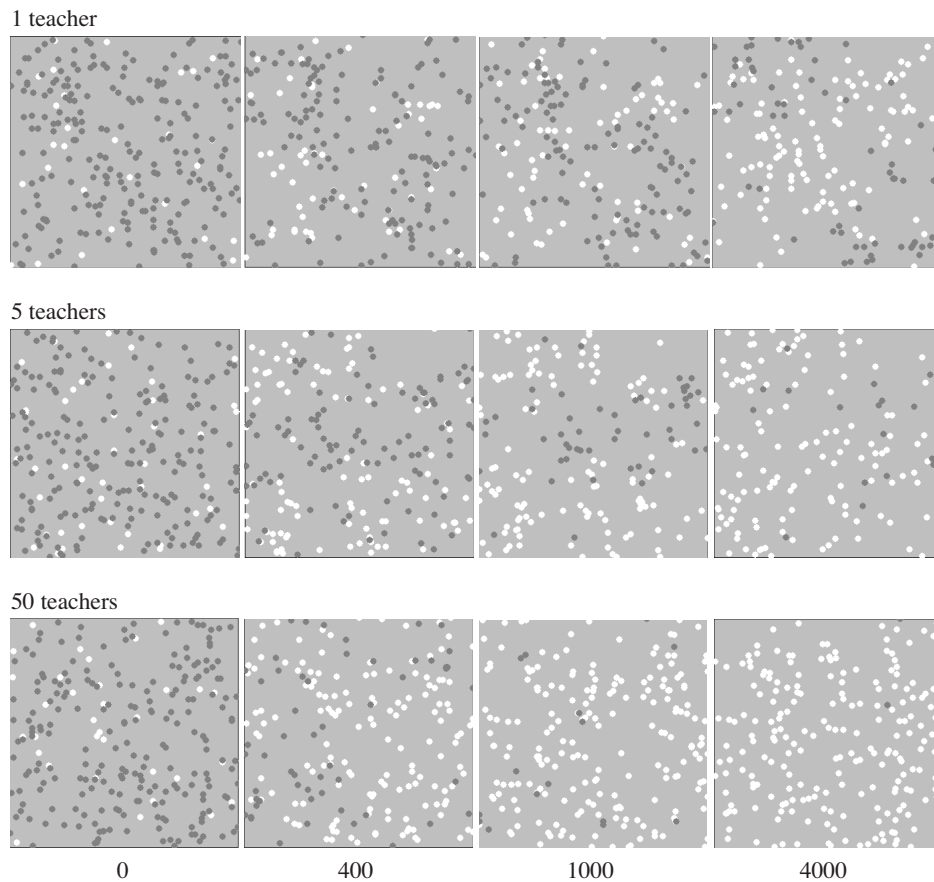


Figure 12.7. Spread of knowledge pieces over the students for typical runs with 1, 5 and 50 teachers at timesteps 0, 400, 1000 and 4000.

Figure 12.7. shows a series of maps of the world displaying the incidence of the required knowledge piece (`[carry wholesome plant] ⇒ eat`) geographically. The three sequences of maps show the spread of knowledge over time for typical runs with 1, 5 and 50 teachers respectively. Students that contain the required knowledge show white, those that don't show dark grey. Teachers and knowers are not shown. Note the logarithmic time-scale.

Again, it is plain that, even with a single teacher to initiate dissemination, the decisive knowledge is spread through a significant part of the population—the population as a whole stores the knowledge effectively and robustly. As could be expected, the knowledge becomes even more widespread for the experiments with 5 and 50 teachers.

While we have seen the behaviour for the student population reach similar levels for the experiments with 5 and 50 teachers, this is not the case for the incidence of the expected knowledge piece. With 50 teachers, practically all students have obtained this knowledge piece after 4000 time-steps, but in the 5 teachers case, a portion of the students remains unaware of this information at that time. Similarly,

there is no appreciable difference between $g(t)$ at time-step 1000 and at time-step 4000 for the 50 teachers experiments, but there is a marked difference in incidence of the required knowledge piece. From this we can conclude that, after a certain level of prevalence has been achieved, further proliferation of the knowledge piece has no perceivable effect on population behaviour in terms of $g(t)$.

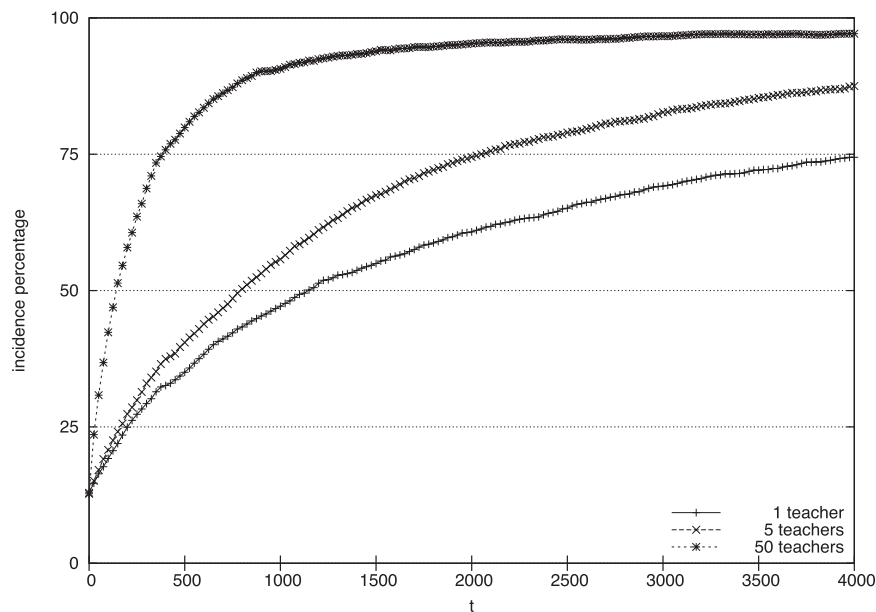


Figure 12.8. Development over time of the percentage of students with the crucial knowledge piece.

Figure 12.8. shows how the percentage of students that have learned the requisite knowledge develops over time with 1, 5 and 50 teachers, respectively, averaged over 20 runs. Because the students spread the knowledge they receive, incidence grows almost exponentially as can be seen from the graph.

Note, that at time-step 0, a portion of the population does contain the knowledge as part of the randomly initialised tree while $g(t)$ for the runs without any teachers doesn't increase over time. This can be explained by the context in which the knowledge piece may be present (i.e., as sub-clause in a more complex, possibly nonsensical rule) and by the fact that the action node's weights (as described in Sec. 12.2.2) aren't sufficiently biased towards actually selecting the eat action.

12.4.5 Discussion

We asked ourselves the questions of whether social learning can provide a successful mechanism to spread knowledge pieces over a population, and is there a minimum requirement to enable the population to create a knowledge repository for otherwise volatile individually acquired knowledge.

The results of the poisonous food experiments clearly show that social learning does provide an efficient mechanism for the dissemination of knowledge pieces through a population of agents. Even from a single agent, the knowledge can

324 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

spread over the majority of the population like an ink-stain on tissue paper. Within the framework of PAS in general and the implementation in *NEW TIES* in particular, this means that social learning is capable of allowing the population to form a knowledge repository for individually acquired knowledge so that such knowledge doesn't necessarily expire with the agent that discovered it.

12.5 Embodied Imitation and Memetic Evolution

This section presents a further and contrasting case study in artificial social learning. The case study focusses on one particular kind of social learning called 'embodied imitation'. Here the artificial agents are physically embodied, i.e. they are robots, and social learning takes place through robot-robot imitation, i.e. one robot imitating the behaviour(s) of another. The work presented here is part of a multi-disciplinary research project called "the emergence of artificial culture in robot societies" whose overall aim is to investigate the processes and mechanisms by which proto-cultural behaviours, or traditions, might emerge in a free running collective robot system. However, at the time of writing, this is an ongoing project and so this section will focus on the processes and mechanisms of embodied robot-robot imitation, and how socially learned behaviours evolve in a robot collective.

This section is organised as follows. First we outline a definition for embodied imitation and place that in the context of the existing research literature on robot imitation. Then we present an outline of the "artificial culture" project and its physical infrastructure, since this also provides essential context for the work on embodied imitation. Next the section describes an implementation of robot-robot imitation on e-puck mobile robots and presents experimental results which focus on variation and the quality of imitation. Finally we describe and analyse an experiment in open-ended memetic evolution.

12.5.1 Embodied Imitation

We define embodied imitation, or robot mimesis, as: the imitation of one robot's behaviour, or sequence of behaviours, by another robot, where the learner-robot uses only its embodied sensors to perceive the teacher-robot. This definition precludes robot-robot 'telepathy', i.e. one robot transferring part of its controller directly to another as in the previous section. Our insistence on embodied perception of one robot by another means that an implementation of embodied imitation needs to solve the so-called 'correspondence problem'; a term which refers to the learner's problem of translating a set of perceptual inputs to motor actions that correspond with the perceived actions of the teacher [Nehaniv and Dautenhahn (2002a)]. This is a hard problem, which begs the question: since robots are capable of telepathy, why do we insist on solving the correspondence problem? The answer is that we are not here proposing solutions to real-world robotics problems but instead interested in modelling and illuminating the processes of social learning, and therefore biological plausibility is important to us. Furthermore, real robots and sensors, less-than-perfect-fidelity embodied perception and the estima-

tion process inherent in solving the correspondence problem, means that embodied imitation will be imperfect and, as we will describe later in this section, imperfect imitation is valuable to us because it leads naturally to variation in socially learned behaviours as they propagate across the population.

The study of imitation and social learning in robots, humans and animals has received cross-disciplinary attention in recent years [Nehaniv and Dautenhahn (2002b, 2007)]. Not surprisingly attention has been given to the problem of humanoid robots imitating humans, since this presents a way of programming a robot by demonstration rather than coding [Scassellati (1999); Mataric (2000)].

There has been less work describing experiments in embodied robot-robot imitation. The earliest is perhaps the work of Hayes and Demiris which describes an approach with one (pre-programmed) robot finding its way through a maze and another following it and observing its actions (turns). The following (learner) robot then associates each observed action with its own (time delayed) perception of the environment and hence learns how to navigate the maze, by imitation; this kind of imitation is called 'matched dependent behaviour' [Hayes and Demiris (1994)]. In what is perhaps the first attempt to develop a generalisable approach to robot-robot imitation Gaussier and Moga, et al, proposed a sophisticated approach to imitation using a neural network architecture able to learn temporal perception-action sequences [Gaussier *et al.* (1998); Moga and Gaussier (1999); Moga (2000)]. Also following a connectionist approach Billard and Hayes proposed the DRAMA architecture (dynamical recurrent associative memory architecture); they provide one case study that interestingly involves the active participation of the teacher robot in the process of imitative learning [Billard and Hayes (1999)].

Following their 1994 work, [Hayes and Demiris (1994)] Demiris et al, went on to propose the HAMMER architecture (Hierarchical, Attentive, Multiple Models for Execution and Recognition). In an important series of papers Demiris et al, developed an imitation architecture based upon the building block of paired inverse and forward models; the inverse model outputs motor commands but when, instead of actually driving the motors, those commands are fed to the forward model and the output of the forward model compared with the input of the inverse model the robot is able to 'rehearse' possible actions and compare these with its perception of the actions it is trying to imitate [Demiris and Hayes (2002); Johnson and Demiris (2004); Demiris and Khadhour (2005)]. The idea that these internal models reflect, in some sense, the simulation 'theory of mind' are further explored in [Johnson and Demiris (2005); Demiris and Johnson (2007)].

Alissandrakis et al, developed the ALICE architecture (Action Learning via Imitation between Corresponding Embodiments) in order to address the problem of robot-robot imitation across dissimilar embodiments; although not tested with real robots, ALICE contributes a powerful generalised solution to the correspondence problem for agents (or robots) with different morphologies [Alissandrakis *et al.* (2002, 2003, 2007)]. ALICE works by creating a correspondence library relating the actions (and importantly effects) of the teacher to actions (or action sequences) that the learner is capable of.

12.5.2 The Artificial Culture Project

The Artificial Culture project aims to address and illuminate the question “how can culture emerge and evolve as a novel property in groups of social animals?” by building an artificial society of embodied intelligent agents (real robots), creating an environment (artificial ecosystem) and appropriate primitive behaviours for those robots, then free running the artificial society. The aims of the project lie primarily in modelling the processes and mechanisms by which we might observe the emergence of ‘artificial traditions’ in a group of embodied agents (robots). Even with small populations (a few tens) of relatively simple robots we see, in a short time, a very large number of interactions between robots. In this project those interactions will encompass social learning.

In the project we are attempting to create the conditions and primitives in which proto-culture, or traditions, might emerge in the robot collective. Robots will, for example, be able to copy each other’s behaviours and select which behaviours to copy. Dawkins coined the term ‘meme’ to describe a unit of cultural transmission [Dawkins (1976)], and we use this terminology here¹². Imitated behaviours (memes) will mutate because of the noise and uncertainty in the real robots’ sensors and actuators, and successful memes will undergo multiple cycles of copying (heredity), selection and variation (mutation). Observed measurable and sustainable differences in the memes across different groups of robots, where those memes can be traced back to common ancestral memes, might be interpretable as evidence for emerging ‘artificial traditions’ in the robot collective.

The project is inspired by the *Copybots* thought experiment suggested by Blackmore, [Blackmore (1999)] pages 106-107, and by Dautenhahn’s visionary 1995 paper ‘Getting to Know Each Other - Artificial Social Intelligence for Autonomous Robots’ [Dautenhahn (1995)]. From a technical perspective the project draws upon a multi-disciplinary body of literature in imitation [Nehaniv and Dautenhahn (2007)]; for instance the work of [Alissandrakis *et al.* (2007)] describing imitation leading to “cultural transmission of behaviours and emergence of proto-culture” between two simulated 2D two-jointed robotic arms. However, we argue that a multi-robot collective is a necessary substrate for this work and bring key concepts from the field of swarm robotics [Beni (2005); Şahin and Winfield (2008)].

12.5.2.1 The Artificial Culture Laboratory

Core to the project is the creation of an artificial environment: the artificial culture lab. The artificial culture lab comprises a physical space (‘arena’) designed for and populated by miniature wheeled mobile robots. The arena is closed in the sense that its physical boundaries define the edges of the robots’ world, out of which they cannot physically stray. The arena is not hermetically sealed, thus robots (since they have both light and sound sensors) are affected by ambient lighting or noise levels. Providing that these external environmental influences do not overwhelm (blind) the robots’ sensors, they are not a problem. Indeed, a certain level of background noise in the environment is considered essential as it will contribute to

¹²Equivalent to the ‘knowledge nugget’ of Sec. 12.2

imperfect robot-robot imitation or communication, and changing light levels (day and night) might be useful in providing the robots with a circadian rhythm.

The robots, called e-pucks, are wheeled, differential-drive, robots capable of moving forwards or in reverse, or turning (including turning on the spot) [Moncada *et al.* (2009)]. They are equipped with a range of sensors, including short-range infra-red and/or ultra-sound proximity and ranging sensors that allow the robots to sense the presence, direction and range of obstacles and other robots close by. Importantly, robots can sense and track the movements of other robots nearby (albeit imperfectly because of their limited sensors); thus robots have the physical means for imitation. They have multi-coloured programmable lights (LEDs), and simple cameras; microphones and speakers. We have a wide range of options for robot-robot interaction. Robots can signal to each other with movement, light, or sound, one-to-one or one-to-many, and with or without active consent (i.e. one robot can eavesdrop on the communication between two others). The robots are not equipped with manipulators (grippers), thus the only way they can physically act upon the world is with their own bodies (i.e. by pushing light objects, or co-operating with other robots to push heavier objects).

The artificial culture lab is fully instrumented. A tracking system allows the movements of all robots to be captured and recorded for analysis and interpretation. Wireless communication with each robot allows data logging, allowing the emerging behaviours to be captured for analysis. Webcams provide video capture for analysis, and importantly video for project web-pages for open access to support interpretation; for discussion of interpretation see [Winfield and Griffiths (2010b)]. Fig. 12.9.(a) shows the artificial culture lab in the Bristol Robotics Lab (BRL); Fig. 12.9.(b) shows one of the e-puck robots fitted with Linux extension board and tracking 'hat' [Liu and Winfield (2010b)].

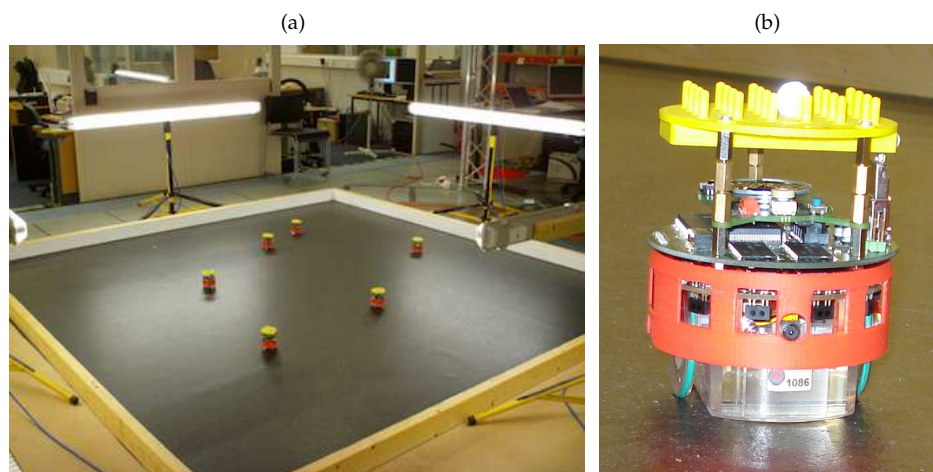


Figure 12.9. (a) Artificial culture lab showing 6 robots in the arena; (b) An e-puck with Linux board fitted in between the e-puck motherboard (lower) and the e-puck speaker board (upper). Note both the red 'skirt' which allows one robot to see and track another with its camera, and the yellow 'hat' which provides a matrix of pins for the reflective spheres which allow the tracking system to identify and track each robot.

328 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

The use of real physical robots in an artificial ecosystem as described above, rather than computer simulated agents, is central to the methodology proposed for this project. The rationale is that real robots provide vastly more scope for emergence in their interactions than simulated agents. The combination of imperfect sensors; sensing errors that occur because of the distance between robots; multiple robots sharing the same environment (i.e. occlusion of robots by each other) and sharing the same communications modality (i.e. all talking at once); small differences between sensors and actuators (motors) which mean that the robots are not all identical; real-world physics which means that each experimental run (even with the same starting conditions) will quickly diverge into a new space of possibilities; noise in the environment and, unexpected non-fatal faults (i.e. a faulty wheel which gives the robot a 'limp'), we argue could not be created in simulation (to do so each of the factors listed would have to be separately modelled, and those models would inevitably lead to simplification thus chronically limiting the space of possibilities). Even in the designed artificial environment we propose here, the use of real physical robots provides vast scope for unexpected emergence. Thus, we argue, behavioural artefacts that might be interpretable as artificial memes – elements of an artificial proto-culture – will emerge for no other reason than that they can.

12.5.3 *Robot–robot Imitation of Movement*

We are concerned here with the embodied imitation of behaviour, but 'behaviour' is too broad a term. Within embodied imitation we can identify at least three types of imitation:

- imitation of actions only, i.e. one robot copying another's sequence of movements, sounds or lights;
- imitation of action and perception, i.e. one robot copying another's interactions with objects or other robots: we label this the 'imitation of interaction';
- imitation of goals, i.e. one robot copying the goals or intentions of another using, perhaps, a completely different set of actions.

The third of these categories, the imitation of goals, is outside the scope of this paper. Because it is the simplest we have first implemented the imitation of actions, and specifically the imitation of movement. We now describe the imitation-of-movement algorithm and experimental results obtained.

Before outlining the imitation algorithm we first need to describe the basic setup and some simplifying assumptions. In this approach one robot performs a sequence of movements (the 'teacher') while another robot watches it (the 'learner') and then attempts to copy the observed sequence. The roles of teacher and learner are not fixed but interchangeable and – since we are interested here in propagation of imitated behaviours – robots alternate between teacher and learner modes. When in teacher mode a robot is ready to perform a movement sequence (i.e. 'enact' a meme) it will signal this by flashing its red LEDs. There is no handshake between teacher and learner, so any robot(s) in learner mode that see the red LED signal will start to observe the sequence, but if no robots see the signal

then the teacher robot will go on to perform the sequence anyway. The learner robot watches the teacher robot with its onboard camera and, in order to facilitate the recognition of the teacher robot and its movements, robots are fitted with coloured skirts that contrast with the background (i.e. arena boundaries), as shown in Fig.12.9.(b). Since robots have only one camera and hence monoscopic vision the learner robot must judge the relative direction of movement and distance of the teacher robot by tracking the position and size of the teacher's coloured skirt in its field of view, rotating if necessary to maintain the teacher in its field of view. Although estimating relative size and position of the teacher robot is straightforward image processing, it is error prone because of the relatively low resolution camera (640x480) and the presence of other robots and, furthermore, the learner robot cannot see the teacher robot's turns – only infer them from changes in direction, thus we simplify the correspondence problem by limiting movement sequences to be composed of turns and straight line segments at a constant velocity.

The imitation of movement algorithm thus has three stages:

- (1) while observing captured visual frames, for each frame estimate the relative position of the teacher robot, storing these positions in a linked list;
- (2) after the teacher's sequence is complete, process the linked list using a regression line fitting approach to convert the estimated positions into straight line segments;
- (3) transform the straight line segments, and their intersections, into a sequence of motor commands (moves and turns).

The imitation-of-movement algorithm outlined here does not have the sophistication or complexity of the architectures outlined above in Sec. 12.5.1, although it does clearly share a number of common elements. There are a number of reasons for the relative simplicity of our approach. Firstly, we are here imitating movement only and not interaction with objects, or other robots: thus the learner needs only to deduce action sequences and not perception-action sequences. Secondly our robots are homogeneous (similarly embodied), thus when the learner robot transforms its estimate of the teacher robot's movement trajectory into ego-centric motor commands it can assume it has the same motion dynamics as the teacher robot. Thirdly, we are interested primarily in meme propagation across the robot collective, so our approach to imitation is only as complex as needed to create the conditions for movement-meme evolution.

12.5.4 Experimental Results

Initial experimental trials have focussed on small numbers of robots, typically 2 or 4, in order that we can test and evaluate the performance of robot-robot social learning by imitation before then scaling up to larger populations of robots. In this subsection we first outline a quality-of-imitation metric, then describe and analyse in detail two instances of robot-robot movement-imitation. Then we show how, with 2 robots, it is possible to demonstrate open-ended memetic evolution.

12.5.4.1 A Quality of Imitation Metric

In order to quantitatively assess the fidelity of imitation (i.e. similarity of learned memes) we need to define a quality-of-imitation function, Q_i . Since movement-memes consist of a series of turn and straight line segments (vectors) we can compare the similarity of two memes by separately estimating three quality indicators: quality of angle (turn) imitation, quality of length imitation, and quality of segment imitation. The quality of angle imitation between original meme (O) and learned meme (L) is calculated as follows:

$$Q_a = 1 - \frac{\sum_m |a_m^L - a_m^O|}{\sum_m a_m^O} \quad (12.2)$$

where a_m is the turn angle of move m . Here we determine the ratio of the sum of angle differences between the moves of original and learned memes, to the total turn angle of the moves of the original meme. If original and learned memes have a different number of segments, N^O and N^L respectively, then we sum only over the number of segments in the smaller: $\min(N^L, N^O)$. A value of 1 indicates perfect fidelity imitation. The quality of length imitation similarly calculates the length errors between original and learned memes:

$$Q_l = 1 - \frac{\sum_m |l_m^L - l_m^O|}{\sum_m l_m^O} \quad (12.3)$$

where l_m is the length of move m . Again a value of 1 indicates perfect fidelity imitation. The quality of segment imitation simply considers the difference between the number of segments (vectors) between original and learned memes:

$$Q_s = 1 - \frac{|N^L - N^O|}{N^O} \quad (12.4)$$

where N^L and N^O are the number of segments of learned and original memes, respectively. We now calculate the weighted sum of the three quality indicators, to arrive at a composite overall quality-of-imitation score:

$$Q_i = AQ_a + LQ_l + SQ_s \quad (12.5)$$

where A , L and S are weighting coefficients, and $A + L + S = 1$. In the results given here we give equal weighting, thus $A = L = S = 0.33$.

12.5.4.2 Robot-robot Imitation with Variation

Figure 12.10. shows two examples of embodied social learning, of movement, by imitation. Each of the three subfigures in Fig. 12.10. plots tracking data recorded, from the experimental infrastructure described in Sec. 12.5.2.1, when an epuck robot enacts a single movement sequence. Here epuck 9 has been initialised with a sequence of three turns and moves that describe an equilateral triangle, with 15 cm sides, and Fig. 12.10.(a) shows epuck 9 enacting the triangle. In this trial epuck 12 watched epuck 9 enact the triangle and, using the procedure outlined above, attempted to learn the movement sequence, by embodied imitation; the result is shown in Fig. 12.10.(b) and it is immediately clear that this is a poor-fidelity copy.

12.5. Embodied Imitation and Memetic Evolution | 331

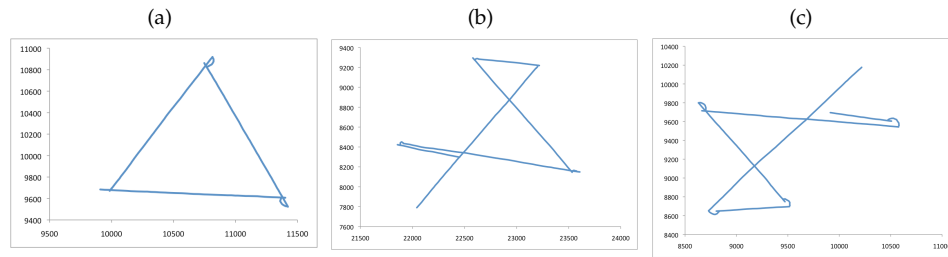


Figure 12.10. (a) Meme 1: initial movement meme enacted by epuck 9. (b) Meme 2: imitation of meme 1 by epuck 12, $Q_i = 0.47$. (c) Meme 3: imitation of meme 2 by epuck 9, $Q_i = 0.94$. Units of x and y axes are 0.1 mm.

Although the copy clearly retains characteristics of the original triangle two additional short segments have been inserted, one at the start, followed by a u-turn, and another at the top apex of the triangle. Given these two additional segments it is not surprising that our quality of imitation score is poor: $Q_i = 0.47$. The quality of length imitation is much higher: $Q_l = 0.75$.

In this trial epuck 9 then watched Meme 2, enacted by epuck 12, and attempted to learn it, thus Fig. 12.10.(c) is an imitation of Fig. 12.10.(b). In contrast with the poor fidelity Meme 1 \rightarrow Meme 2 imitation, we see that Meme 2 \rightarrow Meme 3 imitation is much higher fidelity. Meme 3 is of course rotated with respect to Meme 2, but that is exactly what we would expect. Meme 3 retains the rather complex five segment structure of Meme 2, and gives a very high quality of imitation score of $Q_i = 0.94$; epuck 9 has certainly learned the complex ‘dance’ of epuck 12. We have thus demonstrated both robot–robot social learning, by imitation, and shown that we obtain variation in socially learned behaviours ‘for free’ as a consequence of embodiment.

12.5.4.3 Open-ended Memetic Evolution

An evolutionary process requires variation, selection and inheritance. We already have, as demonstrated above, variable degrees of both variation and inheritance; embodied and noisy sensors together with environmental noise lead to variation and, depending on the fidelity of imitation, we also see that copied memes may inherit characteristics of the original meme. Selection we cannot get ‘for free’ (except in the trivial sense that our robots might simply imitate any robot that happens to come into view). Let us now arrange that our robots memorise each meme they observe and learn, regardless of the fidelity of the learned memes. During an experimental run each robot’s memory (imeme¹³ list) will grow, and each time it is that robot’s turn to enact a meme it will need to select one from its memory. In this scheme there is no need for a robot to compute the similarity between learned memes and robots therefore do not ‘recognise’ previously seen memes.

There are clearly many ways in which we could select which meme to enact. We could, for instance, select on some feature of the stored memes with a fitness function that favours say memes with a given number of moves, or turn angles less

¹³Here we use the word imeme as shorthand for ‘internal representation of a meme’

332 | Individual, Social and Evolutionary Adaptation in Collective Systems

than a given value (for smoother, less angular ‘dances’), or smaller or larger move lengths. However, in initial trials we have sought a selection operator that requires no fitness function in order that we do not constrain the direction of memetic evolution; to achieve – in other words – open-ended memetic evolution. This is of particular interest because we want to see whether the robots’ sensors and morphology (body shape, sensor placement and motors) might influence the direction of memetic evolution. We report here trials with one such selection operator: *select, from the imeme list, at random, with equal probability.*

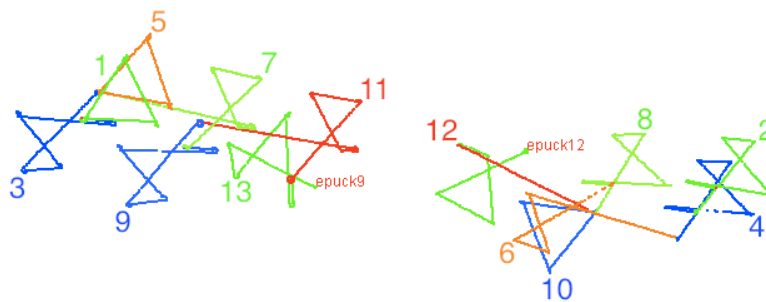


Figure 12.11. Trajectory plot: two robot movement-meme evolution in which all observed memes are stored and meme selection is random, with equal probability. The experiment starts with epuck 9 (left) in teacher mode, following a movement trajectory that describes a triangle with sides of 15 cm.

Figure 12.11. plots the position data captured during a two robot experiment in which each robot alternates between teacher-mode and learner-mode. Each robot learns and stores the meme enacted by the other, but then – when in teacher-mode – chooses which meme to enact using the equal-weighting random-selection operator. For clarity each movement sequence is shown here in a different colour, and labelled with the order in which the movement-memes were enacted by the two robots. In this run each robots memory is initialised with one imeme: a pattern of movements that describe an equilateral triangle with sides of 15 cm, and epuck 9 is initially in teacher mode.

We now apply the graphical meme-tracking approach proposed in [Winfield and Griffiths (2010b)], in order to trace the evolution of memes in the experiment of Fig. 12.11.. Inspection of Fig. 12.11. shows that a ‘figure of eight’ meme appears to dominate, and the meme evolution diagram in Fig. 12.12. explains why.

Figure 12.12. shows the evolution and heredity of memes in the two robot experimental trial of Fig. 12.11.. It does not identify robots, but instead traces the evolution of memes – something which is not obvious from the trajectory plot of Fig. 12.11. and requires deeper analysis of experimental logs to determine, for each selected and enacted imeme, which original (parent) meme it is a learned copy of. Thus, each horizontal grey line in Fig. 12.12. represents a timeline for each single imeme. When that imeme is selected and enacted there are two possibili-

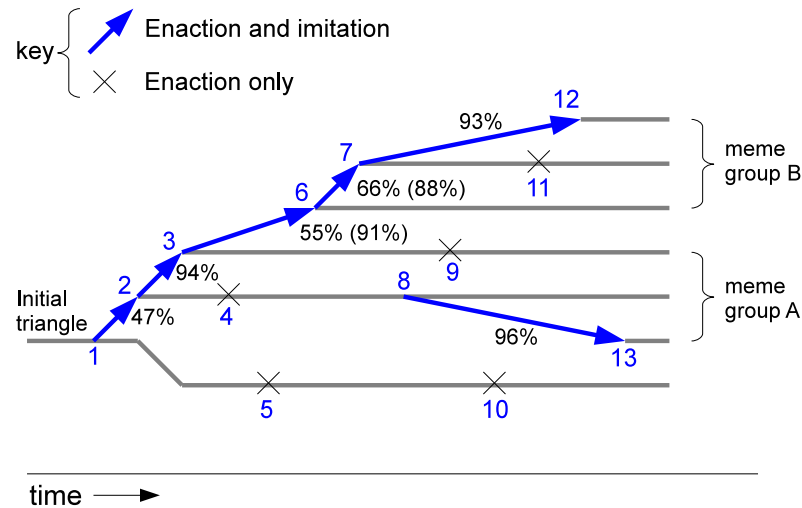


Figure 12.12. A visualisation of meme evolution within the (two) robot collective. At the start of the period just one movement-meme (triangle) is present; horizontal lines represent the ‘life course’ of each meme from left to right. Events (enaction only or enaction and imitation) are labelled with numbers, in blue, which correspond with memes in Fig. 12.11.. Enaction and imitation events (blue arrows) are labelled with the quality-of-imitation score.

ties: one is that the enaction was not, at any time during the experimental trial, imitated (i.e. learned *and* enacted) – these are shown as crosses (and labelled in the key ‘enaction only’). The other possibility is that the enaction was imitated during the experimental trail – these are shown as blue arrows (and labelled in the key ‘enaction and imitation’). The enaction only events (crosses) are labelled with the meme-enaction number in trajectory plot Fig. 12.11.; the enaction and imitation events (blue arrows) are labelled with originating (parent) meme-enaction number, and learned meme-enaction number from Fig. 12.11.. Each enaction and imitation event is also labelled with the quality-of-imitation score Q_i . Note that each enaction and imitation event results in a new imeme timeline which continues throughout the trial. This reflects the fact that our robots, in this experiment, have unlimited imeme memories. If we instead had either a mechanism for robots ‘forgetting’ imemes (according to some criteria) or robots themselves ‘dying’, then some imeme timelines would terminate.

Thus we see, in Fig. 12.12., that Meme 2 is a poor-fidelity copy of Meme 1 (0.47) – the first ‘figure of eight’ movement pattern. Significantly, Meme 3 happens to be a high-fidelity copy of Meme 2 (0.94), and furthermore there are no further enaction and imitation events originating from Meme 1 – just two enaction only events: 5 and 10. Thus, all second and later generation memes have, as an ancestor, Meme 2. This fact, together with the high-fidelity copy of Meme 8 → Meme 13 (0.96) means that Memes 2,3,4,8 and 13 are all either the same or very closely related and

334 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

we label these Meme group A. Consider now imitation event Meme 3 \rightarrow Meme 6, which appears to be relatively poor quality (0.55). However inspection shows that Meme 6, which has four segments, has lost the initial short segment of Meme 3; if we ignore the first segment of Meme 3 and re-calculate Q_i for Meme 3 \rightarrow Meme 6, we obtain 0.91 – which more closely reflects the subjective similarity of Memes 3 and 6. By chance imitation event Meme 6 \rightarrow Meme 7 has inserted a new short segment so that Meme 7 returns to five segments and, by similarly ignoring the new segment in Meme 7 and re-calculating Q_i , we obtain 0.88. Thus we see that Meme group B is both quantitatively and subjectively similar to Meme group A, with strongly inherited characteristics retained across five generations of meme: 2 to 12. We now understand why the emergent figure of eight movement pattern has become dominant.

Of course this particular meme evolution is highly contingent. The emergence of the same kind ('species') of dominant 'figure of eight' movement memes is most unlikely to happen again (and indeed, in repeat trials, has not). But this is exactly what one would expect of an embodied evolutionary process. Perhaps what is surprising is that in an open-ended evolutionary system one kind of meme becomes dominant (at least in this particular trial) – but this is simply explained by the fact that if there is a group of closely related memes in the robots' memories (because of high-fidelity learning) then our equal probability random selection operator is more likely to select one of these. Note also just how important the initial few imitation events are to the later evolution of the system; the happenstance initial sequence of a poor-fidelity imitation event followed by a high-fidelity imitation event strongly (although not completely) determined the later evolutionary course of our trial system. Again this is strongly characteristic of an evolutionary system.

12.5.5 *Discussion and Further Work*

This section has presented a case-study in artificial social learning. The project that this case study draws upon is, at the time of writing, ongoing and it would be premature to draw any general conclusions with regard to the overall project aims of illuminating the processes and mechanisms for the emergence of artificial traditions across a robot collective. However, initial experimental trials have demonstrated robot-robot social learning, by imitation, and shown that embodiment gives rise to variation in socially learned behaviours. With the addition of a simple selection operator we have demonstrated promising open-ended memetic evolution, with just two robots, in which a new behavioural 'meme' can emerge and become dominant (albeit over a short experimental trial), in a 'population' of memes.

The case study described here demonstrates an adaptive system in three different respects. Firstly, because individual robots are able to adapt their behavioural repertoire, by social learning. Secondly, because the whole population evolves and therefore, in a sense, self-adapts its social environment. And thirdly, the selection operator could be designed to adapt the robots socially learned behaviours in a desired direction.

Further work will:

- run longer trials with larger groups of robots to investigate the dynamics, over time, of dominant meme-groups and convergent evolution;
- explore the relationship between embodiment, including sensor characteristics and robot morphology, and the quality of imitation, and address questions relating to the stability of meme transmission;
- extend the imitation algorithm to implement the imitation of interaction so that interactions between robots can be imitated and propagated across the collective, with richer ‘social learning’;
- further explore the mechanisms of meme selection together with environmental variation, in order to model the spatial and temporal dynamics of meme propagation across the robot collective and the possibility of the emergence of artificial traditions.

12.6 Conclusion

We began the chapter by introducing a framework for adaptation in population-based adaptive systems (PAS), positioning and relating evolution, individual and social learning.

The chapter showed examples of each type of evolutionary mechanism as implemented in the *NEW TIES* platform as well as an example of how social learning can be achieved through imitation among robots.

All this serves to illustrate that there are many ways to set up adaptive behaviour in a PAS, be it individually or collectively. The experiments in this chapter show that evolution, individual and social learning all provide powerful mechanisms for initiating and spreading adaptation. Combining mechanisms may further enhance the population’s performance, as indicated by the experiment in Sec. 12.4. There, we saw that social learning can provide an excellent method to share individually acquired adaptations among the population, allowing the whole population to benefit from an individual’s experience and preventing valuable knowledge from being lost when an individual –be it a software agent or a robot– ceases to function.

We saw that social learning can constitute (part of) an evolutionary adaptive system. The ‘telepathic’ version described in Sec. 12.4 can co-operate with individual learning; individual learning provides variation, while social learning implements recombination. Selection is done environmentally (poorly adapted individuals disappear) and/or in social learning. The embodied imitation case study of 12.5 illustrated that (slightly) inaccurate copying of behaviour, together with selection of which behaviour to enact, combine to make a full-fledged evolutionary process.

Indiscriminate combination of adaptive mechanisms, however, carries a danger: it may lead to the emergence of unwanted interactions as shown in Sec. 12.3. There, we saw that individual learning can counteract as well as promote evolution depending on the rewards the learning is based on: a case in point that illustrates how the goals of the adaptation mechanisms have to be in tune with each other. Therefore, careful consideration should be given to the interactions between the

336 | *Individual, Social and Evolutionary Adaptation in Collective Systems*

adaptive mechanisms when designing a PAS: one cannot simply design the mechanisms in splendid isolation.

Investigating these interactions, then, is one of the challenges that future PAS research will have to address to understand how adaptive mechanisms can be combined to enable truly autonomous robots, robots that can indeed learn control without human supervision.

Acknowledgements

Part of the work presented in this chapter was undertaken as part of the NEW TIES project, supported by a European Commission FET grant under contract FP6-502386. Also, the authors would like to thank Paul Vogt, Andras Lörincz, Zoltán Papp, Martijn Schut, Selmar Smit and the other members of the NEW TIES consortium for their contributions. The Artificial Culture project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC), grant reference EP/E062083/1. The authors especially thank Mehmet Dincer Erbas who implemented the imitation algorithm and obtained the experimental results shown in section 12.5 of this chapter.

Bibliography

- Acerbi, A., Marocco, D. and Vogt, P. (2008). Social learning in embodied agents, *Connection Science* **20**, 2, pp. 69–72.
- Acerbi, A. and Nolfi, S. (2007). Social learning and cultural evolution in embodied and situated agents, in *Proceedings of the First IEEE Symposium on Artificial Life (ALIFE '07)* (IEEE Press, Honolulu, HI), pp. 333–340.
- Alissandrakis, A., Nehaniv, C. and Dautenhahn, K. (2002). Imitation with ALICE: Learning to imitate corresponding actions across dissimilar embodiments, *IEEE Trans. Systems, Man and Cybernetics - Part A: Systems and Humans* **32**, 4, pp. 482–496.
- Alissandrakis, A., Nehaniv, C. and Dautenhahn, K. (2007). Solving the correspondence problem in robotic imitation across embodiments: synchrony, perception and culture in artefacts, in C. Nehaniv and K. Dautenhahn (eds.), *Imitation and Social Learning in Robots, Humans and Animals* (Cambridge University Press), pp. 249–273.
- Alissandrakis, A., Nehaniv, C. L. and Dautenhahn, K. (2003). Synchrony and perception in robotic imitation across embodiments, in *IEEE Int. Symp. on Computational Intelligence in Robotics and Automation*, pp. 923–930.
- Belew, R., McInerney, J. and Schraudolph, N. (1990). Evolving networks: Using the genetic algorithm with connectionist learning, in C. L. et al. (ed.), *Proceedings of the Second Conference on Artificial Life* (Addison-Wesley, Reading, MA), pp. 511–547.
- Belew, R. and Mitchell, M. (1996). *Adaptive Individuals in Evolving Populations: Models and Algorithms* (Addison-Wesley).
- Beni, G. (2005). From swarm intelligence to swarm robotics, in E. Şahin and W. M. Spears (eds.), *Swarm Robotics - SAB 2004 International Workshop*, volume 3342 of *LNCIS* (Springer-Verlag, Berlin, Germany), pp. 1–9.
- Best, M. (1999). How Culture Can Guide Evolution: An Inquiry into Gene/Meme Enhancement and Opposition, *Adaptive Behavior* **7**, 3-4, p. 289.
- Billard, A. and Hayes, G. (1999). DRAMA, a connectionist architecture for control and learning in autonomous robots, *Adaptive Behaviour* **7**, 1, pp. 35–63.
- Blackmore, S. (1999). *The Meme Machine* (Oxford University Press).
- Bull, L., Studley, M., Bagnall, A. and Whitley, I. (2007a). Learning Classifier System Ensembles With Rule-Sharing, *IEEE Transactions on Evolutionary Computation* **11**, 4, pp. 496–502.
- Bull, L., Studley, M., Bagnall, A. and Whitley, I. (2007b). Learning Classifier System Ensembles With Rule-Sharing, *IEEE Transactions on Evolutionary Computation* (in press).
- Buresch, T., Eiben, A., Nitschke, G. and Schut, M. (2005a). Effects of evolutionary and lifetime learning on minds and bodies in an artificial society, in *Proceedings of the 2005 IEEE Congress on Evolutionary Computation (CEC 2005)* (IEEE Press), pp. 1448–1454.
- Buresch, T., Eiben, A., Nitschke, G. and Schut, M. (2005b). Effects of evolutionary and lifetime learning on minds and bodies in an artificial society, in D. Corne, Z. Michalewicz, B. McKay, G. Eiben, D. Fogel, C. Fonseca, G. Greenwood, G. Raidl, K. Tan and A. Zalzala (eds.), *Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2005)* (IEEE Press), pp. 1448–1454.
- Cangelosi, A. and Parisi, D. (1998). The emergence of “language” in an evolving population of neural networks, *Connection Science* **10**, pp. 83–93.
- Şahin, E. and Winfield, A. F. T. (eds.) (2008). *Swarm Intelligence: Special Issue on Swarm*

54 | Bibliography

- Robotics*, Vol. 2(2-4) (Springer).
- Curran, D. and O’Riordan, C. (2006). Increasing population diversity through cultural learning, *Adaptive Behavior* **14**, 4, pp. 315–338, doi:http://dx.doi.org/10.1177/1059712306072335.
- Dautenhahn, K. (1995). Getting to know each other - artificial social intelligence for autonomous robots, *Robotics and Autonomous Systems* **16**, pp. 333–356.
- Dautenhahn, K. and Nehaniv, C. L. (2002). *The agent-based perspective on imitation* (MIT Press, Cambridge, MA, USA), ISBN 0-262-04203-7, pp. 1–40.
- Dawkins, R. (1976). *The Selfish Gene* (Oxford University Press).
- Demiris, Y. and Hayes, G. (2002). Imitation as a dual-route process featuring predictive and learning components: a biologically-plausible computational model, in K. Dautenhahn and C. Nehaniv (eds.), *Imitation in Animals and Artefacts*, chap. 13 (MIT Press).
- Demiris, Y. and Johnson, M. (2007). Simulation theory of understanding others: a robotics perspective, in C. L. Nehaniv and K. Dautenhahn (eds.), *Imitation and Social Learning in Robots, Humans and Animals* (Cambridge University Press), pp. 89–102.
- Demiris, Y. and Khadhour, B. (2005). Hierarchical, attentive, multiple models for execution and recognition (HAMMER), in *Proc. ICRA 2005*.
- Denaro, D. and Parisi, D. (1996). Cultural evolution in a population of neural networks, in *Proceedings of the 8th Italian Workshop on Neural Nets* (Springer-Verlag, London), pp. 100–111.
- Eiben, A., Haasdijk, E. and Bredeche, N. (2010). Embodied, on-line, on-board evolution for autonomous robotics, in P. Levi and S. Kernbach (eds.), *Symbiotic Multi-Robot Organisms: Reliability, Adaptability, Evolution*, chap. 7, Cognitive Systems Monographs (Springer Verlag), pp. 361–382.
- Eiben, A. and Smith, J. (2003). *Introduction to Evolutionary Computing* (Springer).
- Eiben, A. E., Elia, D. and van Hemert, J. I. (1999). Population dynamics and emerging mental features in AEGIS, in W. Banzhaf, J. Daida, A. Eiben, M. Garzon, V. Honavar, M. Jakiela and R. Smith (eds.), *Proceedings of the Genetic and Evolutionary Computation Conference*, Vol. 2 (Morgan Kaufmann, Orlando, Florida, USA), ISBN 1-55860-611-4, pp. 1257–1264.
- Epstein, J. M. and Axtell, R. (1996). *Growing artificial societies: social science from the bottom up* (MIT Press, Cambridge, MA.).
- Ficici, S., Watson, R. and Pollack, J. (1999). Embodied evolution: A response to challenges in evolutionary robotics, in J. L. Wyatt and J. Demiris (eds.), *Proc. of the Eighth European Workshop on Learning Robots*, pp. 14–22, URL citeseer.ist.psu.edu/article/ficici99embodied.html.
- Gaussier, P., Moga, S., Banquet, J. and Quoy, M. (1998). From perception-action loops to imitation processes: A bottom-up approach to learning by imitation, *Applied Artificial Intelligence* **12**, 7-8, pp. 701–727.
- Gilbert, N., den Besten, M., Bontovics, A., Craenen, B., Divina, F., Eiben, A. E., Griffioen, A. R., Hévézi, G., Lörincz, A., Paechter, B., Schuster, S., Schut, M., Tzolov, C., Vogt, P. and Yang, L. (2006). Emerging artificial societies through learning, *Journal of Artificial Societies and Social Simulation* **9**(2).
- Hayes, G. and Demiris, J. (1994). Robot controller using learning by imitation, in *Proc. 2nd Int. Symp. on Intelligent Robotic Systems*.
- Hinton, G. and Nowlan, S. (1987). How learning can guide evolution, *Complex Systems* **1**, pp. 495–502.
- Ijspeert, A., Hallam, J. and Willshaw, D. (1998). From lampreys to salamanders: evolving neural controllers for swimming and walking, in R. Pfeifer, B. Blumberg, J.-A. Meyer and S. Wilson (eds.), *From Animals to Animats, Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior* (MIT Press, Cambridge, MA, USA), pp. 390–399.
- Johnson, M. and Demiris, Y. (2004). Abstraction in recognition to solve the correspondence problem for robot imitation, in *Proc. Towards Autonomous Robotic Systems (TAROS)*, pp.

- 63–70.
- Johnson, M. and Demiris, Y. (2005). Perspective taking through simulation, in *Proc. Towards Autonomous Robotic Systems (TAROS)*.
- Kaelbling, L., Littman, M. and Moore, A. (1996). Reinforcement learning: A survey, *Journal of Artificial Intelligence Research* **4**, pp. 237–285.
- Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, ISBN 0262111705.
- Krasnogor, N. (2002). *Studies on the Theory and Design Space of Memetic Algorithms*, Ph.D. thesis, University of the West of England, URL <http://www.cs.nott.ac.uk/~nxx/PAPERS/thesis.pdf>, supervisor: Dr. J.E. Smith.
- Lamarck, J. B. (1809). *Philosophie zoologique, ou Exposition des considérations relatives à l'histoire naturelle des animaux* (H.R. Engelmann).
- Mataric, M. J. (2000). Getting humanoids to move and imitate, *IEEE Intelligent Systems and their Applications* **15**, 4, pp. 18–24.
- Mayley, G. (1996). Landscapes, learning costs, and genetic assimilation: Modeling the evolution of motivation, *Evolutionary Computation* **4**, 3, pp. 213–234, URL citeseer.ist.psu.edu/mayley96landscapes.html.
- Menczer, F. and Belew, R. (1996). From complex environments to complex behaviors, *Adaptive Behavior*, **4**, pp. 317–363, URL citeseer.ist.psu.edu/menczer96from.html.
- Mitchell, M. and Forrest, S. (1994). Genetic algorithms and artificial life, *Artificial Life* **1**, 3, pp. 267–289, URL citeseer.ist.psu.edu/mitchell93genetic.html.
- Moga, S. (2000). *Apprendre par imitation: une nouvelle voie d'apprentissage pour les robots autonomes*, Ph.D. thesis, Université de Cergy-Pontoise.
- Moga, S. and Gaussier, P. (1999). A neuronal structure for learning by imitation, in D. Floreano, J.-D. Nicoud and F. Mondada (eds.), *Lecture Notes in Artificial Intelligence - European Conference on Artificial Life ECAL99*, pp. 314–318.
- Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klapacz, A., Magnenat, S., Zufferey, J.-C., Floreano, D. and Martinoli, A. (2009). The e-puck, a robot designed for education in engineering, in *Proc. 9th Conference on Autonomous Robot Systems and Competitions*, pp. 59–65.
- Moscato, P. (1999). A gentle introduction to memetic algorithms, in D. Corne, F. Glover and M. Dorigo (eds.), *New Ideas in Optimisation* (McGraw-Hill), p. Chapter 14.
- Munroe, S. and Cangelosi, A. (2002). Learning and the evolution of language: the role of cultural variation and learning costs in the baldwin effect, *Artif. Life* **8**, 4, pp. 311–339, doi:<http://dx.doi.org/10.1162/106454602321202408>.
- Nehaniv, C. and Dautenhahn, K. (eds.) (2007). *Imitation and Social Learning in Robots, Humans and Animals* (Cambridge University Press).
- Nehaniv, C. L. and Dautenhahn, K. (2002a). The correspondence problem, in *Imitation in animals and artifacts* (MIT Press, Cambridge, MA, USA), pp. 41–61.
- Nehaniv, C. L. and Dautenhahn, K. (eds.) (2002b). *Imitation in Animals and Artefacts* (MIT Press).
- Nelson, L., Barlow, G. J. and Doitsidis, L. (2009). Fitness functions in evolutionary robotics: A survey and analysis, *Robotics and Autonomous Systems* **57**, 4, pp. 345 – 370.
- Nolfi, S. and Floreano, D. (1999). Learning and evolution, *Auton. Robots* **7**, 1, pp. 89–113, doi:<http://dx.doi.org/10.1023/A:1008973931182>.
- Nolfi, S. and Floreano, D. (2000a). *Evolutionary Robotics* (MIT Press).
- Nolfi, S. and Floreano, F. (2000b). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines* (The MIT Press, Cambridge, MA. / London).
- Nolfi, S. and Parisi, D. (1995a). Learning to adapt to changing environments in evolving neural networks, Tech. Rep. 95-15, Institute of Psychology, National Research Council, Rome, Italy, URL citeseer.ist.psu.edu/nolfi96learning.html.
- Nolfi, S. and Parisi, D. (1995b). Learning to adapt to changing environments in evolving neural networks, Tech. Rep. 95-15, Institute of Psychology, National Research Council,

56 | Bibliography

- Rome, Italy, URL citeseer.ist.psu.edu/nolfi96learning.html.
- Reynolds, R. G. (1999). Cultural algorithms: Theory and applications, in D. Corne, M. Dorigo and F. Glover (eds.), *New Ideas in Optimization* (McGraw-Hill, London), pp. 367–377.
- Richerson, P. and Boyd, R. (2005). *Not by genes alone: how culture transformed human evolution* (University of Chicago Press).
- Ruppin, E. (2002). Evolutionary autonomous agents: A neuroscience perspective, *Nature Reviews Neuroscience* **3**, pp. 132–141, URL citeseer.ist.psu.edu/ruppin02evolutionary.html.
- Scassellati, B. (1999). Knowing what to imitate and knowing when you succeed, in *Proc. AISB'99 Symposium on Imitation in Animals and Artefacts*, pp. 105–113.
- Smith, R., Bonacina, C., Kearney, P. and Merlat, W. (2000a). Embodiment of Evolutionary Computation in General Agents, *Evolutionary Computation* **8**, 4, pp. 475–493, URL citeseer.ist.psu.edu/smith01embodiment.html.
- Smith, R., Bonacina, C., Kearney, P. and Merlat, W. (2000b). Embodiment of Evolutionary Computation in General Agents, *Evolutionary Computation* **8**, 4, pp. 475–493.
- Sutton, R. and Barto, A. (1998). *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA), URL citeseer.ist.psu.edu/sutton98reinforcement.html.
- Todd, P. M. and Miller, G. F. (1990). Exploring adaptive agency ii: simulating the evolution of associative learning, in *Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats*, ISBN 0-262-63138-5, pp. 306–315.
- Todd, P. M. and Miller, G. F. (1991). Exploring adaptive agency ii: simulating the evolution of associative learning, in J.-A. Meyer and S. Wilson (eds.), *From Animals to Animats, Proceedings of the First International Conference on Simulation of Adaptive Behavior* (MIT Press, Cambridge, MA, USA), ISBN 0-262-63138-5, pp. 306–315.
- Trianni, V. (2008). *Evolutionary Swarm Robotics: Evolving Self-Organising Behaviours in Groups of Autonomous Robots* (Springer).
- Turney, P., Whitley, D. and (eds.), R. A. (1996). Evolution, learning, and instinct: 100 years of the baldwin effect, *Special Issue of Evolutionary Computation*, **4**(3).
- Watson, R. A., Ficici, S. G. and Pollack, J. B. (2002). Embodied evolution: Distributing an evolutionary algorithm in a population of robots, *Robotics and Autonomous Systems* **39**, 1, pp. 1–18, URL <http://eprints.ecs.soton.ac.uk/10620/>.
- Winfield, A. F. T. and Griffiths, F. (2010). Towards the emergence of artificial culture in collective robot systems, in P. Levi and S. Kernbach (eds.), *Symbiotic Multi-robot Organisms: Reliability, Adaptability, Evolution* (Springer).