Discovering and Maintaining Behaviours Inaccessible to Incremental Genetic Evolution Through Transcription Errors and Cultural Transmission

James M. Borg[1], Alastair Channon[1] and Charles Day[1]

[1] Research Institute for the Environment, Physical Sciences and Applied Mathematics, Keele University, ST5 5BG, UK
{j.borg, a.d.channon, c.r.day}@epsam.keele.ac.uk

Abstract

In this work the question of whether the introduction of both transcription errors and cultural transmission, in the form of learning by imitation, can enable the evolution of behaviours *inaccessible* to incremental genetic evolution alone is assessed. To answer this a neural network model using a hybrid of two different networks was implemented: one capable of demonstrating reactive qualities, the other controlling deliberative goal selecting behaviours. Animats using this model were evolved in an adaptation of the environment proposed by Robinson et al. (2007) to solve increasingly difficult tasks. Simulations were run on populations with and without learning by imitation to assess the relative success of each strategy, leading to the conclusion that populations with learning by imitation can successfully demonstrate the most complex behaviour, which was empirically found to be inaccessible to non-learning populations.

Introduction

In this paper we present work showing animats in a virtual environment learning behaviours through imitation that are inaccessible to incremental genetic evolution alone. Learning by imitation is often considered to be a mechanism of social information transfer (Cavalli-Sforza and Feldman, 1981; Whiten and van Schaik, 2007), leading to what may be described as social or cultural learning. By combining population learning and individual learning in the same evolutionary system it is possible to make use of both global and local search: global search through the underlying (multi-generational) genetic algorithm and local search through individual (lifetime) learning (Hinton and Nowlan, 1987). It has been demonstrated by Best (1999) that by using cultural learning in place of individual learning on a more challenging version of the Hinton and Nowlan (1987) problem, it is possible to improve the speed at which a population of agents discover an adaptive goal. Cultural learning has the added advantage of allowing individuals to pass on learnt information to other members of the population, and so preserving extra-genetic information for the next generation. Beyond its uses in evolutionary optimisation and search, cultural and social learning is also a well known natural phenomenon with various species using social learning mech-

anisms such as imitation, emulation, teaching and the use of public information to produce adaptive behaviours in dynamic and challenging real world environments (Whiten and van Schaik, 2007; Reader and Biro, 2010).

A number of studies have investigated the effect learning by imitation has on populations of evolving neural networks (Best, 1999; Cangelosi et al., 2006; Acerbi and Parisi, 2006; Acerbi and Nolfi, 2007; Curran and O'Riordan, 2007; Marriott et al., 2010). In much of the literature these imitating neural networks are referred to as agents, with some, as is the case in this work, even taking on the role of animats or autonomous agents in virtual environments (Marriott et al., 2010). It is the aim of this work to investigate whether learning by imitation in a population of neural networks enables behaviours that are deemed to be inaccessible to incremental genetic evolution, to be learned and maintained. In order to test our claims an increasingly complex virtual environment is used in which animats' behaviours are evaluated. It is expected that without learning these animats will only be able to exhibit a limited set of behaviours, whereas animats learning through imitation should evolve in such a way to allow access to all categories of behaviour.

Incremental Genetic Evolution

Long-term incremental evolution necessarily uses converged populations, which can be referred to as species (or quasi species). In genetic algorithms (GAs) this is referred to as the Species Adaptation Genetic Algorithm or SAGA approach (Harvey, 2001). The SAGA approach impacts on the way populations evolve: recombination will have a far smaller effect on the motion of the population than in a standard GA, as each species is already genetically similar, leaving mutation as the primary driving force behind evolution. Mutation can be substantially effective in spaces percolated by neutral networks: pathways of level fitness through the fitness landscape. In this case genotypes can vary while still producing similar phenotypes and behaviours. When phenotypes of higher fitness are found the population converges onto them. This incremental approach enables species of animats to discover and converge upon an easily accessible

solution. However, if there is no neutral or incremental path between the corresponding basic behaviour and fitter ones, the population will struggle to move away from these sub-optimal behaviours. Figure 1 depicts a mock example.

One approach to solving the problem of sub-optimal convergence is to increase the rate at which mutation is applied, potentially allowing the population to explore more of the fitness landscape and so discover new fitness peaks. However, there are problems with this approach: as mutation rates increase, the evolutionary search strategy begins to resemble random search, with larger mutation rates making it increasingly difficult for the population to maintain solutions. The point at which mutation becomes so large that favourable structures discovered by evolution are lost more frequently than they are found is known as the error threshold. Ochoa et al. (1999) and others have demonstrated a link between error thresholds and optimal mutation rates in evolutionary algorithms.

Discovering and Maintaining Inaccessible Solutions: Transcription Errors and Imitation

To solve the issue of sub-optimal population convergence without crossing the error threshold, noise is often added to the fitness landscape via the genotype to fitness map. However, where such noise is in the phenotype to fitness section of that map, its ability to aid in the transition between peaks (or more accurately between neutral networks) is limited. By instead incorporating noise into the genotype to phenotype map, as with transcription errors, behaviours inaccessible to incremental genetic evolution may be exhibited reliably by individuals while leaving the genotype untouched. It can be useful to view such noise as a type of unguided individual learning.

In order to maintain successful behaviours in the population, some form of extra-genetic learning needs to take place. The model employed in this work makes use of imitation through interactions between teachers and pupils to facilitate the transmission of learnt behaviours (Cangelosi et al., 2006; Acerbi and Parisi, 2006; Acerbi and Nolfi, 2007; Curran and O'Riordan, 2007). As in Curran and O'Riordan (2007) pupils follow teachers in a mock evaluation on a set of environments. As both teacher and pupil receive the same environmental input the teacher's output may be used as a target pattern for error backpropagation, reducing the pupil's output error compared to that of the teacher. By learning in this way pupils are able to imitate the behaviours exhibited by teachers, thus maintaining behaviours in the population that would have been lost in incremental genetic evolution.

Neuroevolution of Deliberative Behaviours

This work uses populations of neural networks embodied in animats. The neural network architecture used here is a hybrid of two different networks: the first controlling the high level deliberative behaviours of the animat, and the sec-

ond controlling the animat's reactive capabilities (Robinson et al., 2007). By making use of both reactive and deliberative mechanisms, neural architectures of this sort are able to seek long term goals while also reacting to unforeseen events ultimately enabling the evolution of complex problem solving abilities. To demonstrate these problem solving abilities Robinson et al. (2007) developed a complex problem called the 'river crossing' or RC task. The RC task required animats to find a single reward-giving *Resource* in a 2D grid-world environment containing a number of obstacles. Alongside Resource objects animats could encounter *Water*, *Grass*, *Traps* and *Stones*. Grass objects made up the majority of the environment and where seen as neutral space for the animats to move across; Trap objects were immediately lethal, as were Water objects, which were placed in such a way to resemble an unbroken river cutting the animat's path to the Resource. In order to cross the river animats were required to pick up Stone objects, which could be carried at no cost to the animat, and place them in the same cells as Water thus negating their lethality. Once a continuous bridge of Stones over the river had been built animats could access the Resource. To succeed at the RC task animats were required to evolve with no *a priori* knowledge of the world; each new environment was unique and animats had no concept of co-ordinates, making solutions such as 'move five steps to the right' impossible, instead animats evolved goals and sub-goals such as 'go to resource','avoid traps' or 'head to nearest stone' which then allowed the network to navigate the animat towards these goals. Despite the RC task being reasonably complex, Robinson et al. (2007) demonstrated that it could be solved by initially converged populations of animats using only incremental genetic evolution. To test our hypothesis a more complex version of the RC task has been developed: the RC+ task.

The RC+ Task

An important aspect of the RC task was that individuals were evaluated on increasingly difficult environments. In Robinson et al. (2007), animats were first shown a map with no river blocking their path; then a river with a width of one cell was introduced, followed by a final environment containing a river with a width of two cells. Stone and Trap objects were of a consistent number throughout all tests giving animats equal exposure in each environment. The RC+ task makes the task harder in regard to both river width and exposure to Stone objects. The number of environments an animat is evaluated on is increased from three to five, with environments becoming increasingly difficult to solve due to river width increasing from zero cells to four cells. To add to the difficulty further, the number of Stone objects gradually decreases from twenty in the first environment to zero in the final environment, making each environment more challenging to the point where the final environment cannot be completed by building a bridge. In order to make the final envi-
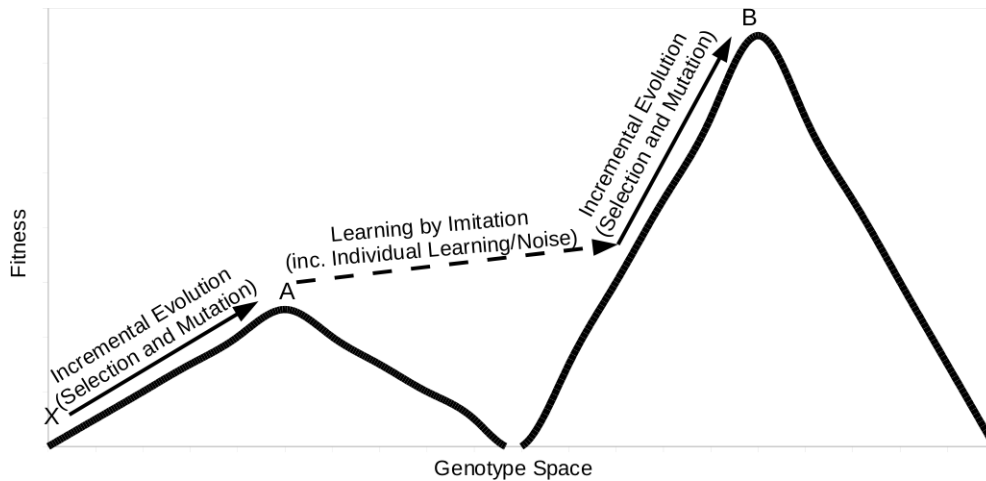
Fig. 1: A species starting from point X on the above mock fitness landscape would achieve peak A by way of the hill climbing strategy adopted by incremental genetic evolution (driven primarily by mutation and selection). Gradient-based learning amongst such a species would ordinarily also be restricted to peak A. The inclusion of both noise in the genotype to phenotype map and learning by imitation can enable the species to jump across areas of lower fitness to higher peaks (inaccessible to hill climbing alone), where incremental genetic evolution and learning can resume hill climbing.

ronment solvable two extra objects, *Object A* and *Object B,* are introduced into the environment. Object A and Object B are rare objects, with only one instance of each found in each environment. Like Stones, Object A and Object B may be carried at no cost to the animat and placed upon any square or object. If an animat happens to place both Object A and Object B on a square containing Water (notionally forming a floating raft that carries the animat to the resource), a reward equal to that of the Resource is received and the animat is considered to have successfully solved the environment. In short, an alternate Resource may be constructed out of the three other objects (Object A, Object B and Water), removing the need to build bridges but still requiring agents to be driven towards the Resource when Water is not present. The RC+ task is impossible to solve with incremental genetic evolution alone. To solve it, animats are required to engage with Water, Object A and Object B while still avoiding Traps and uncovered Water, and to also be able to reach the Resource in the absence of Water (the simplest sub-solution to evolve). The rarity of both Object A and Object B adds to the difficulty of the RC+ task as animats must now evolve to be driven to towards Object A and Object B despite potentially very little exposure during their time in the environment.

The Model

Animat movement is controlled by a hybrid neural network embodying both reactive and deliberative qualities. This hybrid network may be broken down into two network models: a *shunting network* and a *decision network*, with the decision network passing information on to the shunting network which in turn controls the animat's movement. The

shunting network is not directly exposed to any evolution or learning. The deliberative network on the other hand is exposed to both evolution and learning, enabling the evolution and inheritance of animat behaviour.

The Shunting Network

Shunting networks are a specialised form of neural network making use of what is known as the shunting model (Yang and Meng, 2000). The inspiration for the shunting model came from Yang and Meng's (2000) desire to develop motion planning systems capable of reacting quickly in real-time environments, thus allowing robotic agents to exhibit robust and collision-free motion planning behaviours. Instead of directly specifying behaviours, the shunting model maps network outputs onto environmental outputs (within an internal map of the environment) which are propagated across the environment to form an activity landscape. This activity landscape is used by the agent to control movement through the environment, by dynamic gradient ascent of the landscape. In their model, Yang and Meng (2000) demonstrated a neural network composed of an n-dimensional lattice of neurons, with each neuron representing a possible state in the system. By using neurons to represent states in this way it is possible to represent any system which is capable of being fully described by a set of discrete states.

The environment used for the RC and RC+ tasks is a simple 2D grid-world consisting of 20×20 cells, with each cell representing a position in co-ordinate space. Each position in the grid-world may be occupied by any number of objects found in the RC+ environment (Resource, Water, Trap, Grass, Object A and Object B), allowing the system to be fully described by a set of discrete states, thus enabling the

use of the shunting model to direct animat movement across the RC+ environment and ensuring a simple one-to-one relationship between neurons and geographical locations.

In Yang and Meng (2000), two versions of a transition function for specifying inter-neuron dynamics were developed: one which controlled activity saturation in the network and one which did not. Consistent with the findings of Robinson et al. (2007), we found activity saturation not to be a problem exhibited by networks in the RC+ task, enabling the use of the simpler transition function in equation 1.

$$\frac{dx_i}{dt} = -Ax_i + I_i + \sum_{j=1}^{k} w_{ij} [x_j]^+ \qquad (1)$$

Alpha (A) represents the passive decay rate, which determines the degree to which each neuron's activity diminishes towards an idle state. The functions $[x]^+$ is $max(0, x)$. The connection weight (or synapse strength) $w_{i,j}$ between neurons i and j is the Euclidean distance between cells i and j within the receptive field. k is the receptive field size and here is set to 4, corresponding to the four cells orthogonally surrounding cell i. Iota (I) is equal to E in the case of the target, and $-E$ for an obstacle, where E is a large integer.

In the case of the RC and RC+ tasks Iota values are limited to 15, -15 and 0, representing the target resource, an obstacle and neutral space respectively. The result of using a transition function with these values are 2D environments with large peaks at the sites of target states, large troughs in cells occupied by obstacles, and large amounts of neutral space through which neuron activity from targets may spread. Using the shunting model to control animat movement allows for goals such as 'head for resource while avoiding traps' or 'place carried stones on water' to be easily achieved.

The Decision Network

The role of the decision network is to set the Iota values for object states found in the RC and RC+ task. Using the decision network animats can set the desirability of object states in relation to their current environmental inputs, allowing them to manipulate the shunting network's activity landscape and so combine multiple actions such as 'pick up the closest stone' and 'place stone on water' to create complex behaviours.

As in Robinson et al. (2007), the decision network is simply a feed-forward multi-layer perceptron with one hidden layer comprising of four hidden units. The input layer is capable of representing the animat's current state in the environment including whether or not the animat is currently carrying a movable object (Stone, Object A, Object B), with each movable object having a dedicated carrying input. Inputs taken by the input layer are single values of 1 or 0, representing the presence of the object in the same cell as the animat. These input values are fed through to the hidden layer neurons via weighted connections in the range [-1,1]. At each hidden unit the weighted sum of inputs is passed through a hyperbolic tangent activation function to produce hidden layer outputs. In the RC+ task the output layer is made up of sixty-seven neurons representing the Iota values of all sixty-four possible environmental states (excluding Grass objects whose Iota values are always set to 0 and therefore do not need be represented in the decision network) and a pick-up/put-down output for each non-static object (Stone, Object A, Object B). At each output neuron the sum of all weighted connections is passed through a hyperbolic tangent activation function with fixed thresholds: neurons outputting within the range [-0.3:0.3] are set to output 0, while all outputs over 0.3 resolve to 1 and all outputs below -0.3 resolve to -1.

For outputs representing the pick-up/put-down actions output values of -1 cause the animat to put down the specified object they are carrying, values of +1 causing animats to pick up the movable objects they are currently sharing a cell with providing the animat is not already carrying an object of that type. For all other outputs, resolved output values set the Iota values to be used in the shunting network. So if an output neuron has a negative output, all objects of that class found in the environment at that point in time will have their activations set to -15; for positive outputs to +15. Any object resulting in an Iota value of 0 will remain neutral, causing their activation values in the shunting network to be solely based on the propagated activations of other objects. The resulting environment will contain a number of peaks of high activity and troughs of low activity, gradually propagating activity through neighbouring neutral cells.

Figure 2 shows two of the five potential environments an animat may observe in the RC+ task, and the corresponding activity landscapes given certain outputs from the decision network. The first environment represents the initial challenge an animat must complete, where only traps stand in the way of a resource. As can be seen by this environment's activity landscape, the Iota value associated with the resource has been set to be positive resulting activity propagating from the resource over the surrounding neutral space. The second environment represents the second challenge, to cross a river before having access to the resource. In this environment's case, activation propagation from the resource has been impeded by the decision network outputting negative Iota value for Water objects. Negative activity repels animats from objects with negative Iota values; however positive activation can been seen coming from the Object B object, providing a hill-climbing route for the animat to take in activity space.

Evolution of the Decision Network

To evolve the decision network a steady-state genetic algorithm was used. At each iteration two animats were selected from the surviving population to be evaluated in tournament selection, with the worst performing animat being
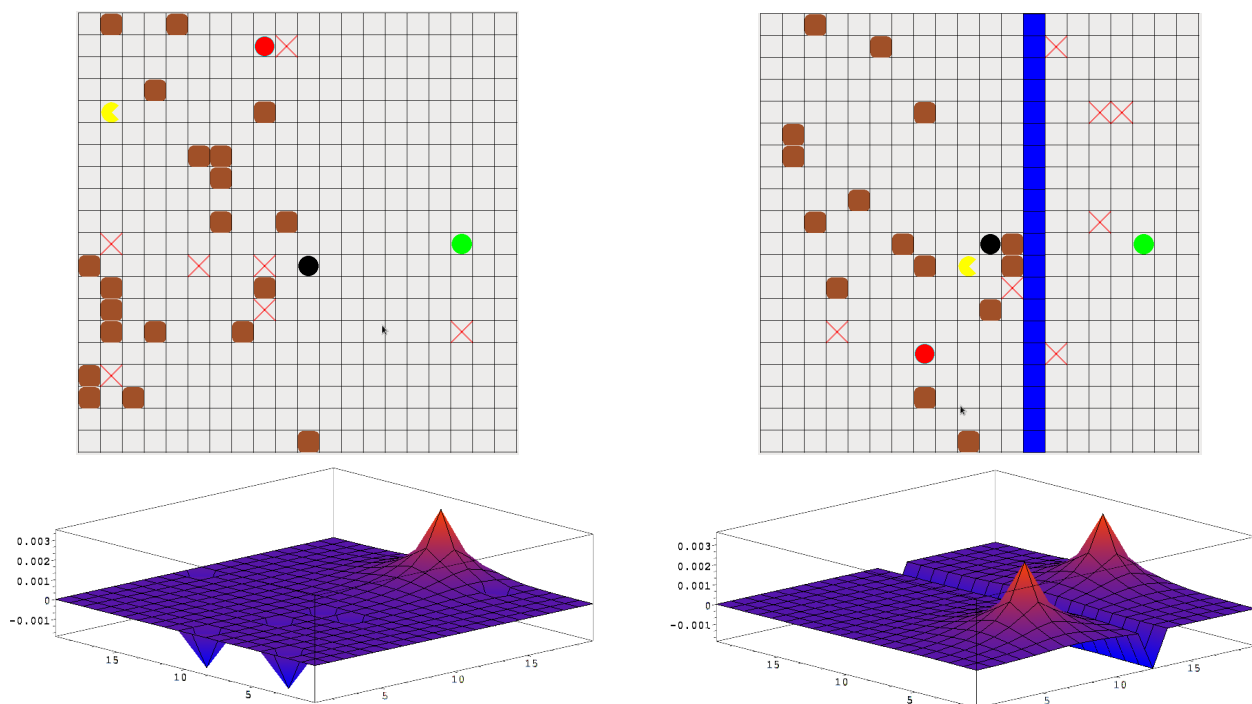
Fig. 2: Two environments with their activity landscapes (given certain outputs from the decision network - see main text). Animat=yellow, Stones=brown, Resource=green, Object A=black, Object B=red, Traps=crosses, Water=blue.

replaced by the progeny of the better performer. The competing animats are evaluated in five increasingly difficult environments. If during evaluation an animat fails to complete an environment, the evaluation is terminated. Fitness is set to be the number of environments successfully completed by an animat during evaluation.

An animat's genotype consists of a set of floating point values each in the range [-1,1], which are transcribed into the connection weights in the animat's decision network. The genotype and the decision network are stored separately, so any learning that may take place during an animat's lifetime will only affect the decision network: no changes are made to its genotype after an animat is initially created. New animats are the offspring of two other animats from the current population: one tournament winning animat and one randomly selected animat. The child's genotype is created first through recombination of the parents' genotypes; for this operation single-point crossover is used with the point of crossover being a randomly selected point in either parent's genotype. Each loci in an animat's genotype represents exactly the same connection weight as in any other animat's genotype, with all genotypes being of length $L = 308$. Mutation follows recombination; each point has a probability $P_{mut} = 1/L$ of having a random value from $N(0,0.4)$ added to it, with the resulting values being bounded within the range [-1,1]. Once the genotype has been constructed it is written to the new animat's decision network; this process is

referred to as transcription. During transcription two randomly selected connection weights are overwritten with a new random value selected from a discrete uniform distribution $U(-1,1)$. The weights now present in the decision network dictate the animat's future behaviours within each environment.

Learning in the Decision Network

Following reproduction new animats are afforded the opportunity to learn from a teacher via error-backpropagation. This method of teacher-pupil backpropagation has been previously employed by Curran and O'Riordan (2007). However, the teacher-pupil scenario used in this work differs in a number of ways. In the learning model used by Curran and O'Riordan (2007), teachers were selected from the population based upon their fitness and then assigned n pupils to teach. We contend that in nature absolute fitness is very difficult to assess. To resolve this issue, the current tournament-winning parent is assigned the role of teacher, with the parent's most recent progeny assigned the role of pupil.

There are also differences in the way error-backpropagation is used to teach pupils in this model compared to that of Curran and O'Riordan (2007). As with our model, Curran and O'Riordan (2007) allowed pupils to hitchhike on the back of the teacher during a mock evaluation, with inputs shared between teacher and pupil and using the teacher's output pattern as a target pattern

for the pupil to learn. The learning method employed by Curran and O'Riordan (2007) permitted pupils to learn from the target pattern until the error between child and parent outputs were minimised to a satisfactory level. In our model pupils are only presented with the current teacher's output once every simulation time step (immediately after the teacher's decision network's inputs, activations and outputs are updated). If a teacher happens to move through the environment in such a way that both inputs and outputs remain the same, the child will be presented with many opportunities to learn a given target input-output pattern. However, if the teacher moves around the environment via many different input combinations, the student will have the opportunity of potentially witnessing many different target outputs but at the cost of having very little time to minimise error. Imitating in this manner enables the population to retain favourable behaviours not coded for genetically, whilst not undermining the incremental genetic evolutionary process.

Experimentation

At each iteration of the model two individuals are taken from the population to be evaluated on a series of five environments/maps. All maps have seven Trap objects placed randomly on the map, one reward-giving Resource, one Object A, one Object B, and $20 - (5 \times riverwidth)$ Stone objects. River width varies from an initial width of zero, increasing by one cell per map. During evaluation individuals must successfully reach the Resource or place Object A and Object B onto a cell containing Water; any animat failing to do so within 100 steps or dying by means of a Trap or uncovered Water is not permitted to attempt the next environment.

Fitness in the model is determined to be the number of maps successfully completed in the current tournament iteration, with individual fitness being set to zero before each evaluation. The individual achieving the highest fitness is allowed to reproduce, with the weaker individual being replaced by the progeny of the tournament winner and a randomly selected animat. This steady-state approach maintains the population at a size of 100 individuals.

After reproduction the child is allowed to learn via error-backpropagation from its tournament winning parent. The child follows its parent in a mock evaluation, with the child's inputs being set to those of the parent. Learning takes place for as along as the parent is being evaluated. Once the parent either fails to complete a map or completes all five environments, learning is terminated. At each step through the evaluation the child attempts, via error-backpropagation with a learning rate of $\delta = 1$, to learn to imitate the parent's output for the current inputs.

Three strategies are used in this model: two without learning and one with learning. Populations of animats with no access to learning fall into two categories. The first, known as *Non-Learners(1)*, having a mutation rate and transcrip-

tion error equal to that use by learning populations. As populations of Non-Learners(1) have no way of assimilating transcription errors back into the genotype it may be seen as giving learning populations, known as *Learners*, an unfair advantage. With this in mind a second of category of non learners, known as *Non-Learners(2)*, are also evaluated. Non-Learners(2) do not have transcription errors, and instead have a mutation rate equal to that of the original mutation rate plus two transcription errors: $P_{mut2} = 3/L$.

To test the ability of each strategy to exhibit the behaviour necessary to complete the most difficult map, fifteen populations of each learning strategy were simulated. Each simulation lasted a maximum of 5,000,000 tournaments. In each simulation the best individual's fitness and the mean population fitness were recorded at intervals of 500 tournaments. The maximum fitness an individual could achieve was five, which directly relates to the successful completion of all five evaluation environments, the fifth environment being impossible to complete by bridge building and so requiring the combination of Object A and Object B on Water. For a population to be considered as adequately completing the fifth map, a fitness of five must have been recorded by the fittest individual at ten recorded tournaments with at least five of these tournaments being unbroken by a sub-optimal result. This ensures that the complex behaviour tested for is not only found but also maintained by the population.

Results

Table 1 shows results from the fifteen populations of animats using the Non-Learners(1) strategy: the mean, best and worst number of tournaments required to solve each map, across the fifteen populations (runs), and the proportion of populations that were successful in solving each map. Of the Non-Learners(1) populations over 90% were able to complete maps 1 to 4 but no population was able to demonstrate a successful solution to map 5. Populations of animats using the Non-Learners(2) strategy also demonstrated a high level of proficiency when completing maps where the bridge building solution is effective, though with a lower proportion of populations able to complete map 4 (see table 2). This may be due to the higher mutations rate used in the Non-Learners(2) strategy causing the destruction of potentially beneficial behaviours before they can proliferate through the population. To complete map 4 animats had to be stricter (more consistent) in their use of Stone objects. Despite this behaviour being reachable using incremental genetic evolution it is within a small area of weight-space, causing it to be potentially lost with higher mutation rates. Neither non-learning strategy was able to discover the precise behaviour necessary to complete map 5, so failures recorded in tables 1 and 2 were not due to a sufficient behaviour being discovered but not maintained: the map 5 solution was simply never found, empirically demonstrating the inaccessibility of map 5 to incremental genetic evolution alone.

Map	Mean	Best	Worst	Stdev	Success
1	1200	500	3500	996	100%
2	502571	11000	2152500	738090	100%
3	1568000	34000	4429500	1501336	93%
4	1613786	58000	4432500	1506065	93%
5	N/A	N/A	N/A	N/A	0%

Tab. 1: Non-Learners(1): Mean, best, worst number of tournaments required to solve each map.

Map	Mean	Best	Worst	Stdev	Success
1	1400	500	3000	784	100%
2	81692	4500	252500	96805	100%
3	1801286	12500	4987000	1502754	93%
4	2193385	41500	4466500	1497156	87%
5	N/A	N/A	N/A	N/A	0%

Tab. 2: Non-Learners(2): Mean, best, worst number of tournaments required to solve each map.

Table 3 shows results from animats using the Learners' strategy. Unlike non-learning strategies, Learners are able to complete map 5 and thus exhibit the complex behaviour tested for in this work a third of the time, proving the hypothesis that learning by imitation is capable of enabling populations of animats to discover behaviours found to be inaccessible to incremental genetic evolution alone. However, Learners are seemingly less likely to discover and maintain solutions to maps 3 and 4 than non-learning animats.

Figure 3 charts the mean fitness of the best performing population from each learning strategy. From this graph it can be observed that Learners bypassed the sub-optimal bridge building solution once the population had (for some time) been evaluated on maps with rivers. The incremental nature of the evolution in this model causes the majority of the population to rapidly converge on the optimal solution once it has been discovered. Without learning, this optimal behaviour cannot be found. In this model incremental genetic evolution leads to convergence on sub-optimal solutions in non-learning populations, making it impossible for the discovery of the optimal behaviour. By combining learning by imitation and incremental genetic evolution in a

Map	Mean	Best	Worst	Stdev	Success
1	1533	500	5000	1302	100%
2	512333	9500	2026000	616376	100%
3	2484455	5600	4340500	1395760	73%
4	2458800	88500	4211500	1861794	33%
5	1843200	83500	3851000	1631808	33%

Tab. 3: Learners: Mean, best, worst number of tournaments required to solve each map.
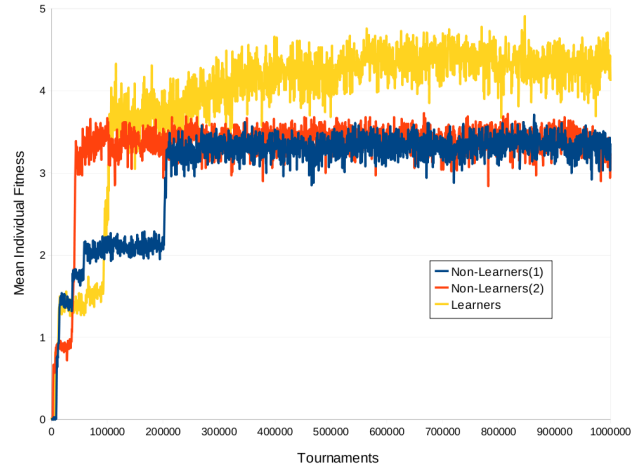

Fig. 3: Graph showing the mean fitness in the best performing populations for each learning strategy. Populations learning by imitation demonstrated the ability to converge on more complex behaviours, thus achieving a higher fitness. Neither non-learning strategy is capable of producing the more complex behaviour.

model such as the one presented here, it is possible to not only discover complex behaviours inaccessible to incremental evolution alone, but also to have rapid convergence to a population exhibiting and maintaining that behaviour, thus creating a behavioural tradition or culture (Whiten and van Schaik, 2007). The results found here are broadly consistent with those of Acerbi and Nolfi (2007), who found that the combination of individual and social learning in artificial embodied agents not only allowed for the development of difficult and costly behaviours, but also provided an adaptive advantage over individual learning alone and lead to cumulative cultural evolution.

Conclusions and Future Work

If a learnt behaviour is exhibited and maintained throughout a population for a number of generations it may tentatively be called a tradition or even a culture. According to Whiten and van Schaik (2007) traditions are "consistent habits" that make use of social information transfer. In the model demonstrated here learning by imitation enables social information transfer with behaviours being maintained by converged populations or species giving rise to traditions. The limited set of behaviours observed in this population do not however constitute the category of culture, which is reserved for the maintenance of multiple behaviours by a species. The incremental nature of the model causes sub-optimal behaviours to be phased out of the population. Were greater environmental diversity to be used, it may be possible to evolve a culture rather than a tradition.

The hypothesis presented here was that the introduction of both transcription errors and cultural transmission in the

form of learning by imitation are sufficient to discover and maintain the most complex behaviour possible in the model, while incremental genetic evolution alone is not. The results prove our hypothesis by demonstrating that without learning by imitation the solution to the final environment is never found but with imitative learning all behaviours can be discovered, exhibited and maintained.

One drawback to the model used in this work is the limited set of behaviours available to animats. By using a larger environment with a greater variety of potential states available to the animats and evolving the size and structure of the decision network, it may be possible to demonstrate the evolution of multiple behaviours leading to the emergence of a culture. To investigate more complex behavioural development and the role of imitative learning in the evolution of traditions and cultures, it would be beneficial to implement larger and more dynamic environments and allow for greater evolution in the decision network. A secondary drawback was the simple vertical social transmission mechanism used. The inclusion of intra-generational or oblique cultural transmission has been shown to be both sufficient (Cavalli-Sforza and Feldman, 1981) and beneficial (Acerbi and Parisi, 2006) for the evolution of complex and robust cultural behaviours. Further investigation and application of oblique transmission within models such as that presented here would further benefit our understanding of and ability to achieve the evolution and maintenance of complex cultural traits.

References

Acerbi, A. and Nolfi, S. (2007). Social learning and cultural evolution in embodied and situated agents. In *Proceedings of the IEEE Symposium on Artificial Life*, ALIFE'07, pages 333–340. IEEE.

Acerbi, A. and Parisi, D. (2006). Cultural transmission between and within generations. *Journal of Artificial Societies and Social Simulation*, 9(1).

Best, M. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Adaptive Behavior*, 7(3–4):289–306.

Cangelosi, A., Hourdakis, E., and Tikhanoff, V. (2006). Language acquisition and symbol grounding transfer with neural networks and cognitive robots. In *Proceedings of the 2006 International Joint Conference on Neural Networks*, IJCNN'06, pages 2885–2891.

Cavalli-Sforza, L. L. and Feldman, M. (1981). *Cultural Transmission and Evolution: Quantitative Approach*. Princeton University Press.

Curran, D. and O'Riordan, C. (2007). The effects of cultural learning in populations of neural networks. *Artificial Life*, 13(1):45–67.

Harvey, I. (2001). Artificial evolution: A continuing saga. In *Proceedings of the International Symposium on Evolutionary Robotics From Intelligent Robotics to Artificial Life*, ER '01, pages 94–109. Springer-Verlag.

Hinton, G. and Nowlan, S. (1987). How learning guides evolution. *Complex Systems*, 1:495–502.

Marriott, C., Parker, J., and Denzinger, J. (2010). Imitation as a mechanism of cultural transmission. *Artificial Life*, 16(1):21–37.

Ochoa, G., Harvey, I., and Buxton, H. (1999). Error thresholds and their relation to optimal mutation rates. In *Proceedings of the 5th European Conference on Advances in Artificial Life*, ECAL '99, pages 54–63. Springer-Verlag.

Reader, S. and Biro, D. (2010). Experimental identification of social learning in wild animals. *Learning and Behavior*, 38(3):265–283.

Robinson, E., Ellis, T., and Channon, A. (2007). Neuroevolution of agents capable of reactive and deliberative behaviours in novel and dynamic environments. In *Proceedings of the 9th European conference on Advances in artificial life*, ECAL'07, pages 345–354. Springer-Verlag.

Whiten, A. and van Schaik, C. (2007). The evolution of animal 'cultures' and social intelligence. *Phil. Trans. R. Soc. B*, 362(1480):603–620.

Yang, S. and Meng, M. (2000). An efficient neural network approach to dynamic robot motion planning. *Neural Networks*, 13(2):143–148.