

The Effects of Cultural Learning in Populations of Neural Networks

Dara Curran*

Dept. of Information Technology

National University of Ireland, Galway

email:dara.curran@nuigalway.ie

Colm O'Riordan

Dept. of Information Technology

National University of Ireland, Galway

email:colm.oriordan@nuigalway.ie

Abstract

Population learning can be described as the iterative Darwinian process of fitness-based selection and genetic transfer of information leading to populations of higher fitness and is often simulated using genetic algorithms.

Cultural learning describes the process of information transfer between

*Corresponding Author

individuals in a population through non-genetic means. Cultural learning has been simulated by combining genetic algorithms and neural networks using a teacher/pupil scenario where highly fit individuals are selected as teachers and instruct the next generation. By examining the *innate* fitness of a population, i.e. the fitness of the population measured before any cultural learning takes place, it is possible to examine the effects of cultural learning on the population's genetic makeup.

Our model explores the effect of a cultural learning on a population and employs three benchmark sequential decision tasks as the evolutionary task for the population: connect-four, tic-tac-toe and blackjack. Experiments are conducted with populations employing population learning alone and populations combining population and cultural learning. The paper presents results showing the gradual transfer of knowledge from genes to the cultural process, illustrated by the simultaneous decrease in the population's innate fitness and the increase of its acquired fitness measured after learning takes place.

Keywords: Cultural Learning, Neural Networks, Sequential Decision Tasks, Games, Artificial Life

1 Introduction

Some research has been performed with regard to the combination of both population and life-time learning approaches[45, 29, 46, 53, 49, 61, 21, 20], thus combining the global search power of the underlying genetic algorithm and the finer local search capabilities of gradient descent techniques. Empirically, the combined approach proves to be successful, as the populations tend to converge faster towards a global optimum.

Cultural learning is an alternative model which combines population learning with a modified version of life-time learning that allows populations to pass on knowledge to the next generation through non-genetic means through a process of communication or artifact creation, often achieved through imitation.

Gene-culture coevolutionary theory describes the effects of culture on the genetic makeup of a population. Typical studies on gene-culture co-evolution have focused on humans, although recent work has highlighted similar effects in populations of other species, such as whales[64]. The theory is built on standard population genetics models such as those employed by Cavalli-Sforza and Feldman [18]. In a famous case-study, Laland *et al.* examined the effects of culture on genetic evolution with regard to the human sex ratio employing a mathematical model that analysed the impact of culturally acquired traditions such as sex-selection on the genetics of a human population, showing that culture can play an important role in shaping the genetics of a population[40].

Much research has been conducted in the field of imitation, particularly with respect to robotics and symbol grounding in animals and artifacts[8, 7, 23, 31,

26] and a number of models have been developed to examine the interaction of culture and evolution[16, 17, 11]. In addition, the simulation of culture in populations of artificial organisms has been focus of much research[66, 25, 3, 27, 57, 47, 35, 54, 36, 6, 24, 14, 12, 10].

Bull *et al.* have examined the effects of altering the relative speed of cultural, or meme, evolution compared to gene evolution and found that as the speed of cultural evolution increases, genetic evolution degrades considerably, even halting completely[13]. In the words of Blackmore, cultural evolution ‘takes hold of the leash’ by becoming increasingly important in the evolution and adaptation of a population[9].

These simulations examine the process of cultural learning using very formal, restrictive mathematical models and do not consider the impact of individual, lifetime adaptation in populations of evolving agents. An alternative and more expressive method of simulating cultural learning is the teacher/pupil approach[8, 28, 15]. Highly fit individuals from the population (teachers) are allowed to instruct the next generation (pupils). In this way, important information gained through the life-time of the previous generation is not lost completely and the fitness of the entire population can be improved.

The aim of this work is to examine the effect of cultural learning on the process of evolution using a more descriptive and flexible model than ones used in previous research. In particular, we are interested in observing the effect of cultural learning on the innate fitness of individuals within the population. In other words, the work is not primarily concerned with the objective problem

solving performance of cultural learning but focuses instead on the effects of culture on populations of evolving agents. In this sense, this work is similar to recent work by Borenstein & Ruppin who explored the effect of cultural learning on the innate fitness of populations for a variety of problem solving tasks[10]. However, while their work allowed individuals to impart only innate knowledge, our framework allows individuals to transmit information acquired culturally.

We are interested in observing whether the population learning mechanism will become less important as the cultural learning process takes hold using simulations that employ more complex agents capable of transferring cultural information using a teacher/pupil mechanism. This would be an important result as it would show that cultural learning is capable of causing a direct effect on the genetic makeup of more complex societies.

We have selected three sequential decision bench-mark problems (tic-tac-toe, blackjack and connect-four) to examine the effect of cultural learning in populations of neural networks. While previous work has attempted to evolve game-playing agents using a variety of games[44, 62, 19, 52], none have explicitly employ game-playing as a test-bed for cultural learning experiments.

A series of experiments were conducted using populations employing population-learning alone and populations employing both population and cultural learning. While the effects of cultural learning have been examined before, the results presented diverge from previous work and shed more light on the subject. The model employed allows individuals to acquire cultural traits through the use of imitation, based on the teacher/pupil scenario and individual agents possess a

genotype that translates to a neural network capable of perceiving their environment. Unlike comparable experiments conducted in the past (see [8, 28, 15, 10]), our system gives the evolutionary mechanism greater freedom by eliminating many neural network architectural restrictions (see section 3.1 for more detail). In addition, our model builds on previous models (such as Borenstein’s [10]) by allowing individuals to impart knowledge that they themselves have acquired during their lifetimes.

The remainder of this paper is arranged as follows: Section 2 summarises related research and background material. Section 3 presents the artificial life simulator employed to conduct the experiments. Section 4 describes the experiments undertaken: tic-tac-toe (Section 4.1), blackjack (Section 4.2) and connect-four (Section 4.3). Section 5 provides a discussion of these experiments and Section 6 presents conclusions.

2 Related Work

The following section outlines some background material including learning models and sequential decision tasks.

2.1 Learning Models

A number of learning models can be identified from observation in nature. These can roughly be classified into population, life-time and cultural learning.

2.1.1 Population Learning

Population learning refers to the process whereby a population of organisms evolves, or learns, by genetic means through a Darwinian process of iterated selection and reproduction of fit individuals. In this model, the learning process is strictly confined to each organism's genetic material: the organism itself does not contribute to its survival through any learning or adaptation process.

2.1.2 Life-time Learning

There exist species in nature that are capable of learning, or adapting to environmental changes and novel situations at an individual level. Such learning, known as life-time learning is often coupled with population-based learning, further enhancing the population's fitness through its adaptability and resistance to change. Another phenomenon related to life-time learning, first reported by Baldwin[1], occurs when certain behaviour discovered through life-time learning becomes imprinted onto an individual's genetic material through the evolutionary processes of crossover and mutation. To quote Hinton and Nowlan whose model ([32]) was the first to demonstrate this effect through simulation, "learning can provide an easy evolutionary path towards co-adapted alleles in environments that have no good evolutionary path for non-learning organisms". Subsequent work has further explored the interactions between evolution and learning and shown that the addition of individual lifetime learning can improve a population's fitness[45, 29, 46, 53, 49, 61, 21, 20].

2.1.3 Cultural Learning

Culture can be succinctly described as a process of information transfer within a population that occurs without the use of genetic material. Culture can take many forms such as language, signals or artifactual materials. Such information exchange occurs during the lifetime of individuals in a population and can greatly enhance the behaviour of such species. Because these exchanges occur during an individual's lifetime, cultural learning can be considered a subset of lifetime learning.

A number of approaches have been implemented to simulate cultural learning including fixed lexicons[66, 15], indexed memory[57], cultural artifacts [33, 14] and signal-situation tables[41]. The approach chosen was inspired by the teacher/pupil scenario[8, 28, 15] where a number of highly fit agents are selected from the population to act as teachers for the next generation. Pupils learn from teachers by observing the teacher's verbal output and attempting to mimic it using their own verbal apparatus. As a result of these interactions, a lexicon of symbols evolves to describe situations within the population's environment.

Experiments conducted by Hutchins and Hazlehurst[34] simulate cultural evolution through the use of a hidden layer within an individual neural network in the population. The hidden layer acts as a verbal input/output layer and performs the task of feature extraction used to distinguish different physical inputs.

In previous work by Parisi *et al.*[28], it was suggested that the addition of

noise to a teacher’s verbal output could enhance a population’s ability to retain culturally acquired information. Experiment conducted in our previous work[22] confirmed that small levels of noise introduced to the communication process improved both the formation of a shared lexicon and agent performance.

2.2 Sequential Decision Tasks

Sequential decision tasks are a complex class of problem that require agents to make iterative decisions at many steps throughout the task. Each decision has a direct effect on the agent’s environment and in turn affects its subsequent decisions. Our selection of a number of games was driven by two main factors: games are good examples of sequential decision tasks and many artificial intelligence implementations exist for ready comparison and analysis.

The games we chose as a test-bed for cultural learning are ordered by perceived difficulty, beginning with tic-tac-toe following with the game of blackjack and concluding with the game of connect-four.

3 Simulator Architecture

The simulator implements population and cultural learning. Population learning is simulated using a genetic algorithm which generates successive generations using three operators: selection, crossover and mutation. The algorithm employs an encoding scheme (described in Section 3.4) to convert genetic codes to neural network structures. Each agent in the population is equipped with a neural network responsible for its perception and response to the environment.

The neural network structure is derived from an individual’s gene code at birth.

Cultural learning is implemented using a *vertical* cultural transmission model [11, 4] inspired by Hutchins and Hazlehurst’s model. The approach employs the last hidden layer of each agent’s neural network as the agent’s verbal apparatus. As an agent encounters stimuli in its environment, it responds both behaviourally (emitting a signal through its output nodes) and verbally (emitting a signal through its verbal units (see Figure 1)). Unlike Hutchins and Hazlehurst’s model, the model employed for this work allows the number of verbal units to evolve along with the network structure. Thus, no limitations are imposed on the communication complexity available to the population.

At the end of each generation, a percentage of the population’s fittest networks are selected and are allowed to become teachers for the next generation. The teaching process takes place as follows: a teacher is stochastically assigned n pupils from the population where $n = \frac{N_{pop}}{N_{teachers}}$, N_{pop} being population size and $N_{teachers}$ the number of teachers.

Each pupil follows the teacher in its environment and observes the teacher’s verbal output as it interacts with its environment. Both teacher and pupil receive environmental stimuli and respond with verbal signals. A teaching cycle occurs when the pupil’s output is corrected to more closely resemble the teacher’s using error back-propagation. Once the number of required teaching cycles is completed, the teachers die and the pupils are released into their environment. At the end of their lifetime, the fittest pupils are themselves selected to become teachers for the next generation and can impart the knowledge ac-

quired through previous cultural exchanges. Thus, information is passed down through successive generations.

Owing to the encoding scheme employed (a description of which follows in the next section), which allows a large number of neural network architecture permutations, the last hidden layers of the pupil and teacher may differ in size. Therefore, teaching takes place using the maximum number of compatible nodes. For instance, if a teacher's verbal layer contains four nodes and the pupil's contains six, the pupil will imitate responses received from the teacher's nodes and leave the remaining two dormant. Conversely, if a teacher has more nodes than the pupil, only a portion of its verbal output will be received by the pupil's verbal layer. As a result, there is a certain pressure on the evolutionary process to produce populations of teacher and pupils that have a similar number of nodes within their verbal layers, so that the maximum amount of information can be transmitted.

In previous work by Parisi *et al.*[28], it was suggested that the addition of noise to a teacher's verbal output could enhance a population's ability to retain culturally acquired information. The experiments undertaken in this work employ this noise setting, dubbed cultural mutation, which distorts the teacher's output as perceived by a pupil with probability p .

3.1 Encoding Scheme

One of the most crucial aspects of the simulator is the translation of genetic codes to neural network structures. Many encoding schemes were considered in

preparation of the simulator, prioritising flexibility, scalability, difficulty and efficiency. These included Connectionist Encoding[5], Node Based Encoding [63], Graph Based Encoding[51], Layer Based Encoding[42], Marker Based Encoding[44], Matrix Re-writing[37, 43], Cellular Encoding[30], Weight-based encoding[58, 38] and Architecture encoding [39].

The scheme chosen is inspired by Marker Based Encoding which allows any number of nodes and interconnecting links for each network giving a large number of possible neural network architecture permutations.

Marker based encoding represents neural network elements (nodes and links) in a sequential list[37, 43]. Each element is separated by a marker to allow the decoding mechanism to distinguish between the different types of element and therefore deduce interconnections (see Figure 2).

In this implementation, a marker is given for every node in a network. Following the node marker, the node's details are stored in sequential order on the bit string. This includes the node's label and its threshold value. Immediately following the node's details, is another marker which indicates the start of one or more node-weight pairs. Each of these pairs indicate a back connection from the node to other nodes in the network along with the connection's weight value. Once the last connection has been encoded, the scheme places an end marker to indicate the end of the node's encoding.

There is no requirement for genotypes to be the same size across the population. Two parents possessing gene codes of different lengths will produce two offspring whose gene code lengths are equal to the gene lengths of their parents.

Thus, if parent A has a large genome of length a and parent B has a smaller genome of length b , one of their offspring will have a genome of length a while the other will have one of length b .

3.2 Crossover

As a result of the chosen encoding scheme, crossover may not operate at the bit level as this could result in the generation of invalid gene codes. Therefore, the crossover points are restricted to specific intervals – only whole node or link values may be crossed over.

Two-point crossover is employed in this implementation. Once crossover points are selected, the gene portions are swapped. The connections within each portion remain intact, but it is necessary to adjust the connections on either side of the portion to successfully integrate it into the existing gene code. This is achieved by using node labels for each node in the network. These labels are used to identify individual nodes and to indicate the location of interconnections. Once the portion is inserted, all interconnecting links within the whole gene code are examined. If any links are now pointing to non-existing nodes, they are modified to point to the nearest labeled node.

3.3 Mutation

The mutation operator introduces additional noise into the genetic algorithm process thereby allowing potentially useful and unexplored regions of problem space to be probed. The mutation operator usually functions by making alter-

ations on the gene code itself, most typically by altering specific values randomly selected from the entire gene code. In this implementation, weight mutation is employed. The operator modifies a weight according to a random percentage value chosen randomly from the range -200% to +200%. Mutation can alter the value of a start or end marker, thereby introducing structural novelty into the evolutionary process.

4 Experiments

The games represent the environment in which the agents live. To be successful, an agent must become sufficiently skilled to play each game adequately.

4.1 Tic Tac Toe

Tic-tac-toe, or three in a row is a very simple two player game played on a 3x3 board. Each player is assigned either the X or O symbol and takes turns placing one symbol onto the board at a time. Each player attempts to place three of his/her pieces in a horizontal, vertical or diagonal line of three.

Agents play games against a minimax player whose first move is randomized, allowing agents to play games of some variety. Fitness is assigned according to the length of the game. In other words, agents are rewarded for bringing the game to as close as possible to a draw.

Each agent's neural network structure contains 18 input nodes, 2 for each board position where 01 is X, 10 is O and 11 is an empty square. Nine output nodes corresponding to each board position are used to indicate the agent's

desired move. The node with the strongest response corresponding to a valid move is taken as the agent’s choice. The simulator allows agents to evolve any number of hidden layers each with an unrestricted number of nodes, giving maximum flexibility to the evolutionary process. During the teaching process, a teacher agent plays alongside the pupil. At each move, both the pupil and teacher emit some verbal output in response to the current board position. At every teaching cycle, the pupil’s verbal output is corrected with respect to the teacher’s using error back-propagation.

Populations of 100 agents were generated for these experiments and allowed to evolve for 250 generations. Crossover was set at 0.6 and mutation at 0.02. The cultural learning settings of teacher ratio and teaching cycles were set at 0.1 and 5 respectively. Cultural mutation was set at 0.05. These parameters were determined empirically to provide good performance.

4.1.1 Experimental Results

Figure 3 shows the average fitness of both populations throughout the experiment run, with error bars showing fitness variance within populations. It is clear from these results that the population employing cultural learning outperforms the population employing population learning alone from the start of the experiment. While the population learning population stabilises at around 0.825, the population employing cultural learning achieves fitness values of 0.9 and above.

Interestingly, it appears that the population employing population learning

alone is less diverse in its fitness variance than that of the population employing both cultural and population learning. Cultural learning appears to be producing individuals with fitness ranges larger than that of population learning alone. On average, even the worst individual in the cultural learning population is performing better than the best individual in the population learning population.

Figure 4 shows the average, maximum and minimum fitness values for the cultural learning population. Minimum and maximum values represent the average best and worst individual in the population. While both maximum and minimum values appear to be stable throughout the experiment run, the population's average fitness tends towards the maximum fitness value. By the second half of the experiment run, the population's average fitness is virtually indistinguishable from the population's maximum fitness value. Individuals that are incapable of improvement are quickly culled from the population, and the cultural learning mechanism is allowing even genetically mediocre individuals to achieve high fitness levels.

Table 1 shows values for the average, average maximum, average minimum and standard deviation for both populations. These figures show that the population employing cultural learning is capable of achieving higher averages, maximum, and minimum average values for the whole interval. There is strong evidence (p value < 0.0001 , 95% confidence) to suggest that the difference between the two populations is statistically significant.

In order to investigate further the effect of cultural learning on the popula-

tion, the population's fitness is measured before and after the teaching cycles begin. Thus, the fitness levels of the population are measured before and after teaching to determine an agent's innate fitness and its fitness acquired through cultural learning. Figure 5 shows three fitness values: one for the population employing population learning alone, one for the cultural learning population prior to teaching and the last showing the cultural learning population after teaching is applied.

The population employing cultural learning performs very differently before and after teaching is applied. Prior to teaching, the cultural learning population's fitness is considerably lower than that of the population learning population. Indeed, the population's genotypic fitness (the fitness measured before any cultural influence is applied) is consistently low and appears to be stable throughout the experiment.

4.2 The Game of BlackJack

Blackjack or twenty-one begins with the dealer dealing two cards face-up to each player and two to his/herself, with one card visible (the *up-card*) and the other face down. Cards are valued by their face value (10 for all picture cards) except for the ace which can be counted either as 11 or 1. The object of the game is to obtain a higher score (the sum of all card values) than that of the dealer's without exceeding 21. Each player can *draw* additional cards until they either *stand* or exceed 21 and go *bust*. Once all players have obtained their cards, the dealer turns over the hidden card and draws or stands as appropriate. Should

the dealer's hand bust, all players win.

The dealer is at considerable advantage because he/she only enters the game once all players have fully completed their play. Thus, it is probable that some players will have bust even before the dealer reveals the hidden card. In addition, the fact that only one of the dealer's cards is visible means that players must make judgements based on incomplete information. As a rule, the dealer follows a fixed strategy, typically standing on a score of 17 or more and drawing otherwise.

All aspects related to betting such as doubling down and splitting have been removed from this implementation in order to facilitate comparison with previous work which employs a similar approach.

In a casino setting, between 3 and 6 six full decks of cards are shuffled at the start of the first hand and the game is played until the cards run out. Up to six players and one dealer may play at a blackjack table. Again for simplicity, this implementation considers only a single player playing against the dealer using a single deck of cards which is shuffled at the start of each hand.

Each agent in the population contains a neural network that allows it to play blackjack. Once cards are dealt to the agents the value of the hand is shown to the network using thermometer encoding[48]. Each agent has a total of 19 input nodes, where 18 express the value of the cards held and the final node flags the presence of an Ace in the agent's hand. The 18 input nodes represent a scale of hand values of 4-21 (Figure 6).

The agent's decision is determined by rounding the output of the single

output node, where draw and stand are represented by 0 and 1 respectively. The number of hidden layers and nodes therein is unrestricted and is determined by the evolutionary process.

4.2.1 Bench-marking

In order to assess the performance of any evolved strategy, a set of benchmarks must be obtained for comparison purposes. While there have been many attempts to calculate the performance of blackjack strategies using simulation and probabilistic techniques[60, 2, 59], the values produced tend to vary by a rather large margin. For instance, the success of a player employing the standard dealer strategy is reported at between 39% and 44% wins. As a result of these discrepancies, it was felt that it may be more meaningful to calculate the values for various strategies using our own simulation. These values will be more readily comparable to the performance of evolved strategies, since a large proportion of the blackjack simulator will also be used by the evolving populations to play games.

The blackjack simulator consists of a dealer, who employs the traditional dealer strategy of standing on 17 or greater, and a single player whose strategy can be set at the beginning of the simulation. As in previous work, both dealer and player hand values are calculated by adding card values where each ace is counted as 11 unless it would cause a bust.

Several strategies were considered:

- Dealer's (Stand on 17 or more, Draw on less)

- Random
- Always stand
- Hoyle's (based on the dealer's up card and the possession of an ace)

```

if (dealer card < 6)
    if (ace is held)
        stand on 15
    else
        stand on 13
else
    stand on 17

```

- Uribe Evolved Strategy (taken from the work of Uribe and Sanchez[50])

```

if (score > 9) or
    [(score > 13) and (score < 19)
     and (an ACE is held)]
    stand
else
    hit with 50% probability

```

In order to produce statistically meaningful results, we performed 1000 runs of 1000 games for each strategy. The results presented in Table 2 are average wins for each strategy. We can see from these results that most strategies are quite poor against the dealer and that Hoyle's strategy performs best.

4.2.2 Experiments

Each experiment allows 100 agents to evolve over 400 generations. At each generation, agents play 100 games against a dealer strategy and an agent's fitness is determined by the percentage of wins obtained scaled to $[0.0,1.0]$. Crossover was set at 0.6 and mutation at 0.02. The cultural learning settings of teacher ratio and teaching cycles were set at 0.1 and 5 respectively. Cultural mutation was also added with probability 0.05.

An agent's fitness is determined by the number of hands won, normalised to the range $[0,1]$. This set of experiments does not allow agents to develop card-counting strategies and therefore the fitness values attained must be put into context with the probabilities of success for non-card-counting strategies. It is exceedingly difficult to beat a blackjack dealer consistently more than 50% of the time, and thus any result approaching this figure should be considered optimal.

Figure 7 shows the average fitness for both populations over the experiment run, with error bars showing fitness variance within populations. Both populations show similar trends, stabilising within 100 generations to their relative maximum values. The population employing both population and cultural learning is clearly achieving a higher average fitness than the population employing population learning alone.

Unlike the previous experiment, the fitness variances of the two populations are similar. This is most likely due to the low probability of any particular individual performing particularly well. However, on average, the worst indi-

viduals of the cultural learning population are performing better than the best individuals of the population learning population.

Figures 8 and 9 show the average, maximum and minimum fitness values for both populations. The minimum and maximum fitness values are those of the average best and worst individuals of each generation. Both minimum and maximum values are slightly higher for the population employing cultural learning. The cultural learning mechanism is not only improving the worst individuals in the population, but is also generating novel, high performing individuals.

Table 3 show the average, average maximum and average minimum fitness values for both population taken over the entire experiment run. It is clear from these results that the population employing cultural learning is superior in its development of strategies than the population employing population learning alone. There is strong evidence (p value < 0.0001 , confidence interval 95%) that the difference in fitness levels of the two populations is statistically significant.

In order to investigate the relative worth of the evolved blackjack strategy compared to the bench-marked strategies described above, the evolved strategy was extracted from the cultural learning population. This was done by presenting the population with every possible card combination and examining the collective decision of the population.

The following resulting strategy was extracted:

```
if (an Ace is held)
{
    if (dealer has a 6 or higher)
        stand on 16
    else
        stand on 17
}
else
{
    if (dealer has a 7 or higher)
        stand on 17
    else
        stand on 13
}
```

The strategy was hard-coded into the blackjack simulator and 1000 runs of 1000 games were played. The averaged results are displayed in Table 4.

There is strong evidence (p value <0.001 , 95% confidence) to support the claim that the evolved strategy and Hoyle's strategy are equivalent in terms of performance, suggesting that the population has evolved an optimum strategy given the information available. It is possible that in order to out-perform Hoyle's strategy it is necessary to keep track of cards that have been played during a game.

As a final investigation into the effect of cultural learning on the population, the cultural learning population’s fitness was examined before and after the teaching cycle is applied. The results for this are presented in Figure 10. As in the previous experiment, the application of cultural learning is clearly producing higher fitness levels than population learning alone. However, the cultural learning population’s fitness levels prior to teaching are slightly poorer than those of the population learning population.

4.3 Connect Four

The game of connect-four is a two-player game played on a vertical board of 7x6 positions into which pieces are slotted in one of seven available slots. Each player is given a number of coloured pieces (one colour per player) and must attempt to create horizontal, vertical or diagonal piece-lines of length four. Players place one piece per turn into one of the seven slots. The piece then falls onto a free position in the chosen column, creating piles, or towers, of pieces. If a column is full, the player must select an available slot.

Some research has been undertaken in the evolution of connect-four players employing a library of existing games to train the neural networks by back-propagation[55] as well as reinforcement learning methods[56]. Our approach allows agents to compete against each other and against a modified minimax player.

Agents play games against a minimax player, whose first move is randomized, allowing agents to play games of some variety. Fitness is assigned according to

the length of the game. In other words, agents are rewarded for bringing the game to as close as possible to a draw.

At each move, the current board position is taken and the agent's pieces are added iteratively into each slot. At each iteration, the network is shown the board position through its 42 input nodes. The neural network has only one output node and the board position with the strongest output response is deemed to be the agent's preferred board position.

4.3.1 Experiments

Populations of 50 agents are allowed to evolve over 400 generations. At each generation, agents play games against a minimax player, once playing as the first player and then as the second second player. Fitness is measured according to how close the game comes to a draw and is scaled to $[0,1]$. Crossover was set at 0.6 and mutation at 0.02. The cultural learning settings of teacher ratio and teaching cycles were set at 0.1 and 5 respectively. Cultural mutation was also added with probability 0.05.

The results illustrated in Figure 11 clearly show the effect of cultural learning on the population. The population employing population learning alone achieves fitness levels of around 0.6, indicating that the population is at least capable of adequately competing against the minimax player in more than half of games played. However, when cultural learning is applied to the population, the performance improvement is evident. The population achieves fitness levels of close to 0.8, significantly higher than population learning alone, indicating

that the population is capable of performing well against a minimax opponent.

Figures 12 and 13 show the average, maximum and minimum fitness values for the two populations. A number of clear distinctions can be observed from these results: firstly, average and maximum fitness values are higher for the population employing cultural learning while minimum values are not altered significantly. This implies that while the cultural learning process is producing high performing individuals, there are still elements in the population that are incapable of successfully interacting with their environment.

Secondly, the average fitness value is higher in the population employing cultural learning, as we have seen from Figure 11. However, the difference between the average and maximum fitness values is significantly reduced in the population employing cultural learning. Clearly, the cultural learning process is not only generating novel, high performing individuals, but it is also causing the entire population to more closely resemble those individuals that are best adapted to their environment.

Table 5 shows the average, average maximum and average minimum fitness values for both population taken over the entire experiment run. It is clear from these figures that cultural learning is producing individuals of average higher fitness, but is also capable of producing novel high performing individuals evidenced by the large differences between the maximum values of both populations. Furthermore, there is strong evidence (p value < 0.0001) that the performance differences between the two populations are statistically significant.

Figure 14 shows the fitness values for the population employing population

learning alone, the cultural learning population prior to teaching and the cultural learning population after teaching takes place. Once again, as in previous experiments, cultural learning appears to be selecting individuals for their genetic ability to learn, rather than for their innate ability to solve a particular task. This is illustrated by the fact that the fitness values for the population employing cultural learning are considerably lower prior to teaching than those of the population employing population learning alone.

Significantly, once teaching is applied to the cultural learning population, the fitness level rises and considerably exceeds that of the population employing cultural learning alone. Thus, the cultural learning process is generating individuals with a genetic predisposition toward learning. If teaching is not applied such individuals perform poorly, but as soon as teaching commences, the innate potential of such individuals is realised in full.

5 Discussion

The results obtained correlate with previous work, showing that the addition of cultural learning is capable of enhancing population fitness. The model of cultural transmission allows individuals to impart information that they themselves have acquired culturally, rather than innate knowledge, leading to some interesting results.

In particular, it is clear that culture is being passed on through generations, as the population's fitness continues to improve despite a significant deterio-

ration in innate fitness. The evolutionary process judges individuals on the basis of their performance after cultural information has been acquired and as a consequence, the genotypic behaviour of individuals becomes less and less important as the cultural exchanges become more successful. Individuals in a cultural learning setting only become competitive once they acquire the population's culture. The innate fitness of such individuals is considerably poorer than that of the population learning population, indicating that most of the knowledge required to survive in the environment is being stored in the culture, not in the genomes.

First generation teachers impart innate knowledge, as they have no teachers to imitate. From then on, pupils acquire knowledge that has itself been acquired by their teachers, cascading back to the first generation. However, the culture is constantly shaped by the influx and outflux of different teachers and therefore changes in character over time. Such information transmission is much faster than population learning, and allows the cultural learning population to achieve higher fitness levels, despite its genotypic deterioration.

A number of further insights can be derived from the results obtained. The original view of cultural learning was that any developments would be kept in check by genetic evolution, that the evolutionary process would favour only those memes that enhance genetic evolution. Thus, culturally acquired information, memes, would be 'kept on a leash' by the evolutionary process [65]. However, it is possible for the cultural process to turn the tables and begin affecting evolution, thus 'getting off the leash'.

Memes that are beneficial are passed on through the population and become more important as time goes on, leading to the degradation of genetic evolution. The only purpose of genetic evolution at this level is to ensure that individuals are capable of communicating effectively. Thus, their communicative apparatus should improve over time - this has parallels with human evolution, particularly with regard to brain size. It has been argued that brain size grew as a result of increased use of memes and that the advantages gained through the ability to share information throughout populations outweigh the considerable cost in terms of childbirth risk and calorific consumption[9].

Once the communicative apparatus has reached its peak however, genetic evolution becomes less important. This is especially true if individuals are capable of passing on knowledge that they themselves have acquired. By doing so, cultural artifacts undergo a process of evolution quite separate from the typical genetic one and are passed on from generation to generation.

If genetic evolution becomes less important over time as the cultural process takes over, it is tempting to ask if beneficial evolution could occur in the absence of population learning. In other words, can the cultural process alone achieve similar results?

On reflection, it seems unlikely that this could be the case because the genetic evolution is necessary firstly to tune the communicative apparatus so that successful cultural exchanges can take place between individuals and secondly because without genetic evolution, there would be a lack of diversity within the population which would severely restrict any further progress.

For instance, a population employing cultural learning alone would select the fittest teachers to teach others in the population. Since learning is based on imitation it is very unlikely that pupils would surpass the teachers and bring the population to higher levels of fitness. Therefore, we posit the best that could be hoped for in such populations is that all individuals would come to resemble the teacher, clearly increasing overall fitness, but not to the levels that would be observed when both population and cultural learning are employed.

The experiments presented in this paper allow cultural learning to take place without any associated costs. It is possible that results would be markedly different if learning had an associated cost, which is perhaps a more realistic scenario. If a learning cost were to be introduced it is likely that information acquired through the cultural process will be more readily assimilated into the genome via the Baldwin effect. Cultural learning relies on the existence of individuals willing to impart knowledge without any direct individual gain. This mechanism is unlikely to exist in the natural world without some form of cost. Therefore an important avenue for future research is to examine the effects of cultural learning in environments where a learning cost can be introduced.

6 Conclusion

This paper presents a model of cultural learning employing vertical transmission of culture in a population of neural network agents presented with a set of sequential decision task problems. Cultural learning gives populations the

opportunity to sample acquired information within the population itself. This allows weaker members of the population to gain access to environmental information which would otherwise be impossible to attain without incurring possible fitness losses.

Our model employs a population of evolving neural networks capable of communicating by assigning teacher agents to instruct members of the next generation. The model allows individuals to culturally impart knowledge they themselves have acquired from previous teachers rather than transmitting only innate knowledge. Furthermore, by allowing the evolutionary process to determine the structure of each agent's communicative apparatus, the model simulates cultural learning interactions that are considerably more expressive than previous work.

As in previous work, the results indicate that cultural learning provides improved performance over population learning in each test-case. However, the results obtained with regard to the population's innate fitness differ somewhat from those previously obtained. While the cultural learning population's fitness continues to improve over time, its innate fitness (the population's fitness prior to acquiring knowledge through teaching) deteriorates significantly. Thus, the majority of the population's knowledge about its environment is stored in the culture, rather than in its genomes.

Future work will focus on the cultural transmission of knowledge in dynamic environments, investigating whether the increased plasticity of acquired culture (as opposed to genetically acquired knowledge) leads to increased robustness.

Acknowledgements

We wish to thank the reviewers for their many constructive comments and suggestions. The first author would also like to acknowledge the support of the Irish Research Council for Science, Engineering and Technology.

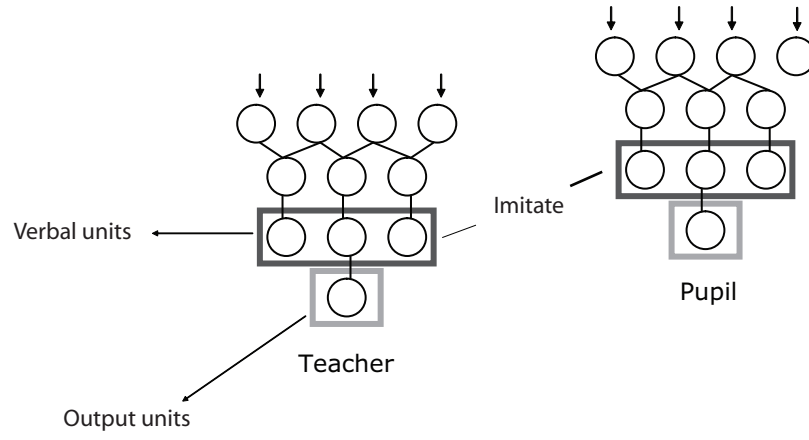


Figure 1: *Agent Communication Architecture*

	Start Marker	Node Label	Threshold	Link to Node	Link Weight	Link to Node	Link Weight	End Marker	
...	SM	5	0.8	4	0.83	3	-0.51	EM	...

Figure 2: *Marker Based Encoding*

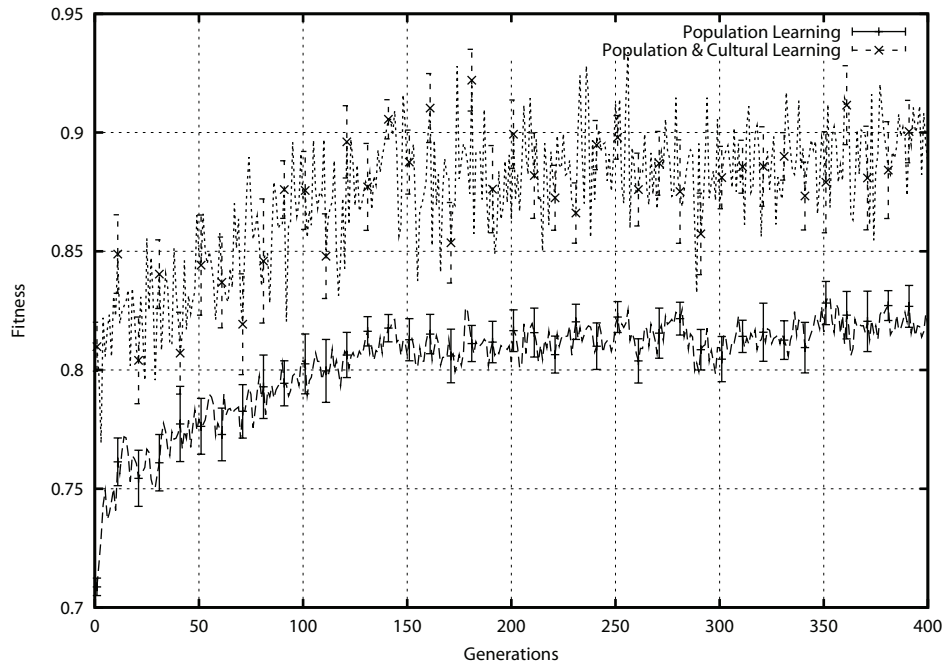


Figure 3: *Tic-Tac-Toe Population Fitness*

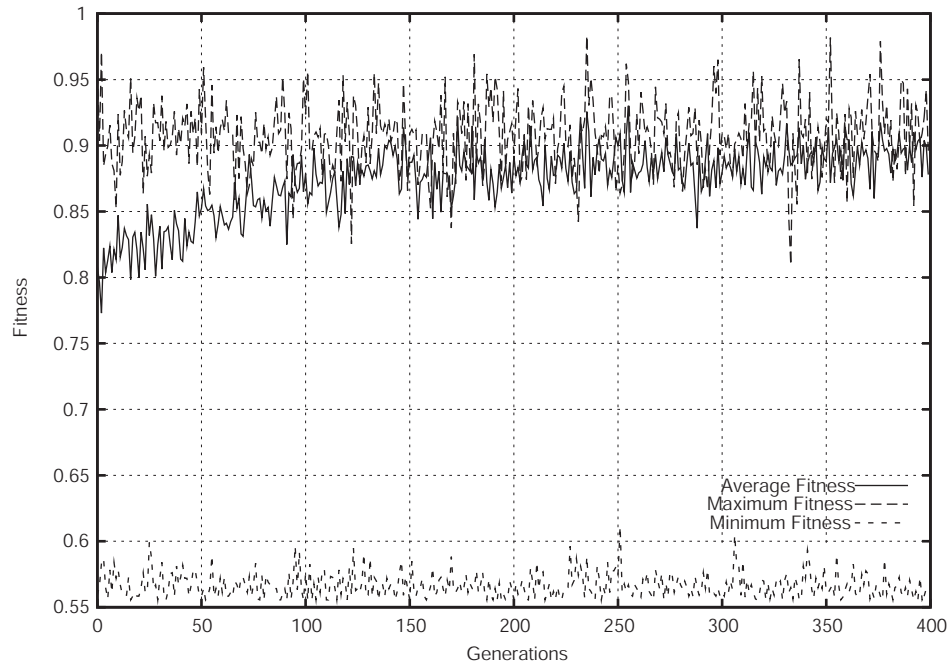


Figure 4: *Tic-Tac-Toe Average, Maximum and Minimum Fitness for Teaching Population*

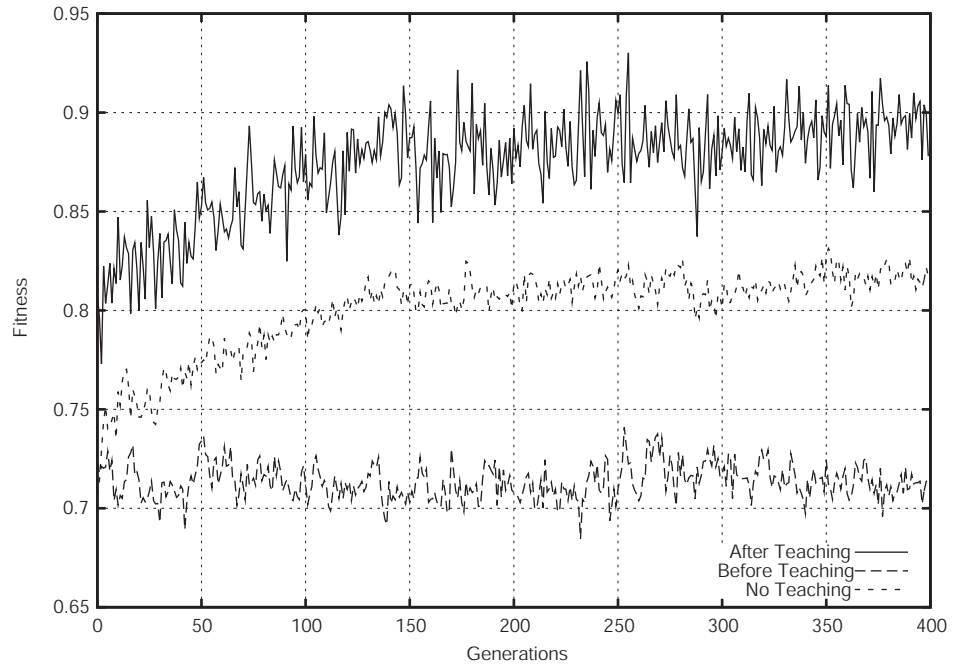


Figure 5: *Tic-Tac-Toe Average Fitness for Population Before and After Teaching*

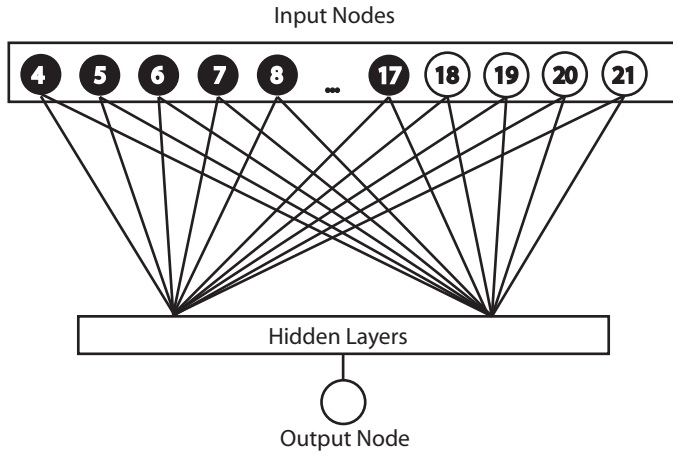


Figure 6: *Thermometer Encoding*

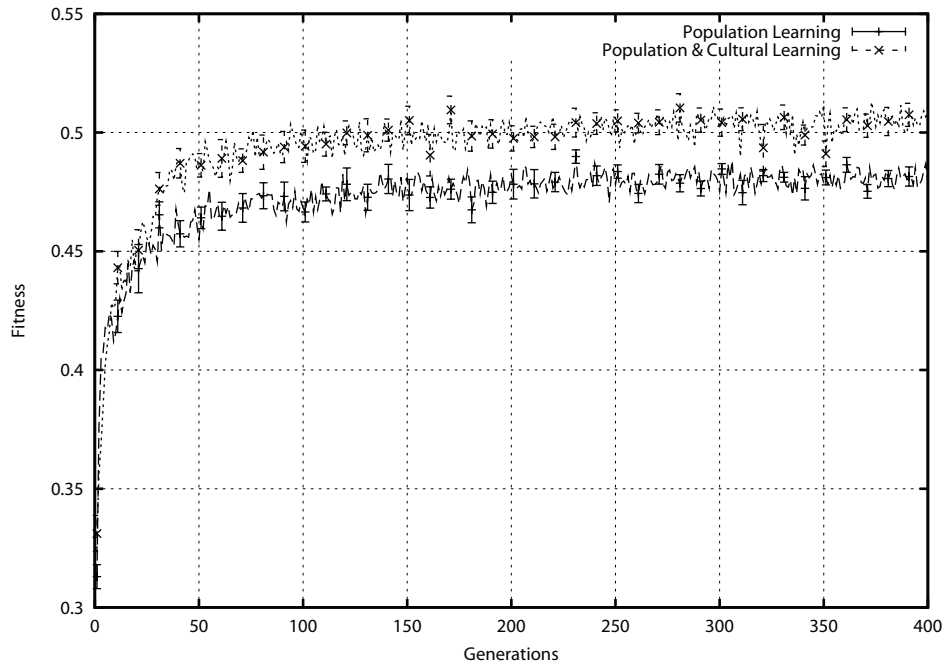


Figure 7: *Blackjack Population Fitness*

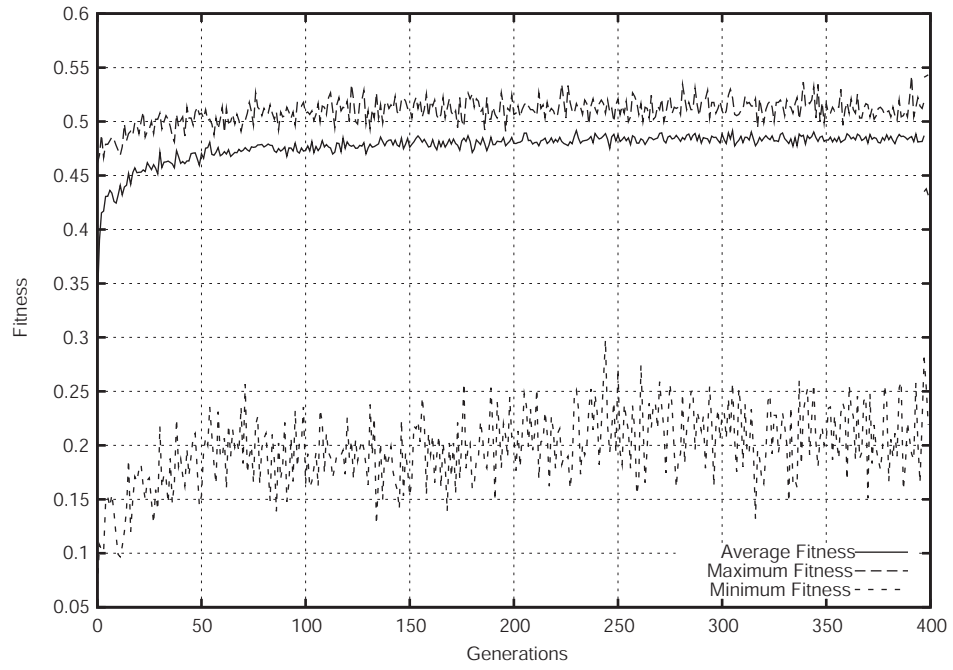


Figure 8: *Blackjack Average, Maximum and Minimum Fitness for Non-Teaching Population*

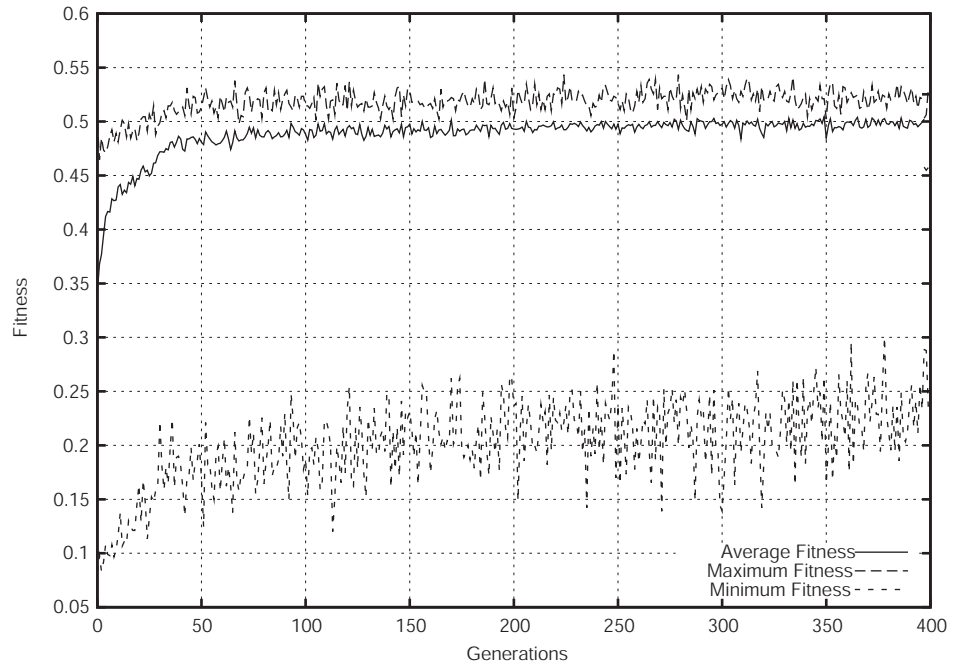


Figure 9: *Blackjack Average, Maximum and Minimum Fitness for Teaching Population*

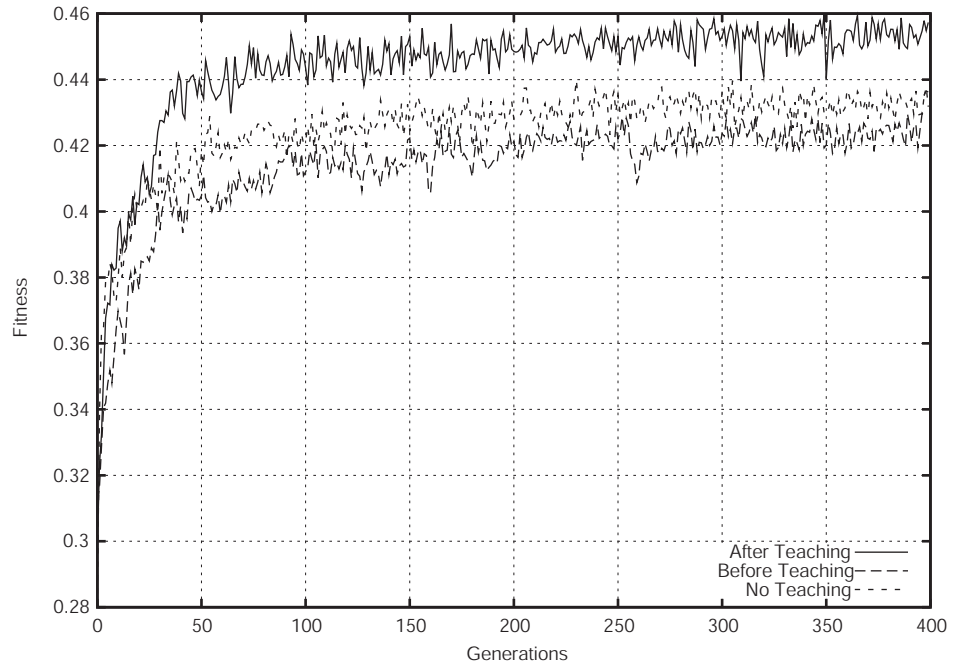


Figure 10: *Blackjack Average Fitness for Population Before and After Teaching*

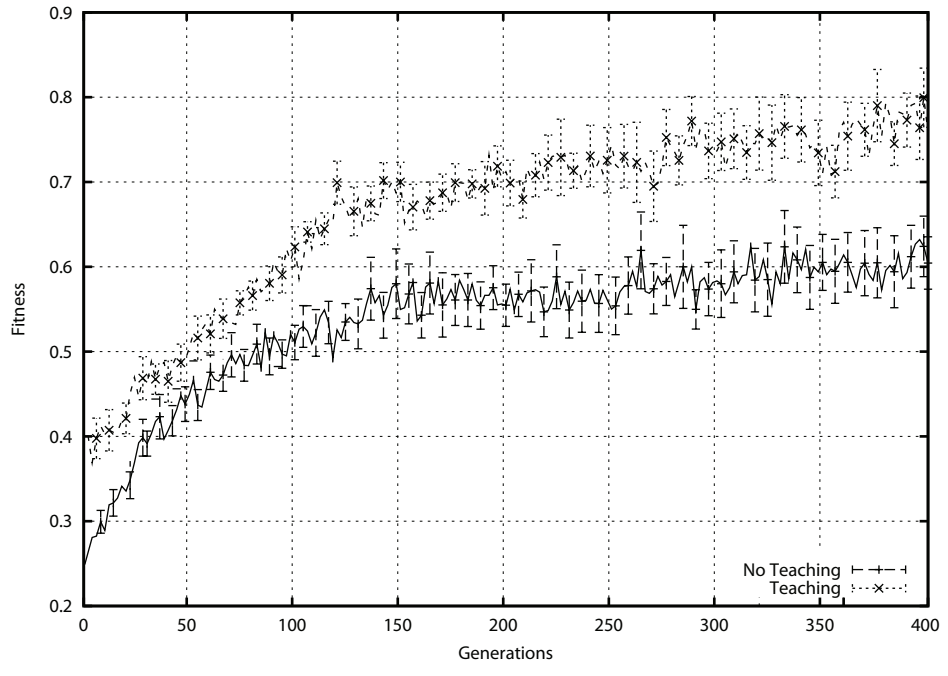


Figure 11: *Connect-Four Population Fitness*

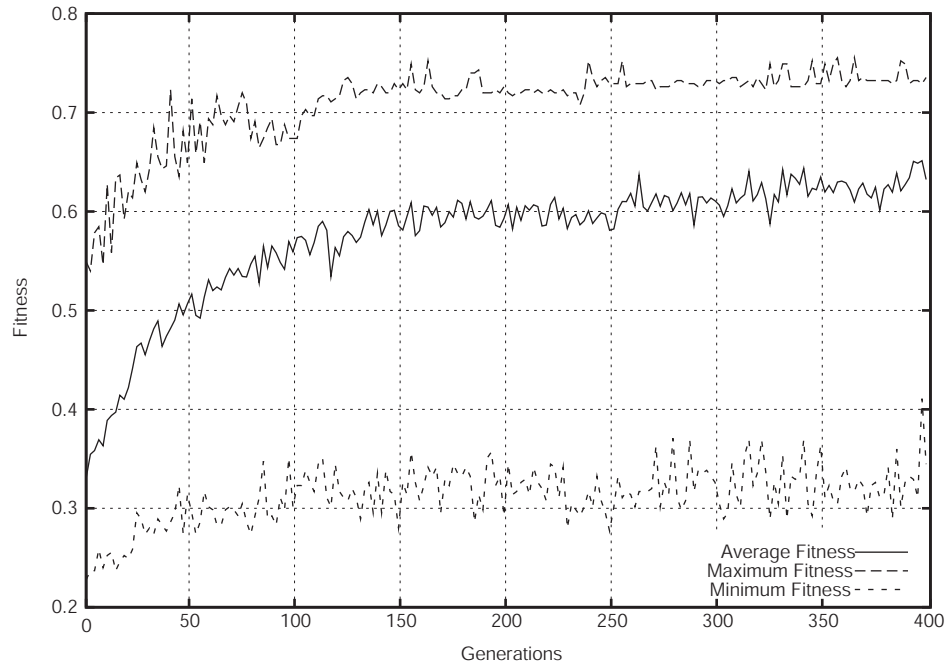


Figure 12: *Connect-Four Average, Maximum and Minimum Fitness for Non-Teaching Population*

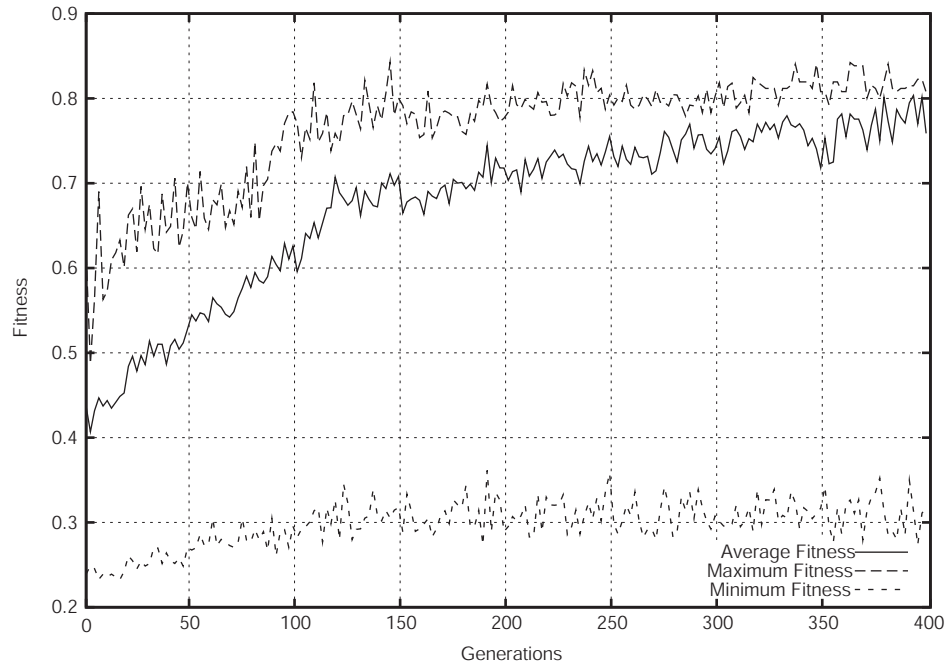


Figure 13: *Connect-Four Average, Maximum and Minimum Fitness for Teaching Population*

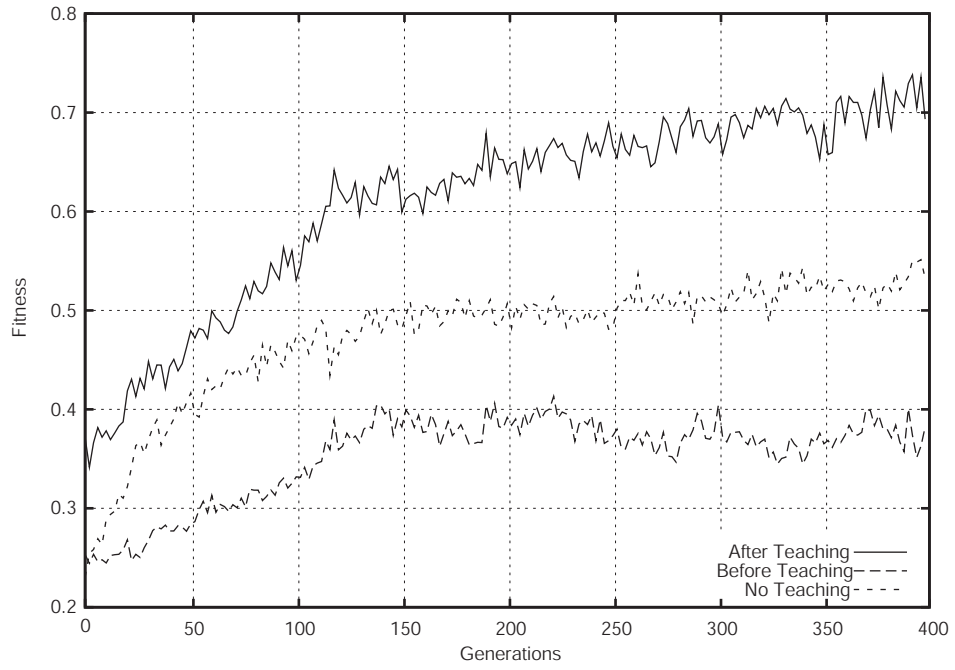


Figure 14: *Connect-Four Average Fitness for Population Before and After Teaching*

Population	Avg. Fitness	Max	Min	S. D.
Pop. Learning	0.80292	0.83087	0.70866	0.00040
Cultural Learning	0.87325	0.93377	0.76922	0.00076

Table 1: Tic-Tac-Toe Average Fitness

Strategy	Percentage Wins	Standard Deviation
Hoyle	43.70	1.587
Dealer	41.55	1.576
Uribe <i>et al.</i>	38.76	1.505
Always Stand	37.91	1.531
Random	30.41	1.511

Table 2: Blackjack Benchmarking

Population	Avg. Fitness	Max	Min	S. D.
Pop. Learning	0.47265	0.48989	0.31296	0.00026
Cultural Learning	0.49416	0.51282	0.33112	0.00043

Table 3: Blackjack Average Fitness

Strategy	% Wins	Standard Deviation
Hoyle	43.69	1.573
Evolved	43.67	1.582
Dealer	41.52	1.571
Uribe <i>et al.</i>	38.43	1.495
Always Stand	38.00	1.529
Random	30.67	1.507

Table 4: Final Blackjack Benchmarking Results

Population	Avg. Fitness	Max	Min	S. D.
Pop. Learning	0.53696	0.63200	0.25423	0.00591
Cultural Learning	0.66488	0.79935	0.36893	0.01092

Table 5: Connect-Four Average Fitness

References

- [1] Baldwin, J. (1896). A new factor in evolution. *American Naturalist*, 30, 441–451.
- [2] Baldwin, R. R., Cantey, W. E., Maisel, H., & McDermott, J. P. (1956). The optimum strategy in blackjack. In *Journal of the American Statistical Association*.
- [3] Batali, J. (1998). Computational simulations of the emergence of grammar. In S.-K. M. Hurford, J. R. & K. C. (Eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge: Cambridge University Press.
- [4] Belew, R. K. (1990). Evolution, learning and culture: Computational metaphors for adaptive algorithms. *Complex Systems*, 4, 11–49.
- [5] Belew, R. K., McInerney, J., & Schraudolph, N. N. (1992). Evolving networks: Using the genetic algorithm with connectionist learning. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II*, (pp. 511–547). Redwood City, CA: Addison-Wesley.
- [6] Best, M. L. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Adaptive Behavior*, 7(3/4), 289–306.

- [7] Billard, A., & Dautenhahn, K. (1999). Experiments in learning by imitation: grounding and use of communication in robotic agents. *Adaptive Behaviour*, 7(3/4), 411–434.
- [8] Billard, A., & Hayes, G. (1997). Learning to communicate through imitation in autonomous robots. In *7th International Conference on Artificial Neural Networks*, (pp. 763–738).
- [9] Blackmore, S. (2000). The power of memes. *Scientific American*, 284(4), 52–61.
- [10] Borenstein, E., & Ruppin, E. (2003). Enhancing autonomous agents evolution with learning by imitation. *Interdisciplinary Journal of Artificial Intelligence and the Simulation of Behaviour*, 1(4), 335–348.
- [11] Boyd, R., & Richerson, P. (1985). *Culture and the evolutionary process*. University of Chicago Press.
- [12] Brighton, H., & Kirby, S. (2001). The survival of the smallest: Stability conditions for the cultural evolution of compositional language. In J. Kelemen & P. Sosk (Eds.), *ECAL01*, (pp. 592–601). Springer-Verlag.
- [13] Bull, L., Holland, O., & Blackmore, S. (2000). On meme-gene coevolution. *Artificial Life*, 6(3), 227–235.
- [14] Cangelosi, A. (1999). Evolution of communication using combination of grounded symbols in populations of neural networks. In *Proceedings of*

- IJCNN99 International Joint Conference on Neural Networks (vol. 6)*, (pp. 4365–4368). Washington, DC: IEEE Press.
- [15] Cangelosi, A., & Parisi, D. (1998). The emergence of a language in an evolving population of neural networks. *Connection Science*, 10(2), 83–97.
- [16] Cavalli-Sforza, L. L. (1981). *Cultural Transmission and Evolution, A Quantitative Approach*. Princeton University Press.
- [17] Cavalli-Sforza, L. L. (1982). Theory and observation in cultural transmission. *Science*, 218, 19–27.
- [18] Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A quantitative approach*. New Jersey: Princeton University Press.
- [19] Chellapilla, K., & Fogel, D. B. (1999). Evolving neural networks to play checkers without relying on expert knowledge. *IEEE Transactions on Neural Networks*, 10, 1382–1391.
- [20] Curran, D., & O’Riordan, C. (2003). Artificial life simulation using marker based encoding. In *Proceedings of the 2003 International Conference on Artificial Intelligence (IC-AI 2003)*, volume II, (pp. 665–668). Las Vegas, Nevada, USA.
- [21] Curran, D., & O’Riordan, C. (2003). On the design of an artificial life simulator. In V. Palade, R. J. Howlett, & L. C. Jain (Eds.), *Proceedings of the Seventh International Conference on Knowledge-Based Intelligent*

- Information & Engineering Systems (KES 2003)*, (pp. 549–555). University of Oxford, United Kingdom.
- [22] Curran, D., & O’Riordan, C. (2004). The effect of noise on the performance of cultural evolution. In *Proceedings of the 2004 Congress on Evolutionary Computation (CEC 2004)*. Portland, Oregon, USA.
- [23] Dautenhahn, K., & Nehaniv, C. (2002). *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press.
- [24] De Boer, B., & Vogt, P. (1999). Emergence of speech sounds in changing population. In D. Floreano, J.-D. Nicoud, & F. Mondada (Eds.), *ECAL*, volume 1674 of *Lecture Notes in Computer Science*, (pp. 664–673). Springer.
- [25] De Jong, E. D. (1999). Analyzing the evolution of communication from a dynamical systems perspective. In *Proceedings of the European Conference on Artificial Life ECAL’99*, (pp. 689–693).
- [26] Demiris, J., & Hayes, G. (1996). Imitative learning mechanisms in robots and humans. In *Proceedings of the 5th European Workshop on Learning Robots*, (pp. 9–16).
- [27] Denaro, D., & Parisi, D. (1996). Cultural evolution in a population of neural networks. In *Proceedings of the 8th Italian Workshop on Neural Nets*, (pp. 100–111).

- [28] Denaro, D., & Parisi, D. (1996). Cultural evolution in a population of neural networks. In *M. Marinaro and R. Tagliaferri (eds), Neural Nets Wirn-96. New York: Springer*, (pp. 100–111).
- [29] Floreano, D., & Mondada, F. (1989). Evolution of plastic neurocontrollers for situated agents. *Animals to Animats*, 4.
- [30] Gruau, F. (1994). *Neural Network Synthesis using Cellular Encoding and the Genetic Algorithm*. Ph.D. thesis, Centre d’etude nucleaire de Grenoble, Ecole Normale Superieure de Lyon, France.
- [31] Hayes, G., & Demiris, J. (1994). A robotic controller using learning by imitation. In *Proceedings of the 2nd International Symposium on Intelligent Robotic Systems*.
- [32] Hinton, G. E., & Nowlan, S. J. (1987). How learning guides evolution. *Complex Systems*, 1, 495–502.
- [33] Hutchins, E., & Hazlehurst, B. (1991). Learning in the cultural process. In *Artificial Life II*, ed. C. Langton et al., (pp. 689–706). Cambridge, MA: MIT Press.
- [34] Hutchins, E., & Hazlehurst, B. (1995). How to invent a lexicon: The development of shared symbols in interaction. In N. Gilbert & R. Conte (Eds.), *Artificial Societies: The Computer Simulation of Social Life*, (pp. 157–189). London: UCL Press.

- [35] Kirby, S., & Hurford, J. (1997). Learning, culture and evolution in the origin of linguistic constraints. In P. Husbands & I. Harvey (Eds.), *ECAL97*, (pp. 493–502). MIT Press.
- [36] Kirby, S., & Hurford, J. (1997). Learning, culture and evolution in the origin of linguistic constraints. In *4th European Conference on Artificial Life*, (pp. 493–502). Cambridge, MA: MIT Press.
- [37] Kitano, H. (1990). Designing neural networks using genetic algorithm with graph generation system. *Complex Systems*, 4, 461–476.
- [38] Kolen, J. F., & Pollack, J. B. (1991). Back propagation is sensitive to initial conditions. *Advances in Neural Information Processing Systems*, 3, 860–867.
- [39] Koza, J. R., & Rice, J. P. (1991). Genetic generation of both the weights and architecture for a neural network. In *International Joint Conference on Neural Networks, IJCNN-91*, volume II, (pp. 397–404). Seattle, WA: IEEE Computer Society Press.
- [40] Laland, K. N., Kumm, J., & Feldman, M. W. (1995). Gene-culture coevolutionary theory: A test case. *Current Anthropology*, 36(1), 131–156.
- [41] MacLennan, B., & Burghardt, G. (1993). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior*, 2(2), 161–188.
- [42] Mandischer, M. (1993). Representation and evolution of neural networks. In R. F. Albrecht, C. R. Reeves, & N. C. Steele (Eds.), *Artificial Neural*

Nets and Genetic Algorithms Proceedings of the International Conference at Innsbruck, Austria, (pp. 643–649). Wien and New York: Springer.

- [43] Miller, G. F., Todd, P. M., & Hedge, S. U. (1989). Designing neural networks using genetic algorithms. In *Proceedings of the Third International Conference on Genetic Algorithms and Their Applications*, (pp. 379–384).
- [44] Moriarty, D. E., & Miikkulainen, R. (1995). Discovering complex othello strategies through evolutionary neural networks. *Connection Science*, 7(3–4), 195–209.
- [45] Nolfi, S., & Parisi, D. (1996). Learning to adapt to changing environments in evolving neural networks. *Adaptive Behavior*, 5(1), 75–97.
- [46] Nolfi, S., Parisi, D., & Elman, J. L. (1994). Learning and evolution in neural networks. *Adaptive Behavior*, 3(1), 5–28.
- [47] Oliphant, M., & Batali, J. (1997). Learning and the emergence of coordinated communication. *The newsletter of the Center for Research in Language*, 11(1).
- [48] Olson, D. K. (1993). *Learning to Play Games from Experience: An Application of Artificial Neural Networks and Temporal Difference Learning*. Master's thesis, Pacific Lutheran University.
- [49] Pereira, F. B., & Costa, E. (2001). How learning improves the performance of evolutionary agents: A case study with an information retrieval system for a distributed environment. In *Proceedings of the International Sym-*

posium on Adaptive Systems: Evolutionary Computation and Probabilistic Graphical Models (ISAS 2001), (pp. 19–23).

- [50] Perez-Uribe, A., & Sanchez, E. (1998). Blackjack as a test bed for learning strategies in neural networks. In *International Joint Conference on Neural Networks, IJCNN'98*, (pp. 2022–2027).
- [51] Pujol, J. C. F., & Poli, R. (1998). Efficient evolution of asymmetric recurrent neural networks using a two-dimensional representation. In *Proceedings of the First European Workshop on Genetic Programming (EUROGP)*, (pp. 130–141).
- [52] Richards, N., Moriarty, D., McQuesten, P., & Miikkulainen, R. (1997). Evolving neural networks to play Go. In *Proceedings of the 7th International Conference on Genetic Algorithms*. East Lansing, MI.
- [53] Sasaki, T., & Tokoro, M. (1997). Adaptation toward changing environments: Why darwinian in nature? In P. Husbands & I. Harvey (Eds.), *Fourth European Conference on Artificial Life*, (pp. 145–153). Cambridge, MA: MIT Press.
- [54] Saunders, G. M., & Pollack, J. B. (1994). *The Evolution of Communication in Adaptive Agents*. Technical report, Department of Computer and Information Science, The Ohio State University.
- [55] Schneider, M. O., & J, L. G. R. (2002). Neural connect four, aconnectionist approach to the game. In *Proceedings of the VIIth Brazilian Symposium on Neural Networks (SBRN'02)*, (pp. 236–241). IEEE Computer Press.

- [56] Sommerlund, P. (1996). *Artificial Neural Networks Applied to Strategic Games*. Unpublished Manuscript, URL: <http://www.diku.dk/student/peso/960513.ps.Z>. (Last accessed 20/05/2005).
- [57] Spector, L., & Luke, S. (1996). Culture enhances the evolvability of cognition. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, (pp. 672–677).
- [58] Sutton, R. S. (1986). Two problems with backpropagation and other steepest-descent learning procedures for networks. In *Proc. of 8th Annual Conf. of the Cognitive Science Society*, (pp. 823–831).
- [59] Thorp, E. O. (1963). *Beat the Dealer. A winning strategy for the game of twenty one*. New York: Vintage Books.
- [60] Thorp, E. O. (1984). *The Mathematics of Gambling*. Hollywood, CA: Lyle Stuart.
- [61] Watson, J., & Wiles, J. (2002). The rise and fall of learning: A neural network model of the genetic assimilation of acquired traits. In *Proceedings of the 2002 Congress on Evolutionary Computation (CEC 2002)*, (pp. 600–605).
- [62] Weaver, L., & Bossomaier, T. (1996). Evolution of neural networks to play the game of dots-and-boxes. In *Artificial Life V: Poster Presentations*, (pp. 43–50).

- [63] White, D. W. (1994). *GANNet: A genetic algorithm for searching topology and weight spaces in neural network design*. Ph.D. thesis, University of Maryland College Park.
- [64] Whitehead, H. (2005). Genetic diversity in the matrilineal whales: Models of cultural hitchhiking and group-specific non-heritable demographic variation. *Marine Mammal Science*, 21(1), 58–79.
- [65] Wilson, E. O. (1978). *On Human Nature*. Cambridge, MA: Harvard University Press.
- [66] Yanco, H., & Stein, L. (1993). An adaptive communication protocol for cooperating mobile robots. In *From Animals to Animats 2. Proceedings of the second International Conference on Simulation of Adaptive Behavior*, (pp. 478–485). Cambridge MA: MIT Press.