Culture and the Baldwin Effect

Diego Federici[1]

Norwegian University of Science and Technology
Department of computer and information science
N-7491 Trondheim, Norway
federici@idi.ntnu.no,
http://www.idi.ntnu.no/~federici

Abstract. It is believed that the second phase of the Baldwin effect is basically governed by the cost of learning. In this paper we argue that when learning takes place the fitness landscape undergoes a modification that might block the Baldwin effect even if the cost of learning is high. The argument is that learning strategies will bias the evolutionary process towards individuals that genetically acquire better compared to individuals that genetically behave better. Once this process starts the probability of experiencing the Baldwin effect decreases dramatically, whatever the learning cost. A simulation with evolving learning individuals capable of communication is set to show this effect. The set of acquired behaviors (culture) competes with the instinctive one (genes) giving rise to a co-evolutionary effect.

1 Introduction

1.1 The Baldwin effect

In the context of the debate between Darwinism and Lamarckism, James Mark Baldwin (1896) proposed that phenotypic plasticity might be regarded as "a new factor in evolution" [1]. Phenotypic plasticity allowing adaptation, would smooth the fitness landscape increasing the efficiency of the evolutionary process [2, 3]. However, phenotypic plasticity has inherent costs associated with the training phase in terms of energy, time and eventual mistakes. For these reasons, in a second phase, evolution may find a way to achieve the same successful behaviors without plasticity.

Thus the Baldwin effect has two phases. During the first phase, adapting individuals can, in same cases, acquire behaviors that help them achieving higher fitness scores. But because of the costs of adaptation, there is an evolutionary advantage towards the discovery of equivalent instinctive behaviors. Thus in this second phase, a behavior that was once learned may eventually become instinctive (see also below and [1–5]) In computer science, the phenotypic plasticity is analog to a local search strategy. The evolutionary process and the local search may be used in combination, often achieving higher efficiency than either of the methods alone [5, 6].

There are three basic requirements for the second phase to take place. First there must be a cost for the local search. In this way, the evolutionary process will have a reason (in terms of inclusive fitness) for the genetic assimilation to take place. This also means that in some settings, i.e. in a fast changing environment, genetic assimilation will never take place. With those setting, plasticity would be *the* optimal strategy. We will refer to this characteristic by *Assimilation Advantage*.

Also, genetic assimilation requires for the optimal strategy, acquired first through local search, to be expressible by the genotype. This might be impossible under some genotype-phenotype mapping strategies in which, the phenotype plasticity is required as part of the developmental process [1]. We will refer to this characteristic as *Genotypic Expressibility*.

In addition, the probability of the assimilation depends on the distance between the genotype using plasticity and the one not using it. The distance would be measured using the metric imposed by the genetic operators. A small distance is possible if there is a strong neighborhood correlation in the transformation from genotypic to phenotypic space. Where the distance is too high, the probability of genetic assimilation could be so little to be considered actually impossible. We will refer to this as *Genotypic-Phenotypic Correlation*

1.2 How learning effects evolution

In a now famous paper, Hinton and Nowlan [6] proved that with the help a local search mechanism it is possible to speed up evolution in a hard fitness landscape.

In the Hinton and Nowlan example, the adaptive solution is ideally placed in the middle between the lowest fitness solutions and the single high one, hence smoothing the fitness landscape. Adaptation in this case is a step towards the discovery of the best non-adaptive solution. The same considerations apply to other examples such as [2, 8] among others.

In these cases adaptation success is not affected by any genetically coded learning strategy. We argue that the search for good learning strategies might distract the evolutionary process from the discovery of fit non-adaptive behaviors. In other words, co-evolution of learning and the non-learning strategies modifies the fitness landscape. The quality of these modifications is an other factor that governs the second phase Baldwin effect.

2 Learning, culture and fitness

Learning can be seen as the process of acquiring behaviors. The difficulty and time lost acquiring behaviors constitute a cost of learning. We have to introduce a clear distinction between instinctive and acquired behaviors. Instinctive are those behaviors that emerge steadily and directly from the genotype, while acquired ones are those that emerge through the interaction with the environment.

[1] Like in the development of the retina [7]

If we consider individuals belonging to a population sharing the same genotype, their individual fitness can be considered the sum of shared population fitness (PopFit), fitness change due to local environmental characteristics (LEFit) and fitness change due to individual specific behavior (IFit): $Fitness = PFit + LEFit + IFit$.

The LEFit can be considered as noise and could be absent in ideal experimental settings (all individuals having the exact same initial conditions or long fitness tests). Fitness deriving from acquired behaviors (IFit) constitutes the value of the learning process and incorporates the cost of learning.

When PFit $\gg$ IFit, the advantage for plasticity is negligible. Otherwise acquired behaviors may provide an advantage to genotypically similar individuals (see figure 1). In this case, there is strong evolutionary pressure towards the discovery of better acquisition mechanisms.

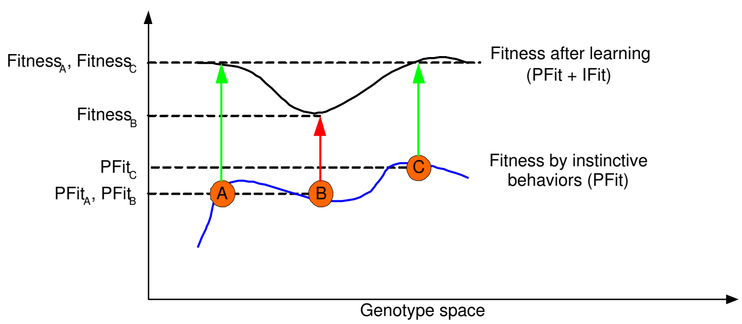


Fig. 1. Two similar individuals (A and B), share similar genotype and PFit values. By acquiring different behaviors they achieve different fitness scores. On the other side, two genotypically different individuals (A and C) reach the same fitness values because the same acquired behavior shadows the instinctive ones.

2.1 Memes

As genes form the transmission medium of biological systems, memes [9] do for acquired behaviors. Memes will be considered behavioral information blocks[2]. Basically memes are those things that "leap from brain to brain" [9] carrying a behavioral content. To strike a comparison to human society, we will call the set of transmittable behaviors *Culture*.

3 Simulation details

We set up a population of learning individuals. Each individual/agent is equipped with a single layer neural network (NN) subject to an evolutionary process and a

[2] Memes usually have a wider definition, but considering only the behavioral ones, the discussion is simplified

classifier-like system (memes), see Figure 3. Agents perceive resources and other bots from all tiles in a hamming distance of 2 (see figure 2), this constitutes the input vector.

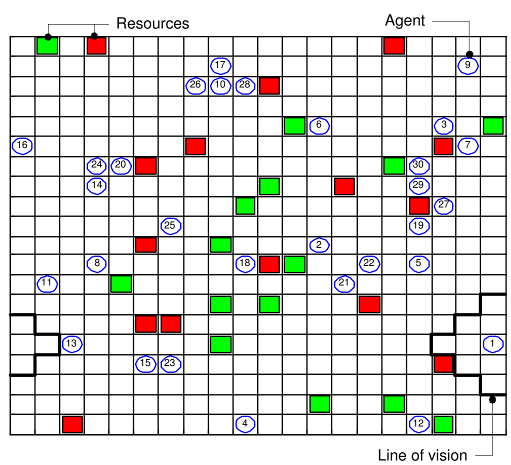
The NN produces an output vector with the expected reinforcement for each of the possible actions: don't move, go north, west, south and east.

Memes remind the agent the reinforcement experienced in the past. They are constituted by an input pattern P, an action a and an expected reinforcement R. If the pattern P matches the present input vector, then the meme replaces the output of the genetically evolved NN with R for the given action a. Basically the meme can recognize a particular sensory context (P) and reminds the agent that in the past he had performed a certain action (a) and the action yielded a given reinforcement (R).

The four expected reinforcements, generated by the NN and eventually modified by the memes, are used to stochastically select the action performed by the agent.

Agents score fitness by collecting resources spread at random in a toroidal map. Each resource type gives a fixed amount of reinforcement. If an agent does not move or collides with another agent it receives a small negative reinforcement[3]. Fitness is the sum of all reinforcements received over a fixed number of iterations. Agents undergo a steady state selection with a replacement fraction of 25%. Surviving individuals keep their memes. The offspring is placed in the proximity of a parent.

Fig. 2. Simulated Environment. The vision range of agent 1 is printed in a thick line. Two different types of resources are present. The resources represented by a darker color give a fixed negative reinforcement and fitness value, while the others give a fixed positive value. Resource types never change value and when consumed are regenerated on a random tile. Newly generated agents are placed in proximity of a parent.



[3] These penalties were added to speed up the evolutionary process

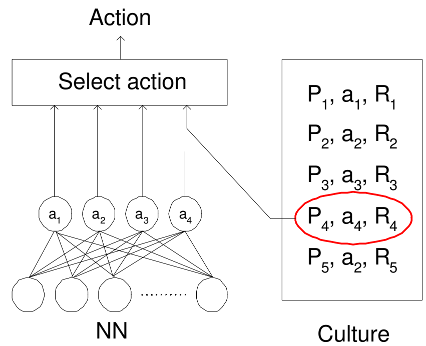


Fig. 3. Agent controller. The genetically evolved NN and the acquired culture are activated in parallel. When a pattern P matches the current input, the corresponding meme is activated (encircled in the figure). Its expected reinforcement R_4 replaces the NN output for action a_4. The performed action is selected stochastically giving a higher probability to higher expected reinforcements.

3.1 Cultural evolution

An agent culture is built by a certain number of memes (20 at maximum in these simulations). Memes can be acquired in two different ways.

First by transmission, whenever two individuals are next to each other. Every iteration a fixed number of memes can be transmitted. These memes are probabilistically selected from the transmitter culture according to their estimated value. The number of memes transmitted when two agents are in contact (communication speed, CS) is varied from 0 (no transmission) to 20 (all the meme pool is transmitted in a single iteration). As the CS increases, the agents can acquire fit behaviors earlier during the fitness evaluation, hence the *Assimilation Advantage* is reduced.

The second way is through operant conditioning.

Operant conditioning is a learning mechanism that has been observed in a variety of animals. When an animal experiences a reinforcement, its brain tries to explain what caused the reward. The effect is that the behaviors that are thought to be responsible of the reinforcement are rewarded. Learning appears to build a relationship between behavior and reinforcement based on two general assumptions: the behavior that steadily is followed by reinforcement is held responsible for it, behavior and reward must fall into a certain time window. These assumptions have strong biological and psychological support [10–12].

Whenever an agent experiences an unexpected reinforcement a meme is generated from this situation. The value of the meme changes as it is used, increasing when it helps predict the expected reinforcement. This inhibition/enhancement is an explicit measure of the memes fitness, and is used to drive the memetic evolutionary process. Unfit memes can be explicitly identified and dropped, fit ones will proliferate through transmission and new variants will eventually be generated.

Memes variants are generated by merging, a stochastic generalization mechanism. Merging is the memetic equivalent of crossover and mutation in genes. It can occur if two memes code the same action and expected reinforcement. In this case, the merging probability is proportional to the hamming distance between the memes input matching patterns. Those parts that are different in the two input matching patterns are replaced by *don't care* symbols.

Merging is a weak simplification of a boolean function:
given $(P_1 \wedge a \mapsto R)$ and $(P_2 \wedge a \mapsto R)$ then with probability $\sim d_H(P_1, P_2)$ replace them with $((P_1 \circledast P_2) \wedge a \mapsto R)$; where $P_i \in$ pattern, $a \in$ action, $R \in$ reinforcement, d_H is the Hamming distance, and $\circledast$ is a bitwise operator $\circledast(b_i, b_j) \ = \{\ b_i$ if $b_i = b_j$, *don't care* if $b_i \neq b_j\ \}$.

If it does not merge, a meme can be added only if the meme pool size does not exceed the maximum. If the maximum is exceeded a meme is dropped, the less general being dropped with higher probability. Because merging of memes can sometimes produce unfit memes, if the expected reward does not match the one experienced, the meme responsible for the error is removed.

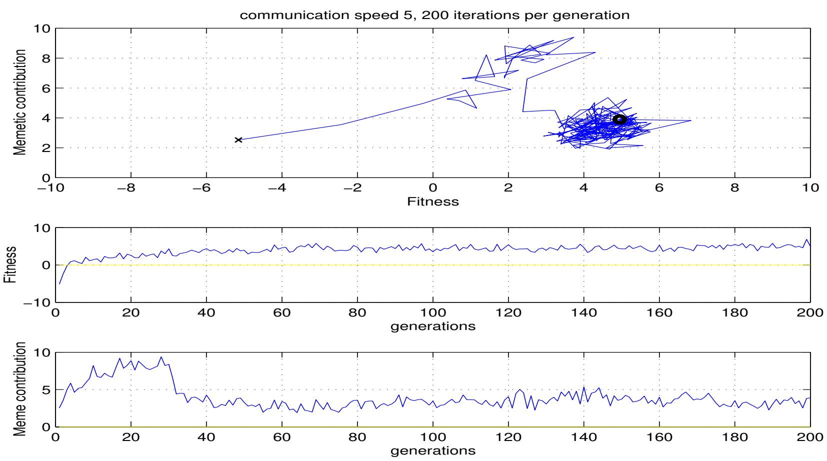


Fig. 4. run showing the Baldwin effect, convergence to the $\mathbb{MG}$ attractor. The memetic contribution to the fitness is at first high. As it decreases, the memetic behavior is partially assimilated in the genes.

4 Results

Agents can transmit some of their memes whenever their are next to each other. The number of memes transmitted when two agents are in contact (communication speed, CS) is varied from 0 (no transmission) to 20 (all the meme pool is transmitted in a single iteration).

As the CS increases, the agents can acquire fit behaviors earlier during the fitness evaluation, hence the *Assimilation Advantage* is reduced.

It is possible to evaluate the amount of fitness generated through acquired behaviors by stopping the simulation every generation and recording how much

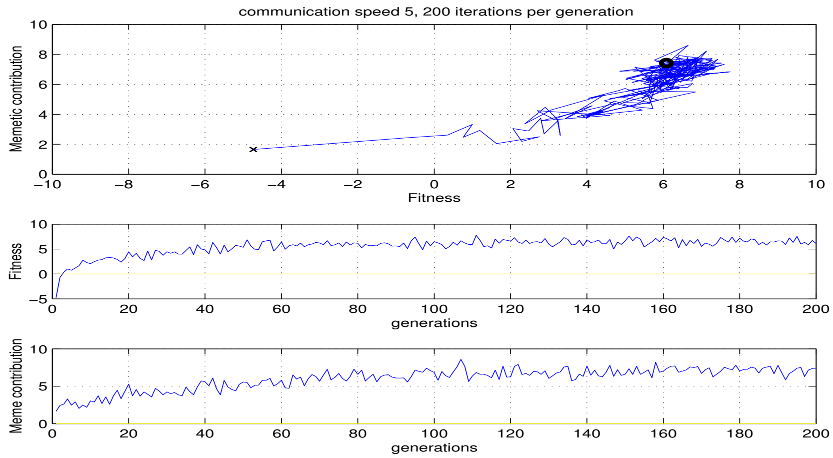


Fig. 5. run not showing the Baldwin effect, convergence to the $\mathbb{M}$ attractor. Genetic assimilation does not take place.

fitness the agents score with and without the help of memes. It is so possible to plot how much the memes contribute to the total fitness score (memetic contribution). Figures 4 and 5 show three different plots. The first one is a state diagram, memetic contribution vs. fitness, showing the trajectory that a population undergoes during evolution. The second one displays the amount of average fitness scored by the population, and the third the quantity of memetic contribution.

The state diagram is particularly useful because it shows the attractors of the evolutionary process. Figure 4 shows the attractor for a typical population that went through the second phase of the Baldwin effect, figure 5 one that doesn't show it.

In almost every setting two attractors, such as those in figure 4 and 5 are present. The two attractors show the convergence basin of different strategies. The first relies both on memes and genes ($\mathbb{MG}$) with the instinctive behaviors capable of scoring some fitness. The second relies on memes only ($\mathbb{M}$).

It is then understandable that while the dynamic towards $\mathbb{MG}$ is associated to the second phase of the Baldwin effect, the path to $\mathbb{M}$ is not.

Figure 6 shows with which frequency a population falls in the $\mathbb{MG}$ attractor.

One would expect to experience a decreasing number of populations falling in the $\mathbb{MG}$ attractor as CS is being increased. In fact, increasing CS reduces the cost of learning and the advantage for genetic assimilation. Instead the frequency decreases to a minimum and then increases again. This proves that the Baldwin effect is influenced by other factors apart the cost of learning.

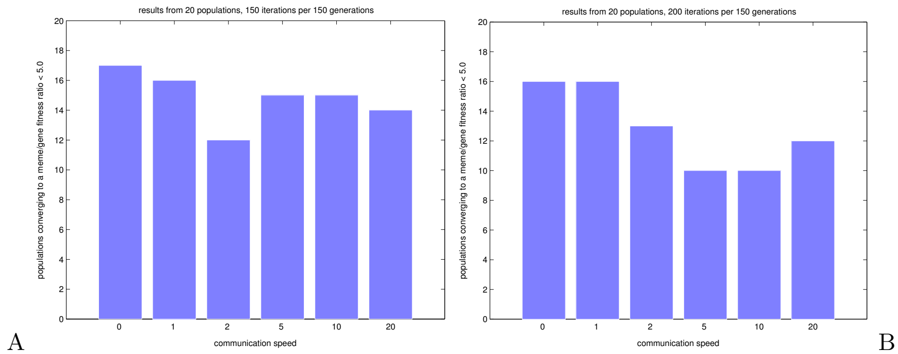


Fig. 6. Populations falling into the $\mathbb{MG}$ attractor with each CS setting

4.1 Different strategies

An agent can acquire fit behaviors in three possible ways: evolving the proper NN, building memes by operant conditioning, and acquiring memes from other agents. CS can affect only the latter.

Being born next to its parent without any meme, an agent is faced with a dilemma. It can either try:

1. first receive as many memes as possible from other agents, and then exploit them to score fitness
2. start scoring fitness immediately and acquire memes by itself

Either choice requires a genetically coded strategy, a social one in the first case, and an asocial one in the latter. The social strategy, scoring fitness mainly by memes, is the one that converges to $\mathbb{M}$. $\mathbb{MG}$ is instead the attractor of the asocial strategy.

CS does not only change the cost of learning, modulating the acquisition speed. It also changes the nature of the two strategies $\mathbb{M}$ and $\mathbb{MG}$.

As CS decreases the social strategy becomes more difficult. In other words it requires a more committing strategy and a more specialized genotype. In fact, it must strongly avoid any asocial behavior, otherwise it will risk to interrupt the acquisition phase. On the other side when CS is maximal, the parent's culture is acquired in a single shot. The social strategy is achieved simply by being born. In this case $\mathbb{M}$ and $\mathbb{MG}$ will be maximally overlapping because an asocial individual will receive the same amount of culture as a social one.

The difference between the two strategies determines the probability of moving from one to the other. The higher the difference the lower the probability.

A second aspect is that the quality of culture (the amount of fitness it can guarantee) is not fixed. As genes evolve every generation, memes do every iteration. If CS is zero, agents cannot share their memes and the culture quality improves in a given generation but not across generations. If CS is high, even if an agent dies some of its memes might survive and continue evolving on other hosts.

The fitness advantage of a strategy or the other, is subjected by the actual level of cultural evolution. When many agents do not socialize, culture improves slowly and the social strategy is less attractive. With many social agents, cultural evolution is faster and offers a greater prize for adaptivity.

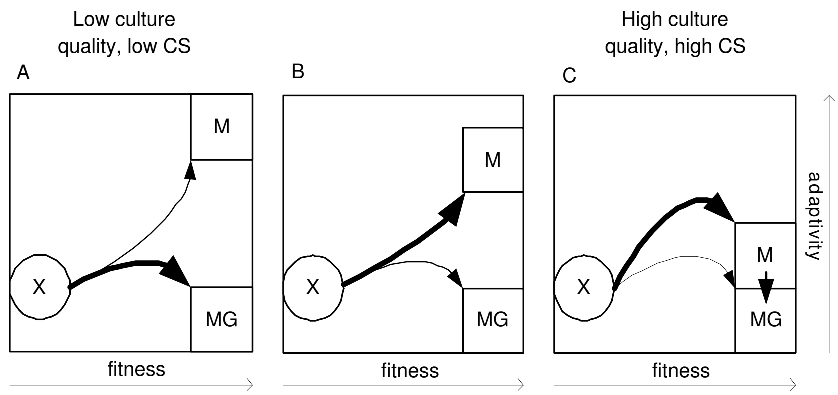


Fig. 7. effects of CS to evolvability

As the cultural quality evolves or fails to evolve, the fitness landscape changes. If cultural evolution proceeds steadily, the $\mathbb{M}$ strategies will be able to dig into their attractor increasing their stability. Still, if $\mathbb{MG}$ lies too close to the $\mathbb{M}$ attractor, this is insufficient to prevent the second phase of the Baldwin effect. Figure 7 summarize these concepts.

Figure 7A With low CS, the quality of culture is not enough so that the asocial strategy $\mathbb{MG}$ is more convenient (the thicker line indicates a higher transition probability).

Figure 7B CS increases and the $\mathbb{M}$ attractor becomes stronger. As the attractors are far away in genotypic space, the probability to move from $\mathbb{M}$ to $\mathbb{MG}$ is very low. The frequency of the second phase of the Baldwin effect is minimal.

Figure 7C CS continues to increase, the social and asocial strategy are very similar and passage from $\mathbb{M}$ to $\mathbb{MG}$ is more probable.

5 Conclusions

We have provided an example in which genetic assimilation cannot be explained by the Baldwin effect alone.

It is suggested that the Baldwin effect considers cases in which the genotype cannot modify directly the adaptation mechanism. Under these conditions, the genetic search can evolve only the non-adaptive part of the phenotype, and only that can provide a continuous fitness improvement. Examples for this case can be found in [6, 2, 8, 13].

We argue that the adaptive behavior must be somehow expressed in the genotype, so that evolution affects also the adaptation mechanism and its quality.

Evolution can than proceed in at least two directions, one towards the discovery of better adaptive strategies (in this case the social behavior), the second towards the discovery of fitter instinctive behaviors (the asocial behavior).

The state of the evolution of adaptivity affects both the cost of learning and the correlation between genotype and phenotype. This can cause that both the cost of learning and the correlation decrease, so that the probability of observing the Baldwin effect is not a monotonic function.

Even in a static environment[4] as the one provided in this paper, the fitness landscape will undergo a dynamic. Under these circumstances, genetic assimilation is ruled more by the quality of the fitness landscape dynamic than by the assimilation advantage.

The simulations presented in this paper are very computational expensive and have been run on ClustIS[5] cluster.

References

1. J.M. Baldwin: A new factor in evolution. American Naturalist **30** 441–451 (1896)
2. K. Downing: Reinforced Genetic Programming. Genetic Programming & Evolvable Machines. Kluwer Publishers 259–288 (2001)
3. P. Turney: Myths and legends of the Baldwin Effect. Proceedings of the ICML-96 (13th International Conference on Machine Learning) 135–142 (1996)
4. P. Turney, D. Whitley, R. Anderson: Evolution, learning and instict: 100 years of the Baldwin effect. Evolutionary Computation **4:3** 206–212 (1996)
5. R.K. Belew, M. Mitchell: Adaptive individuals in evolving populations: models and algorithms. addison-wesley (1996)
6. G.E. Hinton, S.J. Nowlan: How learning can guide evolution. Complex Systems **1** 495–502 (1987)
7. L.M. Chalupa, B. L. Finlay: Development and organization of the retina : from molecules to function. Plenum Press, New York (1998)
8. R. French, A. Messinger: Genes, phenes and the Baldwin effect. In Proceedings of Artificial Life IV 277–282 (1994)
9. R. Dawkins: The selfish gene. Oxford university press (1976)
10. E.R. Kandel, J.S. Schwartz, T.M. Jessell: Principles of neural science. Elsevier Science Publishing Co., New York (2000) 4th edition
11. H. Markram, J. L'bke, M. Frotscher, B. Sakmann: Regulation of Synaptic Efficacy by Coincidence of Post-synaptic APs and EPSPs. Science, **275** 213–215 (1997)
12. T.J. Carew Behavioral Neurobiology: The Cellular Organization of Natural Behavior Sinauer Associates (2000)
13. T. Arita, R. Suzuki Interactions between learning and evolution: The outstanding strategy generated by the Baldwin effect. In Proceedings of Artificial Life VII, 196–205 (2000)

[4] where the best strategy does not require adaptation

[5] information can be obtained from http://ClustIS.idi.ntnu.no/