The Baldwin Effect Revisited: Three Steps Characterized by the Quantitative Evolution of Phenotypic Plasticity

Reiji Suzuki and Takaya Arita

Graduate School of Human Informatics, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
{reiji, ari}@info.human.nagoya-u.ac.jp
http://www2.create.human.nagoya-u.ac.jp/~{reiji, ari}/

Abstract. An interaction between evolution and learning called the Baldwin effect has been known for a century, but it is still poorly appreciated. This paper reports on a computational approach focusing on the quantitative evolution of phenotypic plasticity in complex environment so as to investigate its benefit and cost. For this purpose, we investigate the evolution of connection weights in a neural network under the assumption of epistatic interactions. Phenotypic plasticity is introduced into our model, in which whether each connection weight is plastic or not is genetically defined and connection weights with plasticity can be adjusted by learning. The simulation results have clearly shown that the evolutionary scenario consists of three steps characterized by transitions of the phenotypic plasticity and phenotypic variation, in contrast with the standard interpretation of the Baldwin effect that consists of two steps. We also conceptualize this evolutionary scenario by using a hill-climbing image of a population on a fitness landscape.

1 Introduction

The Baldwin effect[1] is known as one of the interactions between evolution and learning, which suggests that individual lifetime learning (phenotypic plasticity) can influence the course of evolution without the Lamarckian mechanism. This effect explains these interactions by paying attention to balances between benefit and cost of learning through the following two steps[2]. In the first step, lifetime learning gives individual agents chances to change their phenotypes. If the learned traits are useful for agents and make their fitness increase, they will spread in the next population. The learning acts as a benefit in this step. In the second step, if the environment is sufficiently stable, the evolutionary path finds innate traits that can replace learned traits, because of the cost of learning. This step is known as genetic assimilation[3]. Through these steps, learning can guide the genetic acquisition of learned traits without the Lamarckian mechanism in general. Figure 1 roughly shows the concept of the Baldwin effect which consists of two steps described above.

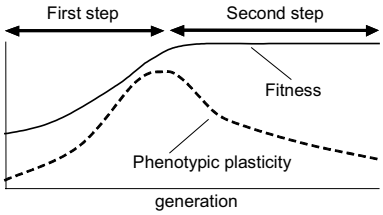

Fig. 1. Two steps of the Baldwin effect

Hinton and Nowlan constructed the first computational model of the Baldwin Effect and conducted an evolutionary simulation[4]. This had a large impact not only on biologists but also on computer scientists. The essential point of this study is that they introduced the quantitative evolution of phenotypic plasticity into their model, in other words, they allowed a population to adjust how much it depends on these two adaptive mechanisms through evolution. They clearly revealed the Baldwin effect by showing the transition of phenotypic plasticity as typically conceptualized in Figure 1, especially in the first step.

Since Hinton and Nowlan, many studies had discussed the interactions between evolution and learning based on various versions of Hinton and Nowlan's model. Harvey observed the existence of the second step in their model by conducting long term experiments and discussed the effect of genetic drift on this step[5]. Turney have discussed the Baldwin effect in view of bias shifting algorithms in machine learning by introducing bias strength (a probability which decides whether each phenotype is determined by learning or genetic information) into a simple model along the line of Hinton and Nowlan's[12]. Watson and Pollack extended their model so as to discuss the evolution of symbiotic interactions by replacing learning (random search) with lifetime interactions among organisms[6]. They demonstrated how symbiotic relationships can guide the genetic make-up of adaptive organisms. Arita and Suzuki discussed the Baldwin effect in dynamic environments by focusing on benefit and cost of learning caused by interactions among agents[7]. They adopted the evolution of strategies for iterated Prisoner's Dilemma, and then introduced a relatively simple learning rule termed as Meta-Pavlov learning into strategies as phenotypic plasticity which is quantitatively allowed to evolve. They found that the population evolved to be cooperative and stable through the two steps of the Baldwin effect.

There have also been many studies which have discussed interactions between evolution and learning with more sophisticated learning mechanism or in more complex environments. Especially, they often adopted the evolution of neural networks. Ackley and Littman's work is another pioneering example of such cases[8]. In their model, each individual seek for food avoiding carnivores in an artificial eco-system using the neural network which maps from the environmental input to the actual behavior. It also has another network which

evaluates the environmental conditions and reinforces the connection weights in the former network. They successfully showed that reinforcement learning and evolution of initial connection weights together were more successful than either alone in producing adaptive population that survived to the end of their simulation. Parisi and Nolfi discussed the effect of the correlation between the learning task and the evolutionary task by evolving a population of neural networks which forage distributed foods in the environment[9]. They pointed out that even if the learning task is not so correlated with the evolutionary task, it can facilitate the increase in the fitness of the evolutionary task. Sasaki and Tokoro also studied the relationship between learning and evolution using a model, where individuals learn to distinguish poison and food by modifying the connection weights in the neural network[10]. They compared the Darwinian and Lamarckian evolutionary mechanism in dynamic environments, and then concluded that the Darwinian evolution is more robust against environmental changes.

In the latter cases of studies, they assumed epistatic interactions among loci and enabled continuous change of them based on learning. However, their discussions were mainly based on two extreme cases, that is, experiments with or without learning process because the quantitative evolution of phenotypic plasticity was not introduced. Thus, the two steps of the Baldwin effect were not clearly discussed in comparison with Hinton and Nowlan's experiment. But it seems essential to clarify how the effect of the quantitative evolution of phenotypic plasticity on evolutionary scenario in such complex environments as discussed in the former cases of studies.

Our purpose is to give a valuable insight into the interaction between evolution and learning by focusing on the quantitative evolution of phenotypic plasticity in complex environments. We adopt the evolution of connection weights in a neural network as a situation where there are epistatic interactions among loci. We introduced the phenotypic plasticity into our model, in which whether each connection weight is plastic or not is genetically defined and plastic weights can be adjusted by back-propagation through learning processes.

The rest of the paper is organized as follows: Section 2 describes a model for investigating the interaction between evolution and learning by evolving connection weights and their plasticity in neural networks. In Section 3, based on the results of experiments, we show that the evolutionary scenario consists of three steps unlike the standard interpretation of the Baldwin effect. Furthermore, we investigate the transition of the benefit and cost of learning in each step by considering the effect of difference in evolutionary state of fitness. Section 4 summarizes the paper and conceptualizes this evolutionary scenario by using a hill-climbing image of a population on a fitness landscape.

2 Model

2.1 Genetic Description of Neural Network

We investigate the evolution of connection weights in a neural network as a situation where there are epistatic interactions among loci. There are P individ-

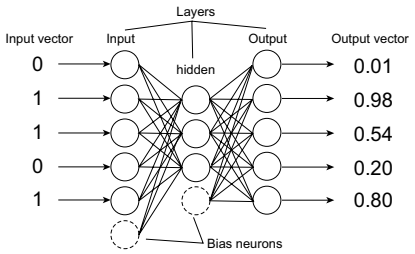

Fig. 2. The topology of neural network in case of $N = 5$, $M = 3$

uals in a population and each individual has a feed-forward multi-layer neural network which consists of $N + 1$ neurons in the input layer, $M + 1$ neurons in the hidden layer and N neurons in the output layer as shown in Figure 2. We represent the bias of each neuron by including extra "bias" neurons into input and hidden layers. Each neuron has a standard sigmoidal activation function as its mapping from a weighted sum of input values to the output value.

Each agent has a set of $(N + 1)M + (M + 1)N$ genes termed GW. Each gene in GW decides initial real value of each connection weight as its phenotype (phenotypic value). Here, we introduce the phenotypic plasticity into our model. Each agent also has another set of $(N + 1)M + (M + 1)N$ binary genes GP which decides whether each corresponding phenotype of GW is plastic ("1") or not ("0"). The connection weights whose corresponding bit in GP equals to "1" can be adjusted based on the learning process.

2.2 N-M-N Encoder Decoder Problem

We adopt a version of the N-M-N encoder-decoder problem so as to evaluate the fitness of each agent. This problem is well known as a benchmark problem of learning algorithms. The objective is to achieve a mapping of N input units to M hidden units (encoding) and a mapping of M hidden units to N output units (decoding), in which each N-length input vector consists of "0" or "1" and its desirable output vector is identical to the input vector itself. As we assume a complete encoder-decoder, all possible 2^N combinations of input and output vectors are evaluated in this model.

2.3 Learning

We use a batch-type back-propagation algorithm which includes the effect of plasticity of connection weights as a learning mechanism of each agent[11]. Each individual updates each connection weight $w^{(t)}$ at time t to learn a correct mapping described before for L times by using the following equation:

$$\Delta w^{(t+1)} = -\eta \cdot p \cdot \sum_{v \in V} \frac{\partial E_v(w)}{\partial w}\bigg|_{w=w^{(t)}} + \alpha \cdot \Delta w^{(t)}, \tag{1}$$

where $\Delta w^{(t+1)}$ represents the difference between $w^{(t+1)}$ and $w^{(t)}$, η is the learning rate, p denotes the phenotypic plasticity ("0" or "1") of w, V is a set of all possible 2^N input vectors, and α denotes the momentum parameter. $\partial E_v(w)/\partial w|_{w=w^{(t)}}$ is the local gradient function of the squared error with respect to w when we compared the input vector v (that is the same as the training vector) with the output vector using connection weights at time t. It can be derived mathematically from back-propagation. The essential point of this equation is that there is an additional value p in the first argument on the right side. If $p = 1$, the weight is updated based on a standard back-propagation mechanism, otherwise it is not updated at all. Thus, we can regard GP as the parameter which decides the learning rate of each connection weight respectively.

2.4 Evolution

After all individuals have finished learning processes, we calculate the fitness of each individual by using the mean square error of all pairs of input and output vectors based on the following equation:

$$fitness = 1.0 - \frac{1}{N \cdot 2^N} \sum_{v \in V} \sum_{i=0}^{N-1} (Out_{v,i} - In_{v,i})^2, \tag{2}$$

where $In_{v,i}$ denotes the input value into the ith neuron in the input layer and $Out_{v,i}$ denotes the output value of ith neuron in the output layer when we inputted the vector v. Note that we do not introduce the explicit cost of learning process itself into the fitness evaluation.

Finally, the population in the next generation is generated by a simple genetic operation as follows: First, the worst individual's sets of genes (GW and GP) are replaced by copies of the best individual's. Then, every gene of all individuals is mutated with a probability p_m, respectively. A mutation in GW adds a randomly generated value within the range $[-d, d]$ to the current weight and a mutation in GP flips the current binary value.

The GP in our model corresponds to a binary case of bias strength in Turney's model[12] as described before. But our model is clearly different from his in the following sense: 1) there are epistatic interactions among loci in our model, 2) we focus on the continuous changes of phenotypic value through learning process, 3) we assume no effects of noise on learning.

3 Experimental Results

3.1 Interaction between Evolution and Learning

We have conducted evolutionary experiments using the following parameters: $P = 20$, $N = 5$, $M = 3$, $L = 10$, $\eta = 0.2$, $\alpha = 0.5$, $p_m = 0.005$ and $d = 1.0$. The initial population was generated on condition that initial connection weights were randomly decided within the range $[-d, d]$ and the proportion of

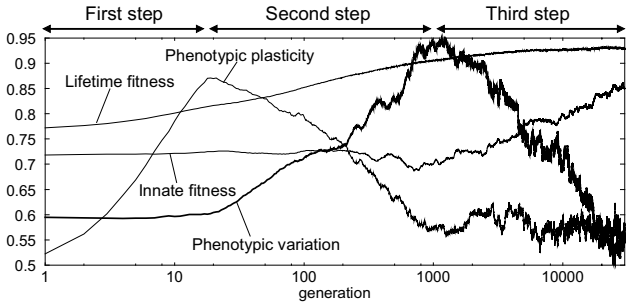

Fig. 3. Evolutionary dynamics of fitness and phenotypic plasticity

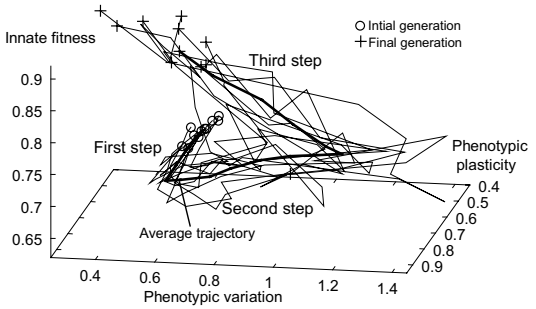

Fig. 4. The evolutionary trajectories of 10 trials

"1" in GP of each individual was uniformly distributed. Figure 3 shows the course of evolution over 30000 generations. The results are averages over 10 trials. The horizontal axis represents the generation in logarithmic scale. The *lifetime fitness* denotes the average actual fitness among all individuals that is calculated after learning process in each generation, and the *innate fitness* is the average potential fitness among all individuals calculated before learning process using initial (genetically defined) connection weights. The *phenotypic plasticity* represents the average proportion of "1" in all GP s and the *phenotypic variation* is the average absolute difference between the initial weight and the resultant weight adjusted by learning process among all plastic phenotypes. Figure 4 shows the evolutionary trajectories of these 10 trials drawn in the space of phenotypic plasticity, phenotypic variation and innate fitness.

By focusing on the transitions of these indices, we have found that the evolutionary scenario consists of three steps unlike the standard interpretation of the Baldwin effect as shown in Table 1, although there is some degree of variation in the actual transitions of these indices among trials as illustrated in Figure 4. Starting from the initial population, we observe an increase in both lifetime

Table 1. Transitions among three steps

step (generation)	1st (-20)	2nd (20-1000)	3rd (1000-)
lifetime fitness	increasing	increasing	slightly increasing
innate fitness	steady	steady	increasing
phenotypic plasticity	increasing	decreasing	steady
phenotypic variation	steady	increasing	decreasing
the standard interpretation	1st	1st and 2nd	2nd

fitness and phenotypic plasticity while the innate fitness remained steady, then the phenotypic plasticity rose to the peak value 0.87 at around 20th generation. This means that more plastic individuals could obtain higher fitness and could occupy the population because of the benefit of learning. Thus, we can regard that the population was in the first step of the Baldwin effect.

Subsequently, while the lifetime fitness still gradually increased, the phenotypic plasticity decreased to about 0.58 until around 1000th generation. This step has both properties of the first and second steps in the standard interpretation of the Baldwin effect in the following two different points of view. When we focus on the transition of the phenotypic plasticity, we can say that the population was in the second step in the sense that the increased fitness by learning was getting dependent of fewer plastic phenotypes. At the same time, in view of the transition of the innate fitness, it still remained flat and the phenotypic variation grew larger. Thus, the population was getting more strongly dependent of remaining plastic phenotypes, then we can also say that the population was still in the first step in this point of view.

Finally, the lifetime fitness still slightly increased and the phenotypic plasticity was approximately kept steady. The innate fitness eventually began to increase, however in contrast, the phenotypic variation decreased. Obviously, the genetic assimilation[3] was occurring in the remaining plastic phenotypes because these initial phenotypic values were getting closer to resultant phenotypic values after learning. Thus, the population entered into the second step in the standard interpretation of the Baldwin effect completely.

3.2 Evolutionary Benefit and Cost of Learning

Here we discuss the benefit and cost of learning in each step by considering the effects of difference in the phenotypic plasticity on fitness so as to clarify the meanings of the three steps. Figure 5 illustrates the effect of difference in evolutionary state on fitness in 10th, 500th and 30000th generations in a trial. For every individual in each generation, we generated copies of the individual whose relative differences in the number of plastic phenotypes between themselves and the original one were ± 1 or ± 2 by flipping the binary values at a randomly selected locus in their GP s for several times, and then evaluated their lifetime fitness. The horizontal axis represents the relative difference in the number of plastic phenotypes, and the vertical axis represents the fitness of individuals after

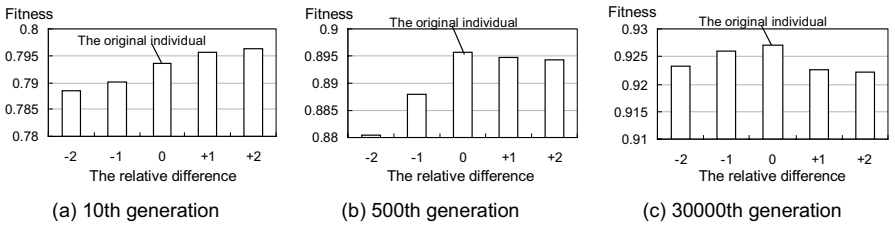

Fig. 5. The effect of difference in evolutionary state on fitness

learning process. For example, an individual whose relative difference equals to "+1" is produced by changing randomly selected a "0" to "1" in a copy's GP. The results are averages over 5 operations described above. We can grasp how the selective pressure was working on the evolution of phenotypic plasticity in each step by comparing these figures.

Figure 5(a) shows that the fitness always increased as the number of plastic phenotypes increased, thus learning worked as a benefit in the first step.

In the second step, we observe the peak of the fitness at the number of the plastic phenotypes of the original individual as shown in Figure 5(b). It is noteworthy that there is a negative effect on the fitness increase as the number of plastic phenotypes increases from original individual, although there is no explicit cost of learning as mentioned before. The reason is supposed to be the implicit cost of learning caused by epistatic interactions. In contrast with Hinton and Nowlan's model, a contribution of each phenotypic value to the individual's fitness strongly depends on the other phenotypic values because each value is one of connection weights in the neural network. Similarly, the learning of a value of plastic phenotype based on back-propagation also affects learning processes of the other values of plastic phenotypes, and such an affection caused by additional plastic phenotypes does not always yield an effective increase in the whole fitness. Thus there is the implicit cost of learning which corresponds to the *genetic costs* in DeWitt's classification of costs and limits of phenotypic plasticity[13]. However, such a detrimental effect could be reduced if more learning iterations were conducted, while we adopted small number of learning iterations $L = 10$ in this experiment. This limitation enlarges the difference of fitness caused by the effect because the learning process is required to achieve higher fitness in the limited number of iterations. In fact, the phenotypic plasticity decreased more gradually on condition that L was larger in our experiment.

Also, the decrease in the number of plastic phenotypes still yielded the decrease in the fitness because the basic benefit of learning obtained through the first step have to be maintained. Thus, we can say that the population evolved to have necessary and sufficient amount of phenotypic plasticity through the second step. In addition, we can also say that the decrease in the phenotypic plasticity was not merely due to the random drift in GP.

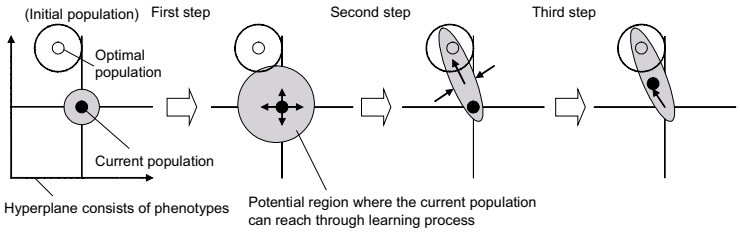

Fig. 6. Three steps of the evolutionary scenario illustrated by a hill-climbing image of a population on a fitness landscape

When the population was in the last step, the fitness tended to decrease more rapidly as the number of plastic phenotypes increased in comparison with the second step as shown in Figure 5(c). It means that the implicit cost of learning got larger in this step. Mayley pointed out that there should be a neighborhood correlation between genotype and phenotype space to guarantee a genetic assimilation to occur[14]. In this model, as the phenotypic plasticity decreases, resultant values of plastic phenotypes obtained by learning tend to correlate more with the genetically specified (initial) values because they are affected by other fewer phenotypes during learning process. Therefore, the decrease in the phenotypic plasticity with the consequent result of genetic assimilation in the third step supports his claim, and suggests that the evolution of phenotypic plasticity automatically enables the population to enter into the second step in the standard interpretation of the Baldwin effect. Besides, if there are no correlations among phenotypes, the evolutionary scenario is supposed to consist of the standard two step process of the Baldwin effect which is approximately similar to that of Hinton and Nowlan's experiment.

4 Conclusion

We have discussed how evolution and learning mutually interact by focusing on the quantitative evolution of phenotypic plasticity based on the evolution of connection weights in a neural network. We observed the Baldwin effect and have found that the evolutionary scenario consists of three steps.

Finally, we conceptualize this evolutionary scenario by using a hill-climbing image of a population on a fitness landscape, although there is another discussion on how to draw an actual landscape of population[15]. As shown in Figure 6, each two dimensional space represents a hyperplane which consists of possible phenotypic values (sets of connection weights). The population can be represented as a particular point in this plane and its height corresponds to the fitness of the population. Let us assume that there was a peak of fitness on the white filled circle and initial population existed on the black filled circle. The gray region represents the potential area where the current population can reach through

learning process. In the first step, the population expands its potential region toward every direction by increasing the phenotypic plasticity of population so as to search for the peak of fitness.

However, the implicit cost of learning limits the maximum area of its potential region. Then, the population transforms the shape of its potential region by whittling away excessive phenotypic plasticity and expands it toward the peak. These correspond to the decrease in the phenotypic plasticity and increase in the phenotypic variation in the second step. Finally, when the potential region reaches the peak of the fitness, the population begins to approach the peak at last, that is, the occurrance of the genetic assimilation in the third step.

References

1. Baldwin, J. M.: A New Factor in Evolution, *American Naturalist*, Vol. 30, pp. 441–451 (1896).
2. Turney, P., Whitley, D. and Anderson, R. W.: Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect, *Evolutionary Computation*, Vol. 4, No. 3, pp. 4–8 (1996).
3. Waddington, C. H.: Canalization of Development and the Inheritance of Acquired Characters, *Nature*, Vol. 150, pp. 563-565 (1942).
4. Hinton, G. E. and Nowlan, S. J.: How Learning Can Guide Evolution, *Complex Systems*, Vol. 1, pp. 495–502 (1987).
5. Harvey, I.: The Puzzle of the Persistent Question Marks: A Case Study of Genetic Drift, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp. 15–22 (1993).
6. Watson, R. A. and Pollack, J. B.: How Symbiosis Can Guide Evolution, *Fifth European Conference on Artificial Life*, pp. 29–38 (1999).
7. Arita, T. and Suzuki, R.: Interactions between Learning and Evolution: The Outstanding Strategy Generated by the Baldwin Effect, *Proceedings of Artificial Life VII*, pp. 196–205 (2000).
8. Ackley, D. and Littman, M.: Interaction between Learning and Evolution, *Proceedings of Artificial Life II*, pp. 487–509 (1991).
9. Parisi, D. and Nolfi, S.: The Influence of Learning on Evolution, *Adaptive Individuals in Evolving Populations*, pp. 419–428 (1996).
10. Sasaki, T. and Tokoro, M.: Evolving Learnable Neural Networks under Changing Environments with Various Rates of Inheritance of Acquired Characters: Comparison between Darwinian and Lamarckian Evolution, *Artificial Life*, Vol. 5, No. 3, pp. 203–223 (1999).
11. Rumelhart, D. E., Hinton, G. E. and Williams, R. J.: Learning Representations by Back-Propagating Errors, *Nature*, Vol. 323, pp. 533–536 (1986).
12. Turney, P.: How to Shift Bias: Lessons from the Baldwin Effect, *Evolutionary Computation*, Vol. 4, No. 3, pp. 271–295 (1996).
13. DeWitt, T. J., Sih, A. and Wilson, D. S.: Costs and Limits of Phenotypic Plasticity, *Trends in Ecology and Evolution*, Vol. 13, pp. 77–81 (1998).
14. Mayley, G.: Landscapes, Learning Costs and Genetic Assimilation, *Evolutionary Computation*, Vol. 4, No. 3, pp. 213–234 (1996).
15. Provine, W. B.: *Sewall Wright and Evolutionary Biology*, The University of Chicago Press (1986).