The Baldwin Effect under Spatial Isolation and Autonomous Reproduction

H. L. Peng and J. C. Tay*
Evolutionary and Complex Systems Programme, Nanyang Technological University
Blk N4, #2a-32 Nanyang Avenue, Singapore 639798
*Email: asjctay@ntu.edu.sg

ABSTRACT
The impact of learning on evolution in dynamic environments undergoes recognized stages of the Baldwin Effect although its cause is not clear. To identify it experimentally, we devise spatial constraints and allowed autonomous reproduction for a multi-agent game play using Iterated Prisoner's Dilemma. In comparison to Arita and Suzuki's model, we exclude the selective bias of noise, reduce the genotypic and phenotypic correlation and explain how mutant strategies can survive under costs assumptions. Simulation results show that plastic agents are exploited by Defecting, non-plastic agents during game-play. In the majority of the simulations, this Nash Equilibrium Phase occurs, resulting in a Defect-oriented population which eventually wipes itself out. However, with spatial isolation, a small number of cooperative plastic and non-plastic agents survive and proliferate in the simulation after the Defecting agents have died out; the two transitions stages ascribed to the Baldwin Effect are then observed to occur.

Categories and Subject Descriptors
I.6.0 [**Simulation and Modeling**]: General; J.3 [**Life and Medical Sciences**] - *biology and genetics*.

General Terms
Design, Experimentation, Theory, Verification.

Keywords
Baldwin effect, learning, evolution, prisoner's dilemma, multi-agent simulation.

1. INTRODUCTION
The Baldwin Effect [2] implies that individual lifetime learning can influence the course of evolution. It proposes a new relationship between the genotype, the genetic makeup of an organism, and the phenotype, the trait displayed by an organism, which may be inborn or a consequence of the environment. The first stage occurs when learning contributes to the fitness of the individual and allows those individuals who acquire a suitable learnt trait to proliferate in the population. This will be evidenced when an increase in the phenotypic plasticity of the population leads to a corresponding increase in the average fitness. The second stage occurs when the environmental stimulus stabilizes and evolution discovers and activates the dormant genotypes to incorporate the learnt traits. Over time, evolution via Darwinian natural selection, will eventually encode these acquired traits into the genotype. This second stage requires two conditions [3]. Firstly, there must be a sufficiently high cost of learning to justify the genetic assimilation of learnt traits. Secondly, there has to be a correlation between genotypic and phenotypic space, without which changes in the phenotype cannot be said to have any effect on the genotype.

GECCO'06, July 8–12, 2006, Seattle, Washington, USA.
ACM 1-59593-186-4/06/0007.

2. ITERATED PRISONER'S DILEMMA
The Iterated Prisoner's Dilemma (IPD) [4] is often adopted to simulate the dynamic interactions between the members of the population for simulating the Baldwin Effect. Players can either "cooperate" or "defect" and are scored according to a possible payoff matrix given in Figure 1.

	Cooperate	Defect
Cooperate	3,3	0,5
Defect	5,0	1,1

Figure 1. Example of an IPD Payoff Matrix

A strategy for IPD may be represented by a string of '0's and '1's, with '1' referring to a move to Cooperate and '0' a move to 'Defect'. As a single game can only have 4 distinct outcomes, strategies that rely only on the history of the previous round can be represented with just a 4-bit string, with each bit telling the agent which move to adopt for the current round based on the outcome of the previous round. Learning is incorporated into the strategy through plastic bits which are represented by an 'X' in the strategy description.

3. A MULTI-AGENT IPD SIMULATION
We implement a Multi-Agent IPD SimulatiON (MAISON) where the agents are free to roam about in a two dimensional grid. Agents are implemented as software threads which carry out various activities such as movement, reproduction and interaction based on playing IPD games with other agents. The agents possess an energy parameter which is proportional to the score of their game performance. There are three types of IPD game playing agents. The most basic type performs random strategies and act as Brownian agents for the experiment. The second type plays according to the game strategy inherited from the parent through asexual reproduction with mutation. The specific strategy does not change throughout the course of the agent's lifetime. The purpose for the second type of agent is to provide a basis of comparison for learning agents, which are the third type.

4. MODEL DESIGN

If noise is present, cooperative relationships may be disrupted if one move to Cooperate was changed to Defect. Thereafter, the opponent will also choose to Defect. Hence, strategies with capabilities to resume cooperative relationships rapidly after defection due to noise will thrive better in a noisy setting. This implies that highly plastic agents may be given an incidental selective advantage; thereby increasing their fitnesses and promoting the first stage of the Baldwin Effect (see Section 1). Without noise, the single cost of learning in the MAISON model is represented by the losses incurred in attempting to achieve a favourable strategy as each new opponent is encountered. We also believe that the stagnation (or stabilisation) of game play is acceptable if the number of rounds for each game is not large. Otherwise, agents which have achieved a cooperative relationship may gain excessively in a single game. In the MAISON model, only new-born agents will carry a strategy phenotype that is based on the genotype. Upon interacting with other agents, the agent then brings forward the new strategy it learns from the previous interactions to future games. Therefore, genotypic change will have a reduced impact while the effect of learning is increased.

5. EXPERIMENTAL RESULTS

We create a null model to provide a basis for comparison with [1]. It is constrained spatially by restricting the simulation to a single cell (i.e. a grid of 1 cell), thereby inhibiting any movement for each agent. As agents can only interact with other agents in the same cell, these constraints enforce direct interaction among all agents. We also remove the autonomous reproduction capability of the MAISON agents. Hence, a centralized control mechanism is required to regulate the reproduction mechanism. This is done by disabling individual reproduction until the population level falls to 25% of the original level. The simulation is then paused and the remaining agents replicate themselves until their numbers return to 100% of the initial population. This method maintains the population at a steady level. With autonomous movement and reproduction enabled for the agents in MAISON, the experimental results differ from the null model. In the early stages, Defect-oriented strategies (0001 or 0000) often prevail, resulting in new generations of agents which inherit these strategies. The proliferation of such strategies lead to drastic falls in the average energy gained from each game played, since most players are defecting. There is also a gradual decrease in plasticity as Defect-oriented play seems to be the most rewarding. We term this the **Nash Equillibrium Phase (NEP)**. Without centralized maintenance of the population level, the population level during NEP falls to zero eventually. The majority of our simulated populations do not progress past NEP. Overall, the Baldwin Effect is not exhibited and phenotypic traits are not genetically assimilated. However, we do observe a small proportion (about 5%) of populations that do survive past NEP. This applies only for the learning agents and can occur in two ways.

High initial plasticity and spatial isolation. Some populations 'discover' the ideal strategy of X00X early by having higher than average plasticity in the initial stages. This only holds true for a population isolated to a single cell within the whole environment. This cell's population will survive past NEP if it is not flooded with Defect-oriented agents from adjacent cells. After which, the other cell populations with predominantly Defect-oriented strategies would die out, leaving the population with the X00X strategy (or those with equivalent cost) free to reproduce and then propagate to adjacent cells. Eventually, the whole environment will be filled with descendents of the X00X strategy and reach its limit of 5000 agents. Figure 2 shows the results of one such run. The simulation starts off with an initial population of 1000 agents.

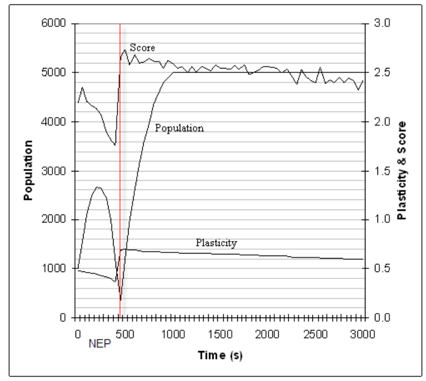

Figure 2. Experimental results of population that do survive past NEP

Low plasticity and spatial isolation. In the second way, the agents which survive the initial defective stage need not be X00X nor have high plasticity. They may consist of strategies such as 1001. These strategies may survive, isolated within their individual cells while the Defect-oriented strategies wipe out each other in the other cells. The popular strategies that will appear in almost all simulations will be 0001 or 000X. These two strategies are inherently "selfish" as they will only cooperate if the opponent does so in the previous round. These strategies often become popular in the simulations just before the end of NEP, resulting in the prevalent defective play that eventually causes the population to destroy themselves. Plastic strategies such as X0XX thrive only after surviving NEP. Of the populations that survive NEP, those agents with more plastic phenotypes will gain a considerable advantage and proliferate within the environment. This can be seen in the simulation where the average plasticity of the agent population rises (Figure 2). This represents the first stage of the Baldwin Effect. Eventually, this highly plastic population will discover the X00X strategy, which incurs less implicit learning costs [1] than more plastic strategies such as X0XX. This represents the second stage of the Baldwin Effect where plasticity falls with genetic assimilation due to the presence of learning costs. The evolutionary choice of the X00X over other strategies with two plastic bits (00XX or X0X1) is such that the particular combination incurs the least losses (learning costs), even with the same level of plasticity. Experimentation with the random play agents produced results that were anticipated.

6. REFERENCES

[1] Arita, T. and Suzuki, R., "Interactions between Learning and Evolution -Outstanding Strategy generated by the Baldwin Effect," In *Proceedings of Artificial Life VII*, 196-205, 2000.

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

[3] Mayley G., "Landscapes, Learning Costs, and Genetic Assimilation." *Evolutionary Computation* 4(3): 213-234, 1997.