(In Proceedings of the 8th International Conference on the Simulation and Synthesis of Living Systems)

Stability and task complexity: A neural network model of evolution and learning

James R. Watson[1], **Nicholas Geard**[1] and **Janet Wiles**[1,2]

[1]School of Information Technology and Electrical Engineering
[2]School of Psychology
The University of Queensland, Australia

Abstract

Since Hinton and Nowlan introduced the Baldwin effect to the evolutionary computation community, agent-based studies of genetic assimilation have uncovered many details of the dynamic processes involved. In a previous paper, we demonstrated genetic assimilation with a simple food/toxin discrimination task using neural network agents that could evolve their learning rate. The study reported in this paper investigated the genetic assimilation of more complex learning tasks.

Kauffman's NK landscape model, which can generate landscapes with a variable degree of correlation, was used to define learning tasks of varying levels of complexity. Simulations indicate an increased tendency of genetic assimilation to occur as the complexity of the learning task decreases and the environmental stability increases. These results are explained in terms of the shifting balance between the evolutionary costs and benefits of learning.

Introduction

The interaction between evolution and learning has been an area of continual interest to the field of artificial life ever since Hinton and Nowlan's first computer simulation of the Baldwin effect (Hinton & Nowlan 1987). What initially appeared to be a relatively simple process has been found to be a complex phenomena with many interacting components.

The Baldwin effect was first described independently by Morgan (1896), Baldwin (1896) and Osborn (1896), and accounts for the tendency of learned behaviours to become genetically specified without resorting to a Lamarckian justification. Hinton and Nowlan's introduction of this biological concept into the evolutionary computation and artificial life communities marked the beginning of a growing body of research.

In Hinton and Nowlan's original simulation, the benefit of being able to learn was that it enabled an individual to solve a 'needle in a haystack' task within their lifetime. However, the advantage of learning was balanced by a fitness function that favoured individuals who were genetically closer to the solution. This cost of learning was sufficient to encourage the population to move, over time, towards a genetically specified solution.

The reason that genetic assimilation occurs at all, given the constraints of Darwinian selection, relies on the balance between the relative benefits and costs associated with the ability to learn (Mayley 1996). The Baldwin effect can be conceptualized in two distinct phases:

(i) initially, the ability to learn a task gives some subset of the population a selective advantage, resulting in subsequent generations becoming increasingly dominated by individuals with the ability to learn;

(ii) once the majority of individuals are able to learn, the costs of learning (e.g., resulting from increased intra-population competition between capable learners) cause selection to favour those individuals who, due to mutation and/or recombination, are more genetically predisposed towards the desired behaviour and therefore don't have as much to learn.

Genetic assimilation occurs when the balance between the benefits and costs of learning shifts so that selective pressures drive the ability to learn out of the population. Empirical demonstrations have shown that genetic assimilation is not always a straightforward process. Hinton and Nowlan's original simulation framework rarely, if ever, results in the complete genetic specification of the solution (see Harvey (1993) for a discussion). Other factors, such as the mutation rate (Fontanari & Meir 1990), selection algorithm (Wiles *et al.* 2001), the cost of learning (Mayley 1996), the amount of phenotypic plasticity (French & Messinger 1994) and population size have been shown to have an effect on the occurrence of genetic assimilation.

In a previous paper, Watson and Wiles (2002) demonstrated an equivalent result to that of Hinton and Nowlan (1987) using a population of evolving neural networks as agents. The task was to learn to discriminate between distinct sets of bit-string representations corresponding to food and toxin. The task was difficult enough that agents who were able to learn initially had a significant advantage. The learning rate was allowed to evolve, providing an indication of the level of phenotypic plasticity in the population. The practical effect of the learning rate is to amplify an agent's corrective action to a given situation, therefore the cost of learning arose

from the noisy responses that resulted from a high learning rate. The magnitude of this cost of learning could be altered by varying the stability of the environment, implemented as a change in the number of consecutive presentations of each representation.

The relationship between landscape complexity and the Baldwin effect has been investigated previously. Bull (1999) found that a high rate of learning was more beneficial as landscapes became less correlated. However, he compared the effect of using different fixed learning rates, rather than learning rates that were able to evolve over time. Bull's study therefore illustrated the benefits of learning in terms of finding a solution more rapidly, but not the relationship between the benefits and costs of learning that drive the occurrence of genetic assimilation.

What remains to be investigated is the relationship between the tendency for genetic assimilation to occur and the complexity of the learning task. To address this question, we extended Watson and Wiles methodology by using multi-layer feed-forward neural networks as agents, with a tunably complex food/toxin discrimination task based on the NK landscape model (Kauffman 1993) as the learning task.

The NK landscape model was originally developed to model the fitness landscapes that result from systems with varying degrees of epistatic interaction between their components. By altering the level of interaction (the parameter K), the resulting landscape can be varied from smooth and unimodal ($K = 0$) to rugged and highly multimodal ($K = N - 1$). In the simulations reported in this paper, the landscape was comprised of length N bit strings corresponding to food or toxin representations.

Each bit of a representation contributes to the total fitness in a manner dependent on its setting and on the settings of the other K bits to which it is linked. It is assigned a fitness table mapping each of the 2^{K+1} possible combinations to a random fitness value - in this case, a real number in the range [0, 1]. The fitness of the entire representation is given by the average of the N fitness contributions and also falls in the range [0, 1] (see Kauffman (1993) for further details). In this study the epistatic linkages were specified to be to the K nearest elements.

The complexity of the learning task was therefore derived from the correlation of the underlying NK landscape. If the landscape was highly correlated, the food and toxin representations would form easily separable subsets of the problem space (Figure 1a). As the landscape became less correlated, the partitioning of the problem space became more complicated (Figure 1b). A major advantage of this approach was that it resulted in a decoupling of the genotype space (defined by the agent's neural network weights – see next section) and the phenotype space (defined by the NK landscape).

Methodology

The simulation consisted of a population of agents, represented by fully-connected, multi-layer feed-forward neural networks. Each network had N input units (corresponding to the N elements of a representation), $N/2$ hidden units and a single output unit (whose output values were determined by a sigmoidal activation function). The (real-valued) connection weights between the input units and the hidden units, between the hidden units and the output unit, the unit biases, and the value of the learning rate, formed each agent's genotype.

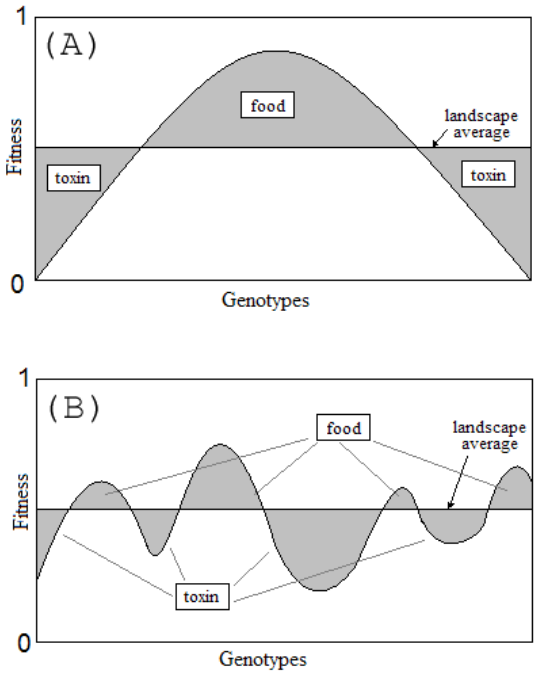

Figure 1: Problem spaces for (a) a simple ($K = 0$) learning task and (b) a more complex ($K > 0$) learning task.

Each agent lived for a fixed number of 'days'. For each day, an agent's task was to differentiate between two strings; one representing food and the other toxin (see below). The food/toxin representations were length N bit-strings chosen randomly from the set of 2^N possible representations of the pre-generated NK landscape. If an agent's response to an input string was higher than a fixed threshold, the agent was deemed to have 'eaten' the representation.

If a representation was eaten, the agent's score was adjusted according to the fitness of the corresponding point on the associated NK landscape, normalized by

the landscape average. That is, representations above the landscape average were considered to be food and made a positive contribution to an agent's score, while representations below the landscape average were toxins and made a negative contribution to an agent's score. The magnitude of the contribution was based on the difference between the value and the landscape average.

Thus, each day consisted of: (i) evaluating each agent's output given the current (length N bit-string) representation of food; (ii) updating the agent's score accordingly; (iii) using the backpropagation learning algorithm, updating agent weights to bring the next output closer to what was expected for the current representation (i.e. above or below the threshold); and then repeating (i), (ii) and (iii) for the current toxin representation.

The level of environmental stability (i.e., the length of time between changes to the current food and toxin representations) was varied according to parameter C. For example, if $C = 1$, a new food and toxin representation were chosen each day; if $C = 10$, the same representations were presented for 10 days before being changed. This allowed the benefits and costs of learning to be altered independently of the complexity of the learning task.

The simulations were run with $N = 12$ and a range of values for K and C. Each simulation was run for 5,000 generations, with each generation consisting of 100 neural-network agents (as described above) with a lifespan of 50 days. The score for each individual was given by the cumulative score over these 50 days. Tournament selection (see Mitchell (1996)) was used to create each successive generation. Neural network populations were asexual, and selected agents reproduced by passing on their inherited connection weights (*not* the learned weights) to the next generation. The network connections, the biases and learning rate were mutated by adding uniform random noise with a probability of 0.1.

Results and Discussion

After an initial period of low performance, the agents acquired the ability to discriminate between food and toxin via learning. Marked increases in the evolved learning rate corresponded to significant improvement in population performance (with the exception of maximally complex learning tasks with highly unstable landscapes).

Subsequent genetic assimilation (i.e. the learning rate falling to zero after the task was learned) occurred most frequently when the learning task was simple and the level of environmental stability was low. For learning tasks based on maximally correlated ($K = 0$) landscapes, complete genetic assimilation occurred when the environment was highly unstable ($C = 1$ and $C = 2$, see Figure 2). When the environment was more stable

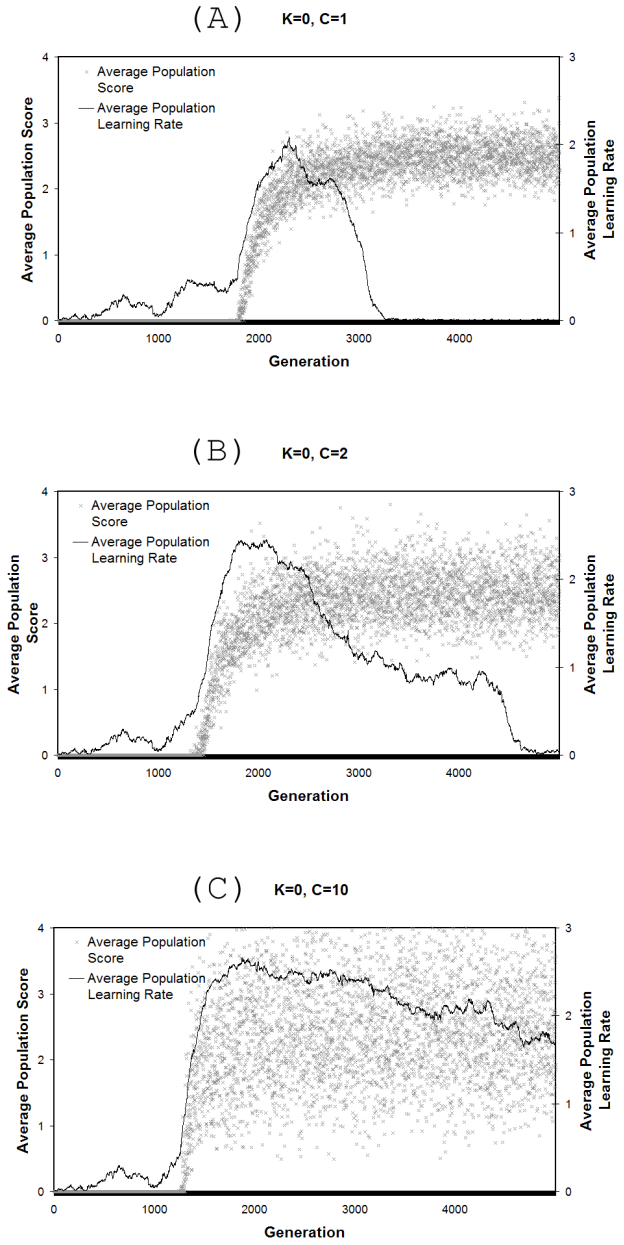

Figure 2: Performance and evolved learning rate over a simple landscape ($K = 0$) with increasing environmental stability. (a) maximally unstable conditions ($C = 1$): the rise and fall of the learning rate demonstrates the 2 stages of the Baldwin effect; initial acquisition via learning followed by genetic assimilation of the food/toxin discrimination task. Note that the average population score rises with the steep increase in learning rate and remains high even when the learning rate returns to 0. (b) and (c) display similar initial behaviour (acquiring the task through learning) as conditions became more stable. However, the rate of genetic assimilation was markedly slower for $C = 2$ and incomplete for $C = 10$.

($C = 10$), the learning rate began to fall, but was still quite high after 5,000 generations (more extensive simulations are required to investigate whether the learning rate does eventually fall to zero or if this residual learning remains).

These results can be explained in terms of the balance between the benefits and costs of learning. At a low level of environmental stability, each agent saw a constantly changing sequence of representations. In this situation, having an instinctive (i.e., genetically specified) response was more important than being able to learn the appropriate response during the agent's lifetime – the benefits of learning were low relative to the costs, and genetic assimilation occurred. As the level of environmental stability increased, the same representation was seen repeatedly during the agent's lifetime. The resulting increase in the benefits of learning, relative to the costs, was such that selective pressure was insufficient to drive the learning rate to zero. Consequently, genetic assimilation did not occur.

As the learning task became more complex (i.e., as K approached $N - 1$) and environmental stability was low, the task could not be achieved. As environmental stability was increased, learning became more beneficial and agents were able to discriminate between food and toxin via learning. The low costs of learning in this situation prevented genetic assimilation from occurring.

A further observed trend was an increase in the variance of the average population score as the representation change rate decreased (with increasing C). This trend was due to the fact that, as the change rate decreased but the agent's lifespan remained constant, a smaller sample of the total problem space was seen by each generation.

Conclusions

The acquisition and subsequent genetic assimilation of phenotypic traits occur when the relative costs and benefits of learning are balanced so that selective pressures first favour plasticity, then shift over evolutionary time to favour innate specification of what has been learned. The results from this study demonstrate how this balance is affected by the complexity of the learning task and level of environmental stability.

The primary benefit of learning, performing correctly in the absence of innately correct network weights, far outweighs the costs while the task is initially being acquired. Once the majority of a population can learn the task, the costs begin to outweigh the benefits and genetic assimilation can occur. The likelihood of such assimilation occurring has been found to depend on task complexity and environmental stability.

When task complexity is low and the representation change rate is high, learning is selected against and genetic assimilation occurs. As task complexity increases and the change rate decreases, the benefit of learning increases and genetic assimilation does not occur.

References

Baldwin, J. M. 1896. A new factor in evolution. *American Naturalist* 30:441–451 536–553. Reproduced in Belew, R. K. & Mitchell, M. (Eds.), *Adaptive Individuals in Evolving Populations*. Addison-Wesley, Reading, MA.

Bull, L. 1999. On the Baldwin effect. *Artificial Life* 5:241–246.

Fontanari, J. F., and Meir, R. 1990. The effect of learning on the evolution of asexual populations. *Complex Systems* 4:401–414.

French, R., and Messinger, A. 1994. Genes, phenes and the Baldwin effect: Learning and evolution in a simulated population. In Brooks, R. A., and Maes, P., eds., *Artificial Life IV*, 277–282.

Harvey, I. 1993. The puzzle of the persistent question marks: A case study of genetic drift. In Forrest, S., ed., *Proceedings of the Fifth International Conference on Genetic Algorithms*, 15–22. San Mateo, CA: Morgan Kaufmann.

Hinton, G., and Nowlan, S. 1987. How learning can guide evolution. *Complex Systems* 1:495–502.

Kauffman, S. A. 1993. *The Origins of Order - Organization and Selection in Evolution*. New York: Oxford University Press.

Mayley, G. 1996. The evolutionary cost of learning. In Maes, P.; Mataric, M. J.; Meyer, J.-A.; Pollack, J.; and Wilson, S. W., eds., *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, 458–467. MIT Press.

Mitchell, M. 1996. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.

Morgan, C. L. 1896. On modification and variation. *Science* 4:733–740.

Osborn, H. F. 1896. Ontogenic and phylogenic variation. *Science* 4:786–789.

Watson, J., and 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)*, 600–605.

Wiles, J.; Schulz, R.; Bolland, S.; Tonkes, B.; and Hallinan, J. 2001. Selection procedures for module discovery: Exploring evolutionary algorithms for cognitive science. In Moore, J. D., and Stenning, K., eds., *Proceedings of the 23rd Annual Conference of the Cognitive Science Society (CogSci 2001)*, 1124–1129. Lawrence Erlbaum Associates.