EMERGENT DYNAMICS OF BENEFIT AND COST OF LEARNING IN THE COURSE OF EVOLUTION

Reiji Suzuki and Takaya Arita

Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
{reiji, arita}@nagoya-u.jp

ABSTRACT

An interaction between evolution and learning called the Baldwin effect is known as the two-step evolution caused by the balances between benefit and cost of learning. However, little is still known about dynamic properties of these balances in complex environments. Our purpose is to clarify the key factors and conditions which bring about the dynamic evolution of the benefit and cost of learning by focusing on the quantitative evolution of phenotypic plasticity under the assumption of epistatic interactions. We have constructed an evolutionary model of quantitative traits by using an extended version of NK fitness landscape, in which plasticity of each phenotype is genetically defined and plastic phenotypes can be adjusted by learning iterations. The simulation results have shown that the dynamic emergence of the benefit and cost of learning brought about the three-step evolution of the Baldwin effect (observed in our previous study [8]) and it depends on the balances between the degree of epistasis and the number of learning iterations.
Keywords: Baldwin effect, evolution and learning, epistasis, phenotypic plasticity, genetic robustness, artificial life.

1. INTRODUCTION

Effects of nongenetic factors on genetic evolution are now drawing unprecedented attention due to the progresses in the molecular and developmental biology. An interaction between evolution and learning (phenotypic plasticity) have become one of controversial issues in this field since the emergent dynamics of interactions between evolution and learning called the Baldwin effect was clearly demonstrated by using a simple computational experiment based on a constructive approach [4].

The Baldwin effect is known as one of the evolutionary scenario which suggests that individual lifetime learning can influence the course of evolution without the Lamarckian mechanism [1]. This effect explains these interactions by paying attention to balances between benefit and cost of learning through the following two steps [11]. In the first step, lifetime learning (or phenotypic plasticity in a broader sense [3]) 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. 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 [12]. Through these steps, learning can guide the genetic acquisition of learned traits without the Lamarckian mechanism in general.

Recently, there are many studies on the Baldwin effect which stem from Hinton and Nowlan's extremely simplified version of a network connection model [4]. In their model, 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. They revealed the existence of the Baldwin effect by showing the increase and subsequent decrease in the phenotypic plasticity. However, the learning mechanism in their model was too simple on the ground that its benefit was approximately proportional to the number of plastic phenotypes and the cost of learning was explicitly introduced. Thus, there are many discussions on what kind of the benefit and cost of learning exist and cause this effect in realistic contexts such as the evolution of the strategies for iterated Prisoner's Dilemma [10].

The benefit and cost of learning caused by epistasis are interesting topics in this field because epistatic interactions are ubiquitous in modern genetics and evolutionary biology. For instance, Mayley conducted an evolutionary experiment using the Kauffman's NK fitness landscape [7]. He pointed out that there should be a neighborhood correlation between genotype and phenotype space to guarantee a genetic assimilation to occur. Bull also discussed the evolution on NK landscape by comparing different genetic inheritance mechanisms [2]. However, previous studies assumed the condition that all phenotypes were plastic and the quantitative evolution of phenotypic plasticity was not introduced into their models. In this sense, the two steps of the Baldwin effect were not clearly discussed in these models in compar-

Table 1: Three-step evolution of the Baldwin effect characterized by the four indices.

step	lifetime fitness	innate fitness	phenotypic plasticity	phenotypic variation	the standard interpretation
1st	increasing	steady	increasing	steady	1st
2nd	increasing	steady	decreasing	increasing	1st and 2nd
3rd	slightly increasing	increasing	steady	decreasing	2nd

ison with Hinton and Nowlan's case.

We have previously investigated the quantitative evolution of phenotypic plasticity in a neural network as a realistic situation [8]. We observed that the evolutionary scenario consists of three steps unlike the standard interpretation of the Baldwin effect. However, the adopted model was insufficient to clearly show the general effects of epistasis and the cause of the three-step evolution, because it was based on a specific situation and was difficult to conduct the experiments with or without epistasis. Thus, the analyses based on more general model are necessary to understand the complex effects of epistasis on the Baldwin effect. Our purpose is to clarify key factors and conditions which bring about such complex evolutionary scenarios by comparing the experimental results in various cases of parameters which are supposed to affect the benefit and cost of learning. Especially, we focus on the effects of degree of epistatic interactions and the number of learning iterations on the quantitative evolution of phenotypic plasticity.

For this purpose, we construct an evolutionary model based on Kauffman's NK fitness landscape [5] in which we can explicitly adjust the degree of epistasis. We adopt the evolution of quantitative traits which are similar to the connection weights in our previous model by extending the fitness evaluation of the NK model. We introduce the phenotypic plasticity into our model, in which the plasticity of each phenotype is genetically defined and the plastic phenotype can be adjusted by iterations of simple learning process. By conducting experiments with various degrees of epistasis and the different numbers of learning iterations, we show that the occurrence of the three-step evolution of the Baldwin effect depends on the balances between these parameters. Finally, we summarize the emergent dynamics of the benefit and cost of learning in this scenario.

2. THREE-STEP EVOLUTION OF THE BALDWIN EFFECT

We investigated the evolution of connection weights in a neural network as a situation where there are epistatic interactions among loci [8]. It was observed that the evolutionary scenario consists of three steps by focusing on the transitions of four indices as shown in Table 1. The *lifetime fitness* represents the fitness after learning in the population and the *innate fitness* is the potential fitness before learning based on initial phenotypes. The *phenotypic plasticity* represents the proportion of plastic phenotypes in the population. The *phenotypic variation* is the absolute difference in phenotypic values between before and after learning among plastic phenotypes. Note that it does not mean the variance of phenotypic values in the population in our study.

The first step, that is the increase in both lifetime fitness and phenotypic plasticity, was simply caused by the benefit of learning. It is important that the second step has both properties of the first and second step in the standard interpretation of the Baldwin effect. The decrease in the phenotypic plasticity corresponds to the second step in the standard interpretation of the Baldwin effect in the sense the increased fitness by learning becomes dependent on fewer plastic phenotypes. At the same time, the increase in phenotypic variation means that the population becomes strongly dependent on the remaining plastic phenotypes. Thus, we can also say that the population was still in the first step in this point of view. This phenomenon is supposed to be due to the implicit cost of learning caused by the epistatic interactions among plastic phenotypes through the learning processes. The third step corresponds to the second step in the standard interpretation because the genetic assimilation occurred in the remaining plastic phenotypes (see [8] for detailed discussions).

3. MODEL

3.1. NK Landscape with Real Valued Traits

We have constructed an evolutionary model based on Kauffman's NK fitness landscape [5], so as to discuss the evolution of phenotypic plasticity in quantitative traits with / without epistatic interactions among loci. There are P individuals in a population and each individual has N traits of which initial phenotypes are determined by genes in a N-length chromosome GI. Each gene represents the quantitative trait t_i (i=0, $\cdots$, N-1) which consists of a real value within the range $[0.0, 1.0]$. We adopt NK fitness landscapes for evaluation of fitness because we can explicitly adjust the degree of epistasis by using the parameter K. It represents the number of other traits that affect the fitness contribution of each trait. However, the standard NK fitness landscape only assumes the binary traits ("0" or "1"). Then we extended the definition of the fitness evaluation so as to deal

with the fitness contributions of quantitative traits.

Each trait t_i has epistatic interactions among other K traits $t_{i+jmodN}$ (j=1, $\cdots$, K). For each t_i, we prepare a lookup table which defines its fitness corresponding to all possible (2^{K+1}) combinations of interacting traits when these phenotypes consist of only binary values ("0" or "1"). The value of each fitness in the lookup table is randomly set within the range $[0.0, 1.0]$. These tables are similar to those of the standard NK landscape.

The fitness for quantitative trait is defined as the linearly interpolated value among the fitness for binary combinations of interacting phenotypes using the following equation:

$$f(t_i) = \sum_{c \in C_i} [f_{i,c} \cdot \prod_{j=i}^{i+K} \{(1.0 - b_{jmodN}) \cdot (1.0 - t_{jmodN}) + b_{jmodN} \cdot t_{jmodN}\}], \quad (1)$$

where $f(t_i)$ is the fitness of the trait t_i, C_i denotes the all possible 2^{K+1} combinations of binary traits, $f_{i,c}$ is the fitness of t_i when the combination of traits is c. b_j represents the j th binary phenotype in c. Figure 1 shows an example of the interpolation of the fitness of quantitative traits for N=2 and K=1. The table on the right side represents the lookup table which determines the fitness of t_0 corresponding to four binary combinations of t_0 and t_1. The left figure shows the interpolated fitness of t_0 generated by the right table and the equation (1). The individual fitness is regarded as the average fitness over all traits. Note that if we assume only binary phenotypic values, this model is equivalent to the standard NK fitness landscape.

3.2. Learning

Each agent has another N-length chromosome GP which decides whether the corresponding phenotype of GI is plastic ("1") or not ("0"). Each trait whose corresponding bit in GP equals to "1" is adjusted by iterating the following procedure L times. First, for each plastic trait t_i, we calculate the difference in t_i between time t and t+1 (Δt_i) using the following equation:

$$\Delta t_i = \begin{cases} -\beta(F_0 - F_c) & \text{if } max(F_0, F_c, F_1) = F_0, \\ \beta(F_1 - F_c) & \text{if } max(F_0, F_c, F_1) = F_1, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where F_c represents the individual fitness of the current combinations of traits and F_0 is the individual fitness when t_i is set to 0, F_1 is the individual fitness when t_i is set to 1. Next, we actually adjust all values of the plastic traits by adding Δt_i at the same time. This process means that the individual gradually adjust its own plastic phenotypes toward fitter extreme phenotypic value ("0.0" or "1.0") in proportion to the increase in the fitness. The gray arrows in Figure 1 show examples of Δt_0 and Δt_1. The black arrow corresponds to the resultant direction and distance of learning process.

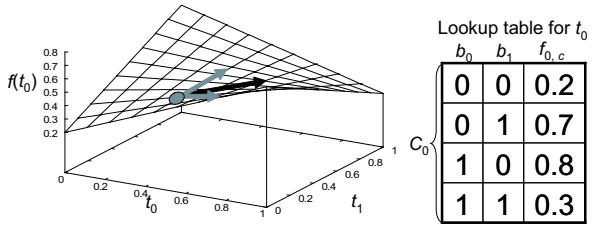

Figure 1: The example of the interpolated fitness of the trait t_0 for N=2 and K=1.

3.3. Evolution

After all individuals have finished their learning iterations, the population in the next generation is generated by a simple genetic operation as follows: First, the worst individual's chromosomes (GI and GP) are replaced by copies of the best individual's. Then, every gene for all individuals is mutated with a probability p_m. A mutation in GI adds a randomly generated value within the range $[-d, d]$ to the current value and a mutation in GP flips the current binary value. If a mutated phenotypic value in GI exceeds the domain of the phenotypic space, another mutation is operated on the original value again. We adopted these procedures so as to observe the gradual transitions of four indices explained before.

4. EFFECTS OF THE DEGREE OF EPISTASIS (K)

4.1. Experiments without Epistasis (K=0)

We have conducted the experiments using the following parameters: P=20, N=15, L=5, β=10.0, p_m=0.003, d=0.03 and K=0 or 4. The initial population was generated on condition that initial values in GI were taken at random within the range $[0, 1]$ and the proportion of "1" in GP for each individual was uniformly distributed within the range $[0, 1]$.

First, we have conducted the experiments without epistasis. Figure 2 shows the course of evolution over 20000 generations with K=0. The results are averages over 10 trials. The horizontal axis represents the generation in logarithmic scale. The lines represent the four indices that we have explained above. Specifically, the *lifetime fitness* denotes the average fitness among all individuals calculated after the learning process, and the *innate fitness* is the average potential fitness calculated before the learning process

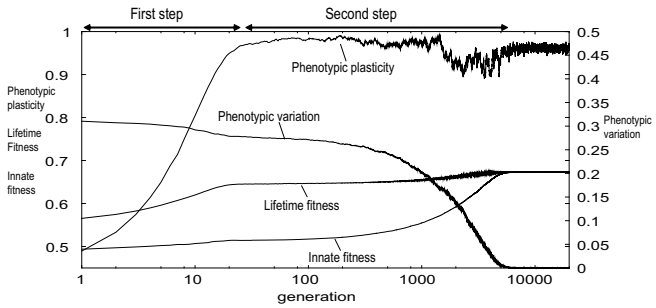

Figure 2: Evolutionary dynamics of fitness and phenotypic plasticity for K=0.

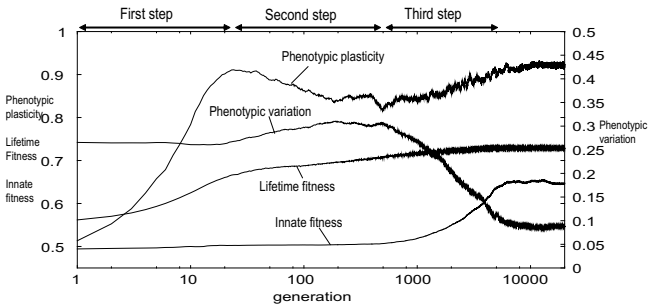

Figure 3: Evolutionary dynamics of fitness and phenotypic plasticity for K=4.

using initial phenotypic values. 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 value and the resultant value adjusted by the learning process among all plastic phenotypes.

As shown in the transitions of these indices, the evolutionary process basically consists of the standard two-step evolution of the Baldwin effect. From the initial population, we observe an increase in both lifetime fitness and phenotypic plasticity while the innate fitness remained steady. The phenotypic plasticity rapidly rose and exceeded 0.97 at around the 28th generation. This means that more plastic individuals could obtain higher fitness and could occupy the population due to the benefit of learning.

Subsequently, the innate fitness slowly increased and the phenotypic variation gradually decreased until around the 6000th generation. We can regard that the genetic assimilation occurred on the learned traits because the initial phenotypic values were getting closer to resultant phenotypic values after learning. The main reason for this phenomenon is due to the limitation in the learning iterations (L). Besides, as the innate fitness got closer to the lifetime fitness, the phenotypic plasticity was slightly decreased until around the 2300th generation. It is due to the genetic drift because there is no explicit cost of learning when K=0.

In contrast with the evolutionary scenario in Table 1, the phenotypic plasticity increased again and kept high even after the genetic assimilation had completely finished, despite the fact that the learning did not increased the fitness of the population at all. Thus, another role of learning must occur after the Baldwin effect. We focused on the implicit benefit of learning, that is, the genetic robustness against mutations which was approximated by measuring the average decrease in the fitness of individuals caused by mutational perturbations on GI. The results have shown that the genetic roustness was kept high by regaining plastic phenotypes after the Baldwin effect (see [9] for detailed results).

4.2. Experiments with epistasis (K=4)

Second, we have conducted the experiments with epistatic interactions among loci. Figure 3 shows the course of evolution for K=4. From the initial population, we observe approximately the same transitions as those for K=0 during the first step, but the peak value of the phenotypic plasticity, 0.91 (at around the 24th generation), was relatively smaller than that for K=0.

However, a clearly different scenario caused by epistatic interactions among loci was observed further on. While the lifetime fitness still slowly increased, the phenotypic plasticity gradually decreased to about 0.82 and then the phenotypic variation increased until around the 500th generation. This phenomenon corresponds to the second step in the three-step evolution of the Baldwin effect, in which the benefit and cost of learning worked together as previously described. The cost of learning is considered to bring about the decrease in the phenotypic plasticity. A contribution of each phenotypic value to the individual's fitness strongly depends on the other phenotypic values when there are epistatic interactions. Similarly, the learning in a plastic phenotype also affects the learning processes of the other plastic phenotypes. However, when we calculate Δt_i for each plastic trait t_i respectively, we do not consider any changes in the other plastic traits. Thus, the learning in too many plastic phenotypes does not always yield an effective increase in the whole fitness. This is the implicit cost of learning caused by the epistatic interactions among loci. At the same time, the benefit of learning is reflected in the steady transition of the innate fitness and increase in the phenotypic variation, because these transitions mean that the lifetime fitness increased by learning was getting more strongly dependent on the remaining plastic phenotypes.

Finally, the innate fitness eventually began to increase, however in contrast, the phenotypic variation decreased. Thus, the genetic assimilation occurred in the remaining plastic phenotypes because these initial phenotypic values were get-

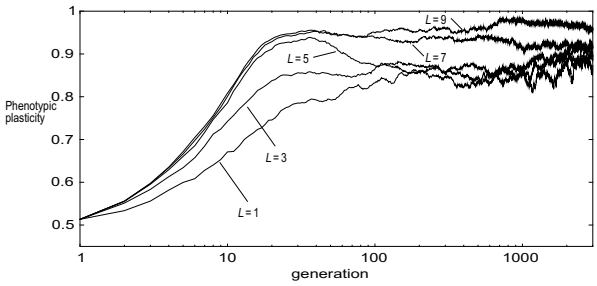

Figure 4: The evolution of phenotypic plasticity in various cases of L (K=4).

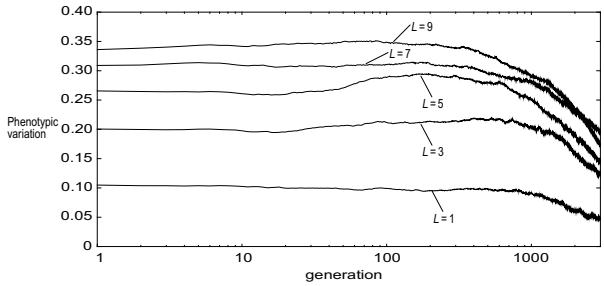

Figure 5: The evolution of phenotypic variation in various cases of L (K=4).

ting closer to resultant phenotypic values after learning. In comparison with K=0, the innate fitness converged to around 0.64 and the genetic assimilation did not occurred completely. This step approximately corresponds to the third step in our three-step evolution of the Baldwin effect, except that the phenotypic plasticity gradually increased again to high values as observed for K=0. This increase in the phenotypic plasticity in the third step and after the Baldwin effect was caused by the evolution of the genetic robustness of GI for the similar reason to the case of K=0 [9].

5. EFFECTS OF THE NUMBER OF LEARNING ITERATIONS (L)

Next, we discuss the effects of the number of learning iterations (L) on the course of evolution. This is another parameter which is supposed to affect balances between the benefit and cost of learning, because individuals have more chances to modify its phenotypic value in order to increase their own lifetime fitness as L becomes larger. We adopted the following parameters: P=20, N=15, L=1, 3, 5, 7, 9, β=10.0, p_m=0.003, d=0.03 and K=4.

Figure 4 shows the transitions of the phenotypic plasticity when L=1, 3, 5 (the default value in previous experiments), 7 and 9. These results are averages over 10 trials for initial 3000 generations which is sufficient to discuss the effect of L on the three-step of the Baldwin effect. This figure shows that only when L=5, the clear increase and subsequent decrease in phenotypic plasticity, which corresponds to the transition in the first and second step, occurred. When L=1 and 3, the phenotypic plasticity slowly and monotonously increased and we could not find its peak until the end. In this case, the small number of learning iterations did not bring about so significant an increase in the lifetime fitness as to the rapid increase in the first step.

On the other hand, the phenotypic plasticity rapidly increased from the initial generation when L= 7 and 9, which is approximately similar to the case of L=5. However, we could not observe its apparent decrease in later generations in these cases. It is due to the fact that the sufficient amount of the benefit of learning compensated for the effects of implicit cost of learning caused by the epistatic interactions among loci. Thus, the evolutionary scenario became similar to the two-step evolution of the Baldwin effect which was observed in the experiments without epistasis.

Figure 5 shows the evolution of phenotypic variation in the same experiments as shown in Figure 4. We observe the average phenotypic variation became larger as L increases as a whole. However the apparent increase and subsequent decrease in the phenotypic variation which corresponds to transition in the second and third step was observed only when K=5. These results imply that the three-step evolution of the Baldwin effect is not only sensitive to the implicit cost of learning caused by epistasis but also sensitive to the benefit of learning caused by learning iterations

6. CONCLUSION

Recent progresses in the modern genetics have reported the empirical results that the expressions of quantitative traits are based on the complex regulations controlled by the quantitative trait loci and many evironmental factors[6]. It is now essential to investigate into evolutionary models based on the constructive approaches with epistatic effects in conjunction with experimental biology in order to understand these mechanisms in real environments.

We have discussed the quantitative evolution of phenotypic plasticity based on an extended version of the NK fitness landscape by focusing on the two different parameters: the degree of epistasis and the number of learning iterations. The results have shown that the existence of epistasis caused the implicit cost of learning and the iterations of learning processes brought about the explicit benefit of learning. What needs to be emphasized is that the dynamic emergence of different kinds of benefit and cost of learning can guide the complex evolution of phenotypic plasticity.

Table 2: The benefit and cost of learning which caused the three-step evolution of the Baldwin effect.

step	benefit of learning	cost of learning
1st	search for an adaptive phenotypic configuration on a phenotypic space in every direction by adjusting many phenotypic values	———
2nd	directional and long-distance search by adjusting the small number of phenotypic values	epistatic interactions among plastic phenotypes with limitation of learning iterations
3rd	genetic robustness against mutations	limitation of learning iterations
post-Baldwin effect	genetic robustness against mutations	———

Table 2 shows the benefit and cost of learning which caused the three-step evolution of the Baldwin effect in each step. Here we regard a learning as a searching process for an adaptive phenotypic configuration on the phenotypic space and it starts from a genetically defined phenotypic configuration. The evolutionary scenario is summarized as the following images. The learning in many plastic phenotypes in the first step corresponds to the search in every direction by adjusting many phenotypic values with their plasticity. It brings about the explicit benefit of learning, that is, the increase in the lifetime fitness.

However, the implicit cost of learning, which is caused by epistatic interactions among plastic phenotypes with the limitation of learning iterations, occurs in the second step. It limits the searchable area on the phenotypic space. Therefore, the population adopts the directional and long-distance search by drastically adjusting the small number of phenotypic values. This corresponds to the decrease in the phenotypic plasticity and increase in the phenotypic variation.

If the learned phenotypic configuration obtains a sufficiently large fitness, the learning guides the initial (genetically defined) configuration to get close to the learned one due to the cost of learning resulted from the limit of the learning iterations through the third step. At the same time, the population slowly increases its phenotypic plasticity again so as to avoid the decrease in the fitness caused by mutational perturbations and keeps it high even after the Baldwin effect. This scenario is dependent on the balances between the benefit and cost of learning in the second step.

Future work includes comparisons between these scenarios and real biological examples which show complex interactions among plastic phenotypes.

7. REFERENCES

[1] J. M. Baldwin, "A New Factor in Evolution," *American Naturalist*, 30: 441–451, 1896.

[2] L. Bull, "On the Baldwin Effect," *Artificial Life*, 5(3): 465–480, 1999.

[3] T. J. DeWitt and S. M. Scheiner (eds), *Phenotypic Plasticity: Functional and Conceptual Approaches*, Oxford University Press, New York, 2003.

[4] G. E. Hinton and S. J. Nowlan, "How Learning Can Guide Evolution," *Complex Systems*, 1: 495–502, 1987.

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

[6] T. F. C. Mackay, "Quantitative Trait Loci in Drosophila," *Natture Reviews Genetics*, 2(1): 11–20, 2001.

[7] G. Mayley, "Landscapes, Learning Costs and Genetic Assimilation," *Evolutionary Computation*, 4(3): 213–234, 1996.

[8] R. Suzuki and T. Arita, "The Baldwin Effect Revisited: Three Steps Characterized by the Quantitative Evolution of Phenotypic Plasticity," In *Seventh European Conference on Artificial Life*, Springer-Verlag, Berlin, pp. 395–404, 2003.

[9] R. Suzuki and T. Arita, "Drastic Changes in Roles of Learning in the Course of Evolution," In *Artificial Life IX*, MIT Press, Cambridge, MA, pp. 369–374, 2004.

[10] R. Suzuki and T. Arita, "Interactions between Learning and Evolution: The Outstanding Strategy Generated by the Baldwin Effect," *Biosystems*, in press.

[11] P. Turney, D. Whitley and R. W. Anderson, "Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect," *Evolutionary Computation*, 4(3): 4–8, 1996.

[12] C. H. Waddington, "Canalization of Development and the Inheritance of Acquired Characters," *Nature*, 150: 563–565, 1942.

[13] B. H. Weber and D. J. Depew (eds), *Evolution and Learning - The Baldwin Effect Reconsidered* -, MIT Press, Cambridge, MA, 2003.