Moderate Environmental Variation Promotes Adaptation in Artificial Evolution

Nicola Milano*, Jônata Tyska Carvalho*, Stefano Nolfi*

*Institute of Cognitive Sciences and Technologies
National Research Council (CNR), Roma, Italia,

#Center for Computational Sciences (C3), Federal University of Rio Grande (FURG), Av. Italia, km 8, Rio Grande, Brasil
jonata.carvalho@istc.cnr.it

Abstract

In this paper we analyze the role of environmental variations in the evolution of artificial agents situated in an external environment and we demonstrate how environmental variations promote the evolution of better agents. The beneficial effect is maximized at intermediate rates of variations, i.e. when the dynamics of the environment displays a sufficient level of stability and variability. The analysis of the obtained results indicate that the adaptive advantage provided by environmental variations is due to the fact that it increases the rate with which evolving agents change phylogenetically. The performance of the adaptive agents and the rate with which agents change phylogenetically are maximized at moderate rate of variations of the environment which provide a good tradeoff between environmental variation and stability.

Keywords—environmental variations; evolvability; stability; artificial evolution.

Introduction

The last two decades have seen an increasing recognition of the role of environmental variations in evolution.

The interaction between environmental conditions and the expression of genetic variation influences the evolutionary dynamics. Genes influencing a trait in one environment may not be important in a different one (Viera et al., 2000). Mutations often have environment-dependent effects (Kawecki, 1994; Szafraniec, Borts and Korona, 2001). The environmental conditions influence the genetic interactions among traits, i.e., the correlation between the genetic influences on a trait and the genetic influences of another trait, which are known to influence the evolutionary dynamics (Sgro and Hoffmann, 2004). For instance, the genetic correlations among certain traits can be positive in an environment and negative in another one. Consequently environmental variations influence evolutionary trajectories in populations (Sgro and Hoffmann, 2004).

Moreover, as stressed by West-Eberhard (2003), phenotypic variation arises not only as a result of genetic variations but also as a result of environmental variations. “Environmentally induced phenotypic changes can give rise to adaptive evolution as readily as mutational induced changes; both are equally subject to
genetic accommodation.” (West-Eberhard, 2003, pp.498).

The role of environmental variations has also been investigated in studies addressing the evolution of artificial agents situated in an external environment.

Nolfi et al. (1994) and Jacoby (1997) showed how robots evolved in simulation in noisy conditions display a greater level of robustness with respect to agents evolved in non-noisy environments. Thanks to this robustness, these robots are more likely to cross the reality gap, namely to maintain their ability to operate properly once they are moved from the simulated environment, in which they have evolved, into the real physical environment, in which they should operate.

Some authors demonstrated how the exposure to recurrent environmental variations can speed-up evolution, i.e. facilitate the adaptation to new varied environmental conditions that present similarities with environmental conditions experienced in previous generations. In particular Anderson (2013) and O’Donnell et al. (2014) showed how diploid digital organisms can readily adapt to environmental variations thanks to standing genetic variation, i.e. alleles playing an adaptive role in previous generations that were maintained in the population and that play again an adaptive role in new environmental conditions. Janssen et al. (2016) showed a similar phenomenon in haploid predator robots evolved against prey robots that periodically change their behavior throughout generations. Indeed, the exposure to this periodic variation leads to the evolution of genetic organization characterized by switching genes, i.e. single genes interacting with several other genes so to produce two well-differentiated behavioral strategies that can be selected as a result of a single point mutation affecting the switching gene.

Other studies investigated the relation between environmental variation and ontogenetic plasticity (Nolfi and Floreano, 2011). The results of these experiments indicate that the possibility to adapt ontogenetically, during the interaction with the environment, enables evolving agents to adapt to environmental variations at a faster time rate with respect to agents that adapt phylogenetically only.

Finally, other studies investigated the role of environmental variations in more abstract models in which the fitness of the agent is proportional to the distance from a certain target point in the search space and in which environmental variations are modeled as variations of the target point (Palmer and Feldman, 2011; Kashtan, Noor, and Alon, 2007).

The fact that variable environmental conditions can promote the evolution of better solutions has already been demonstrated in the context of a competitive co-evolution (Nolfi and Floreano, 1998). In this study the authors demonstrated how predator robots evolved against co-evolving prey outperform robots evolved against fixed pre-evolved prey. Another related reported by Bongard (2011) demonstrates how agents developing into an anguilliform and then into a legged body plan evolved the required behavioral capabilities more quickly than agents that developed into a legged body plan only. Morphological variations, in fact, can be considered as variations of the agents’ internal environment. However, in these studies the adaptive advantage provided by environmental variations can be explained by the fact that the environment varies in a directional manner supporting the incremental evolution of more complex behaviors.

In this paper we demonstrate that non-directional environmental variations promote the evolution of better solutions in population of artificial agents, and that this adaptive advantage is maximized when the environment varies at a moderate rate.

The analysis of how evolving individuals change phylogenetically indicates that environmental variations facilitates phylogenetic changes in the evolving individuals which in turn facilitates the discovery of more adapted solutions. The fitness of the evolving agents and that rate at which agents change phylogenetically is maximized at intermediate rates of environmental variations, i.e. when the environment varies phylogenetically and when the new environmental conditions persist over generations long enough to enable selective reproduction to discover and retain the appropriate adaptive changes.
Method

To investigate the impact of environmental variations on evolution we evolved artificial agents situated in fixed environments and in environments that varied ontogenetically and phylogenetically. The amount of variation experienced by agents during their lifetime and the rate at which the environment varies over generation were manipulated systematically.

Agents are constituted by carts with two poles of different length attached on their top through passive joints. They are situated over a flat surface that can be inclined of a varying angle and can be constituted of materials with varying friction coefficients. Agents are provided with a neural network controller that determines the force applied by the motor of the cart at any step on the basis of the current and previous state of the agent’s sensors. The connection weights of the neural network, which determine the behavior of the agents, are encoded in artificial genotypes and evolved. The fitness of the evolving agents rate their ability to move so as to keep the two poles balanced as long as possible.

Evolving agents are evaluated during a certain number of trials (NT) during which they experience different environmental conditions. The experimental conditions experienced during these trials can vary over generations. The amount of environmental variability experienced during agents’ lifetime can be manipulated by varying the number of trials (NT). The rate at which the environment varies over generations can be manipulated by varying the number of generations after which the characteristics of the environment are modified (N). In both cases the variations of the environment are generated randomly.

The experiments have been replicated by varying systematically the number of trials (NT) and the number of generation after which the characteristics of the environment are varied (N). Each trial is characterized by a vector of 8 values that encode the characteristics of the plane (inclination and friction coefficient) and the initial states of the cart and of the poles (the position and the velocity of the cart, the angular position and the velocity of the two poles). Details are provided in the appendix. We will use the term fixed, intermediate, and always varying experimental conditions to indicate the experiments in which the environment remains constant over generations, varies every N generations, and varies every generation, respectively. In the fixed experimental condition, the NTx8 matrix used during the NT corresponding trials is chosen randomly once and is maintained constant during the entire evolutionary process. In the always-varying experimental conditions, the NTx8 matrix is generated randomly every generation. In the intermediate conditions, the matrix is changed randomly every N generations.

The evolutionary process is continued until a maximum number evaluations (trials) or generations were performed. Since the evaluation of the agents/environmental behavior constitutes the mayor computational cost of these experiments, maintaining the number of evaluations constant permits to maintain the computational cost approximately constant.

The performance of an agent, i.e. the ability of an agent to solve the problem in all possible environmental conditions is estimated by post-evaluating agents for 1000 trials in which the characteristics of the environment are varied randomly. To ensure a proper comparison the same matrix of 1000x8 values generated randomly is used for post-evaluating agents evolved in different experiments. Performance measures, therefore, rate the ability of the evolving agents to solve the problem in varying environmental conditions.

A detailed description of the agent and of the environment, of the agent’s neural controller, of the evolutionary algorithm, and of the fitness and performance measures is provided in the section 1-4 of the Appendix. The experiments performed can be replicated by downloading and installing FARSA simulator.
Moderate environmental variation promotes the evolution of better solutions

Agents exposed to variable environmental conditions during their lifetime (i.e. agents evaluated for multiple trials during which the characteristics of the environment varied) outperform agents exposed to non-variable conditions (i.e. agents evaluated for a single trial), see the appendix. The optimal number of trials varies depending on whether the environment varies also throughout generations or not. Indeed, performance are maximized by using 200 trials in the experiments in which the environment does not vary over generations and by using 25 trials in the case of the experiments in which the environment varies every generations or every 100 generations. The fact that the optimal number of trials is greater in the case of the experiments performed in environments that do not vary over generations indicates that the lack of environmental variation throughout generations can be compensated by increasing the variation of the environment during agents’ lifetime. The effect of the lack of environmental variation throughout generations, however, can be compensated only in part by increasing the amount of environmental variation during agents’ lifetime. This is demonstrated by the fact that the utilization of more than 200 trials in the fixed environmental condition produces a reduction of the obtained performance (see), and by the fact that agents evolved in environments that vary every generation or every 100 generations outperform the agents evolved in environments that do not vary over generations independently from the number of trials (Welch’s t-test p-value < 10⁻³, comparison performed by using the parameters that resulted optimal in each condition).

The performance achieved in the intermediate condition in which the environment varies every 100 generations are significantly better than the performance achieved in the fixed and always varying experimental conditions (Welch’s t-test with Bonferroni correction p-value < 0.01667, comparison performed by using the parameters that resulted optimal in each condition). The condition that leads to the best results, therefore, is the one in which the environment varies moderately both during agents’ lifetime and during generations. The intermediate condition in which the environment varies every 100 generations outperforms the fixed and always varying experimental conditions both the experiments continued for 50 millions evaluations and in the experiments continued for 5,000 generations (Welch’s t-test with Bonferroni correction p-value < 0.0123). Moreover, the intermediate condition in which the environment varies every 100 generations outperforms the fixed and always varying experimental conditions independently from whether the number of trials is to 25, 50, or 200 (Section 5 of the Appendix, Welch’s t-test with Bonferroni correction p-value < 0.0098, comparison performed by using the parameters that resulted optimal in each condition).

The fact that the performance of the agents are maximized in experiments in which the environment changes at a moderate rate over generations is confirmed by the analysis of the performance achieved in a series of experiments in which we systematically vary the rate at which the environment changes over generations (see Fig. 1). Data obtained by keeping the other parameters equal. Mutation rate was set to 1% (the value that resulted optimal in all conditions, see Section 5 of the Appendix). The number of trials and the stochasticity level (that regulates the selection pressure) were set to 50 and 0%, respectively, i.e. to values that produced high performance in all conditions (see Section 5 of the Appendix). The performances obtained when the environment varies every 10, 25, 50, 100 and 200 generations do not differ among themselves (Kruskal-Wallis ANOVA p-value=0.09) but they differ statistically from the other experimental conditions (Kruskal-Wallis ANOVA p-value p-value < 0.001, for all comparisons). In the rest of the paper, we will use the term intermediate condition to refer to the experiments in which the environment varies every 100 generations since it corresponds approximately to the average value of the best conditions.
Fig. 1. Performance in the fixed experimental condition, in eight intermediate conditions in which the environment varies every 5000, 2500, 1000, 500, 200, 100, 50, 25, and 10 generations (5000-10), and in the always varying experimental condition. Boxes represent the inter-quartile range of the data and horizontal lines inside the boxes mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. Circles indicate the outliers. The evolutionary process was continued for 50 millions evaluations. Average results of 10 replications.

The beneficial effect of moderate environmental variations persists on the long term as demonstrated by the fact that the agents evolved in the intermediate condition in which the environment varies every 100 generations outperform agents evolved in fixed and always varying conditions both after 50 and 100 million evaluations (Fig. 2, Welch's t-test with Bonferroni correction, p-value < 0.0095).

Fig. 2 Performance of agents evolved in the fixed environmental condition (fixed), in the condition in which the environment change every generations (always), and in the condition in which the environment change every 100 generations (interm) after 50 and 100 millions evaluations. Results obtained with the best parameters, i.e. mutation rate 1% in all conditions, number of trials 200 in the fixed environmental condition and 25 in the intermediate and always variable experimental conditions, stochasticity 30%, 0%, and 20% in the case of the fixed, intermediate and always variable experimental conditions. Boxes represent the
Part of these data was previously reported in Milano, Carvalho and Nolfi (in press).

### On the role of fortunate specific environmental conditions

A possible hypothesis that could explain why agents evolved in varying environmental conditions achieve better performance is that environmental variations enable evolving agents to experience, sooner or later, specific environmental conditions that promote the evolution of better solutions.

This hypothesis is based on evidences collected in incremental evolutionary experiments in which the problem and/or the environment become progressively more challenging throughout generations (Mikkulainen and Gomez, 1997; Nolfi and Floreano, 1999). The possibility to experience easier conditions during the first generations facilitates the evolution of solutions that can be then adapted to also to the most challenging conditions. Moreover it is based on evidences collected in neural network learning literature that indicate that the relative distribution of qualitatively different stimuli in the training set strongly affect the outcome of the learning process (Hare and Elman, 1995; Zhou and Liu, 2006). Environmental variations in our experiments are random and consequently cannot lead to a progressive complexification of the adaptive problem. On the other hand, the possibility to experience environmental conditions that are easier to handle with respect to average conditions and/or that include a well-balanced distribution of sensory states could boost evolution.

To verify whether the evolution of better solutions is promoted by the occurrence of specific fortunate environmental conditions we measured how the performance of evolving agents varies while they are evolved in different environments. The analysis was conducted on 30 populations of agents evolved in the intermediate experimental condition in which the environment varied every 100 generations after one, five, and ten thousand generations. One hundred copies of these populations were evolved for 500 generations in 100 different corresponding environments for 50 trials. The performances of these populations were post-evaluated every 100 generations on 729 trials in which agents were exposed to systematically varied environmental conditions (see Section 6 of the Appendix).
As can be seen, the performance of the agents after one and five thousand generations increases of a similar amount in all environments (Fig. 3 top and middle pictures). Later on, i.e. after ten thousands generations, performance variations become much smaller (Fig. 3 bottom, notice that the scale used for displaying variation is one order of magnitude smaller in the case of the bottom picture). Moreover, the ability of the agents to solve the 729 trials increases or decreases slightly while they are evolved in different environments, although the number of environments that lead to performance increase is higher than the number of environments that lead to performance loss.

This data does not show evidences of fortunate specific environmental conditions capable of boosting the evolutionary process. During the initial evolutionary phase, the evolving agents improve their ability of a
similar amount in all environmental conditions. Moreover, the evolving agents improve their ability in the majority of the environmental conditions also during successive evolutionary phases.

**Environmental variations promotes phylogenetic variation**

Another factor that could explain why environmental variation promote the evolution of better agents is constituted by the possible impact that environmental variation has on the rate at which evolving agents change phylogenetically during the evolutionary process.

To understand the possible role of this factor we should consider that adaption depends on the generation of phenotypic changes and on the retention of the changes that are adaptive. As stressed by West-Eberhard (2003), phenotypic variations arise not only as a result of genetic variations but also as a result of environmental variations. Moreover, adaptations can arise both as a result of mutational induced changes and environmentally induced changes since both are subjected to genetic accommodation (West-Eberhard, 2003). The occurrence of environmentally induced changes resulting from environmental variations in addition to genetically induced changes induced by mutations can therefore enable evolving agents to vary more phylogenetically which in turn can facilitate the discovery of better solutions.

The fact that genetic accommodation requires time can potentially explain why the advantage of environmental variations is maximized at intermediate rate of variations. Environmental variations that occur too often, e.g. every generation, do not provide enough stability over time to permit genetic accommodation. On the other hand, environmental variations that occur too rarely provide less opportunity for change. An intermediate variation rate can thus represent an optimal tradeoff between the contrastive needs of variation and stability. This hypothesis is in line with West-Eberhard’s claim that the most important reason that explains why environmentally induced changes are evolutionarily important is indeed their time persistence with respect to genetically induced changes:

*Perhaps the most compelling argument for the superiority of environmental induction over mutations in term of recurrence and persistence has to do with the inexorable persistence of an environment immune to natural selection: environmental inducers might be not only immediately widespread without necessity for positive effects on fitness sufficient to spread them due to differential reproduction of their bearers (selection), but they are inexorably present.*

West-Eberhard (2003), pp. 504

A possible hypothesis that could explain why environmental variations enable the evolution of better agents and why the performance of evolved agents is maximized at intermediate rates of environmental variations is that environmental variations increase the rate at which agents change phylogenetically and that the rate of agents’ change is maximized at intermediate rates of environmental variations.

To verify the impact of environmental variations on the rate at which agents change phylogenetically at the behavioral level we measured how often the behavior of the ancestors of the best evolved agents vary over a certain number of generations.

Fig. 4 shows the fraction of trials in which the ability to solve the double-pole navigation problem from 729 systematically differentiated environmental conditions (see Section 6 of the Appendix) varies every 100 generations. Variations include both the case in which the agent at generation N+100 is unable to solve a trial that its ancestor was able to solve and the case in which it is able to solve a trial that its ancestor was unable to solve. To ensure a proper comparison, the analysis was performed on experiments terminated after 5000 generations. As in the case of the analysis reported in Fig. 1, the number of trials were set to 50, the mutation rate to 1%, and the stochasticity level to 0%.
As expected, the rate of variation decreases through out generation as a result of the evolution of better and better agents (Fig. 4). Interestingly, the intermediate experimental condition in which the environment change every 100 generations (blue line) displays a greater amount of variation through out generations than the fixed experimental condition (black line) and the always-varying experimental condition (red line). (Kruskal Wallis test p value $< 3 \cdot 10^{-6}$).

![Graph showing variation in generations](image)

Fig. 4. Fraction of trials in which the ability to solve the double-pole navigation problem varies every 100 generations. The black, blue, and red curves display the data obtained in the experiments carried out in the fixed, intermediate (in which the environment changes every 100 generations), and always-varying experimental conditions, respectively. Data obtained by analyzing the evolutionary lineage of the fittest evolved individual of each experiment. Each curve displays the average results of 30 replications terminated after 5000 generations.

The analysis of how the evolving agents change phylogenetically at the genetic level indicates that, also in this case, the agents evolved in the intermediate condition accumulate more variations through out generations than the agents evolved in the fixed and always varying experimental conditions (Fig. 5, Kruskal Wallis test p value $< 0.005$). To ensure a proper comparison, also in this case the analysis was performed on experiments carried out on experiments terminated after 5000 generations.
Fig. 5. Fraction of genes varied every 500 generations. The black, blue, and red curves display the data obtained in the experiments carried out in the fixed, intermediate (in which the environment changes every 100 generations), and always-varying experimental conditions, respectively. Data obtained by analyzing the evolutionary lineage of the fittest evolved individual of each experiment. Each curve displays the average results of 10 replications of each experiment.

The analysis of these data indicates a significant correlation between the performance of evolved agents and the rate with which the ancestors of these evolved agents varied at the behavioral (Spearman $r = 0.52 \ p<10^{-7}$) and genetic level (Spearman $r = 0.53 \ p<10^{-8}$).

These results confirm the hypothesis that environmental variations increase the rate at which agents change phylogenetically at the genetic and behavioral level and the hypothesis that the rate of phylogenetic variation is maximized in the intermediate condition, i.e. the condition that produces the best performing agents.

The presence of a significant correlation between the performance of the evolved agents and the rate at which their ancestors varied phylogenetically, at the genetic and behavioral level, indicates that the enhanced performance of the agents evolved in the intermediate condition is due, at least in part, to the increased rate with which the ancestors of these agents varied over generations.

**Conclusions**

Our results demonstrate that the performance of evolving agents is influenced by the variability of the environmental conditions experienced by individual agents and the variability of the environmental conditions over generations.

Agents exposed to variable environmental conditions during their lifetime achieve better performance than agents exposed to less variable environmental conditions. Moreover, agents evolved in environments that vary throughout generations achieve better performance with respect to agents evolved in environments that remain constant over generations. The two forms of variations play partially distinct roles as demonstrated by the fact that agents exposed to both types of variations outperform agents exposed to environmental variations during their lifetime only independently from the amount of experienced variation. The advantage
of environmental variations is maximized at intermediate rates, i.e. when the evolving agents are exposed to a limited number of different experimental conditions during their lifetime and when the environment change over generations at a moderate rate.

The analysis performed indicate that the fact that the agents evolved in environmental conditions that vary over generations outperform agents evolved in environmental conditions that remain stable over generations is due to the fact that the former agents change more phylogenetically than the latter agents. The fact that the agents evolved in environmental conditions that vary at a moderate rate outperform agents evolved in environmental conditions that vary at a higher rate is due to the fact that the rate of phylogenetic variation is maximized at moderate rates of environmental variations. Environmental variations create adaptive opportunities for change. However, the exploitation of these opportunities requires that the new environmental conditions remain stable enough to enable evolution to select the genetic variations that are adaptive with respect to the new conditions.

Overall the obtained results support the hypothesis of West-Eberhard (2003) that environmentally induced changes play an important role of evolution. The fact that the adaptive advantage is maximized at intermediate rate of variations also confirms the hypothesis that one of the main reasons that explain why environmentally induced changes play an important complementary role with respect to genetic changes is their tendency to persist over time (West-Eberhard, 2003).

From an evolutionary computation perspective our results demonstrate that environmental variations do not only enable to evolve solutions that are more robust or that adapt faster to environmental variations. Environmental variations, including non-directional random variations, also enable evolution to discover better solutions.

**Appendix**

1. **The agent and the environment**

![Fig. 1 The extended double-pole balancing problem.](image-url)
The cart has a mass of 1 Kg. The long pole and the short pole have a mass of 1.0 and 0.1 Kg and a length of 0.5 and 0.05 m, respectively. The cart can move along one dimension within a track of 4.8 m. It is provided with five sensors that encode the current position of the cart on the track ($x$), the current angle of the two poles ($\theta_L$ and $\theta_S$) with respect to the cart, the angle of the inclined plane ($\alpha$) and the friction coefficient ($\mu$). The motor controls the force ($F$) applied to the cart along the X axis. The goal of the agent is to move so to maintain the angle of the poles and the position of the cart within a viable range (see below).

The behavior of the agent has been simulated on the basis of equations 1-5. This is an extended version of the equations proposed by [16] for the standard problem, in which the inclination of the plane and the friction between the cart and the plane were not considered.

$$\ddot{x} = \frac{F + \mu_c M_c g + M_c g \sin \alpha + \sum_{i=1}^{n} \ddot{F}_i}{M_c + \sum_{i=1}^{n} \tilde{m}_i}$$ \hspace{1cm} (1)

$$\ddot{\theta}_i = -\frac{3}{4l_i} \left( \ddot{x} \cos \theta_i - g \sin \theta_i + \frac{\mu_p \dot{\theta}_i}{m_i l_i} \right)$$ \hspace{1cm} (2)

$$\ddot{F}_i = \mu_c \left[ \frac{3}{4} m_i g \sin^2 \theta_i - \frac{3}{4} \mu_p \frac{l_i}{m_i} \dot{\theta}_i \sin \theta_i + m_i l_i \dot{\theta}_i^2 \cos \theta_i \right] - \frac{3}{4} \left[ m_i g \sin \theta_i \cos \theta_i + \mu_p \dot{\theta}_i \cos \theta_i \right]$$ \hspace{1cm} (3)

$$\tilde{m}_i = \frac{3}{4} [\cos^2 \theta_i - \mu_c \cos \theta_i \sin \theta_i]$$ \hspace{1cm} (4)

$$\tilde{M} = m_c \cos \alpha + \sum_{i=1}^{n} m_i$$ \hspace{1cm} (5)

where $n$ is the number of poles on the cart, $g$ is the acceleration due to gravity, $m_i$ and $l_i$ are the mass and the half length of the $i^{th}$ pole, $M_c$ is the mass of the cart, $\mu_c$ is the coefficient of friction of the cart on the track, $\mu_p$ is the coefficient of friction for the $i^{th}$ hinge, $F$ is the force applied to the cart, $\ddot{F}_i$ is the effective force from the $i^{th}$ pole on the cart, $m_i$ is the effective mass of the $i^{th}$ pole, and $\tilde{M}$ is the effective mass of the cart.

2. The neural network controller of the agent

The controller of the agent is constituted by a neural network with five sensory neurons, ten internal neurons with recurrent connections, and one motor neuron. The sensory neurons are fully connected with the internal neurons, and the internal neurons are fully connected with the motor neurons and the internal neurons. The sensory neurons encode the position of the cart ($x$) expressed in meters, the angular position of the two poles ($\theta_1$ and $\theta_2$) radians, the inclination of the plane ($\alpha$) radians, and the friction coefficient of the plane/cart ($\mu_c$). The state of all sensors is normalized in the range [-0.5, 0.5]. The activation state of the motor neuron is normalized in the range [-10.0, 10.0] N and is used to set the force applied to the cart. The state of the sensors, the activation of the neural network, the force applied to the cart, and the position and velocity of the cart and of the poles are updated every 0.01 s.
The neural network's architecture is fixed. The activation state of the internal and motor neurons is updated on the basis of the logistic function. The connection weights and the biases of the network are encoded in agent’s genome and evolved. More specifically each genome consists of a vector of $171 \times 8 = 1368$ bits that encode the 160 connection weights and the 11 biases of the corresponding neural network controller.

3. The evolutionary method

To evolve the agents we used a steady state method that is widely used for the evolution of autonomous agents (e.g. Harvey, 2001; Nolfi, et al., 2016). It consists of a simple ($\mu + 1$) evolutionary strategy [18] that operates on the basis of populations formed by $\mu$ parents. During each generation, each parent generates one offspring, the parent and the offspring are evaluated, and the best $\mu$ individuals are selected as new parents. When the environmental conditions do not change with respect to the previous generation, the fitness of the parent is set equal to the fitness measured during previous evaluations and the evaluation of the parents is skipped.

The genome of the initial population is composed by a matrix of ($\mu \times 1368$) bits that are initialized randomly. Each block of 8 bits is converted into a floating-point number in the range $[-5.0, 5.0]$ that is used to set the weight of the corresponding connection or bias of the neural network controller. Offspring are generated by creating a copy of the genotype of the parent and by subjecting each bit to mutation with a $MutRate$ probability. Mutations are realized by flipping the mutated bit.

The selection pressure is regulated by adding to the fitness of individuals a value randomly selected in the range $[-$Noise, Noise$]$ with a uniform distribution [19], where Noise corresponds to the theoretical maximum fitness multiplied by the value of the Stochasticity parameter. When Stochasticity is set to 0.0 only the best $\mu$ individuals are allowed to reproduce. The higher the level of stochasticity, the higher the probability that the worse individuals reproduce is.

This method requires to set two parameters: MutRate and Stochasticity. To identify the optimal values of the parameters we carried a series of control experiments in which the two parameters were varied systematically (see the following sections). The method operates well on small populations, e.g. populations formed by 100 individuals [20].

4. Fitness function and performance measure

Agents are evolved for the ability to solve an extended version the non-markovian version of the double-pole balancing problem [14] in which the plane over which the agent is situated can be inclined of varying angles and the friction between the plane and the cart varies.

Each evolving agent is evaluated for NT trials that vary with respect to the characteristics of the plane and with respect to the initial state of the cart and the poles. More specifically, at the beginning of each trial the inclination of the plane ($\alpha$), the friction coefficient between cart and plane ($\mu_c$), the initial position of the cart on the plane ($x$), the velocity of the cart ($\dot{x}$), the angular position of the two poles ($\theta_1$ and $\theta_2$) and the angular velocity of the two poles ($\dot{\theta}_1$ and $\dot{\theta}_2$) are set to values selected within the following ranges, respectively: $[0.0, 0.2617]$, $[0.0, 0.30]$, $[-1.5, 1.5]$, $[-1.2, 1.2]$, $[-0.1047, 0.1047]$, $[-0.1350, 0.1350]$.

Trials terminate after 1000 steps or when the angular position of one of the two poles exceeded the range $[-36^\circ, 36^\circ]$ or the position of the cart exceed the range $[-2.4, 2.4]$ m.

The fitness of the agent during a trial ($f_i$) corresponds to the fraction of time steps in which the agent maintains the cart and the poles within the allowed position and orientation ranges and is calculated on the basis of the following equation:
\[ f_i = \frac{t}{1000} \]  

(6)

where \( t \) is the time step in which the cart or the pole exceeded the allowed range or 1000 in case they are maintained in the range until the end of the trial. The total fitness (F) is calculated by averaging the fitness obtained during the different trials:

\[ F = \frac{\sum_{i=1}^{NT} f_i}{NT} \]  

(7)

where NT indicates the number of trials.

The performance of an agent, i.e. the ability of an agent to solve the problem in all environmental conditions, is measured by post-evaluating agents for 1000 trials in which the characteristics of the environment and the initial state of the cart are set randomly with a uniform distribution in the ranges described above. To ensure a proper comparison the same matrix of 1000 x 8 values generated randomly is used for post-evaluating agents evolved in different experiments. Performance measures thus rate how the evolving agents are capable of generalizing their abilities to solve the double-pole problem in environmental conditions that differ from the conditions experienced during evolution.

5. Performance achieved with systematically varied parameters

Table 1, 2 and 3 reports the results obtained by systematically varying the number of trials, the mutation rate and the stochasticity level in experiments carried out in fixed environment, in always varying environments, and in environments that varied every 100 generations, respectively. The population size is always set to 100. The evolutionary process was continued for 50 millions evaluations\( (n_{individuals} \times n_{generations} \times n_{trials}) \). Each number indicates the average results of 10 replications.

| 1 Trial | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
|---------|---------|---------|---------|---------|---------|
| Mut 1%  | 78.9    | 85.6    | 85.3    | 86.3    | 75.3    |
| Mut 2%  | 68.3    | 75.3    | 81.3    | 74.2    | 70.2    |
| Mut 4%  | 55.3    | 63.2    | 83.3    | 56.8    | 55.2    |
| 50 Trials | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
| Mut 1%  | 562.9   | 578.6   | 529.8   | 613.4   | 646.4   |
| Mut 2%  | 606.9   | 628.4   | 625.7   | 645.2   | 619.6   |
| Mut 4%  | 537.2   | 504.0   | 537.3   | 506.8   | 525.2   |
| 100 Trials | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
| Mut 1%  | 625.6   | 595.9   | 522.0   | 615.9   | 634.1   |
| Mut 2%  | 658.1   | 629.9   | 623.3   | 574.5   | 606.5   |
| Mut 4%  | 479.2   | 442.6   | 506.1   | 477.0   | 463.0   |
| 150 Trials | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
| Mut 1%  | 649.9   | 639.7   | 535.9   | 644.1   | 643.9   |
| Mut 2%  | 601.6   | 605.3   | 524.2   | 629.8   | 602.9   |
| Mut 4%  | 431.8   | 461.8   | 466.1   | 462.0   | 427.0   |
| 200 Trials | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
| Mut 1%  | 641.6   | 644.2   | 634.0   | 658.7   | 636.8   |
Table 1. Performance of the best agents evolved in the fixed environmental condition obtained by systematically varying the number of evaluation trials, the mutation rate, and the level of stochasticity. Each number indicates the average results of 10 replications. The evolutionary process was continued for 50 millions evaluations. Data obtained by post-evaluating evolved agents for 1000 trials.

| Trials | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
|--------|----------|-----------|-----------|-----------|-----------|
| 1 Trial |          |           |           |           |           |
| Mut 0.5 | 91.9     | 108.6     | 102.9     | 105.8     | 90.7      |
| Mut 1%  | 119.9    | 130.2     | 118.9     | 111.8     | 103.7     |
| Mut 2%  | 103.9    | 113.2     | 109.9     | 101.8     | 94.7      |
| 15 Trials |          |           |           |           |           |
| Mut 0.5 | 581.9    | 589.2     | 592.9     | 595.8     | 590.7     |
| Mut 1%  | 669.9    | 670.2     | 678.9     | 681.8     | 673.7     |
| Mut 2%  | 623.9    | 643.2     | 639.9     | 661.8     | 662.2     |
| 25 Trials |          |           |           |           |           |
| Mut 0.5 | 594.6    | 606.3     | 608.8     | 601.6     | 588.2     |
| Mut 1%  | 696.9    | 707.6     | 709.5     | 705.3     | 698.5     |
| Mut 2%  | 631.9    | 656.3     | 658.8     | 681.6     | 668.2     |
| 50 Trials |          |           |           |           |           |
| Mut 0.5 | 570.4    | 575.7     | 598.8     | 591.6     | 588.2     |
| Mut 1%  | 665.6    | 675.6     | 698.9     | 697.8     | 675.7     |
| Mut 2%  | 593.5    | 624.7     | 628.8     | 627.3     | 626.2     |
| 200 Trials |         |           |           |           |           |
| Mut 0.5 | 521.2    | 525.3     | 535.3     | 530.5     | 529.3     |
| Mut 1%  | 604.2    | 615.2     | 617.2     | 613.5     | 611.6     |
| Mut 2%  | 503.4    | 504.2     | 514.5     | 517.7     | 513.1     |

Table 2. Performance of the best agents evolved in the always-varying environmental condition obtained by systematically varying the number of evaluation trials, the mutation rate, and the level of stochasticity. Each number indicates the average results of 10 replications. The evolutionary process was continued for 50 millions evaluations. Data obtained by post-evaluating evolved agents for 1000 trials.

| Trials | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
|--------|----------|-----------|-----------|-----------|-----------|
| 25 Trials |          |           |           |           |           |
| Mut 1%  | 760.0000 | 755.0000  | 731.3000  | 739.5000  | 725.5000  |
| Mut 2%  | 714.4000 | 704.2000  | 680.0000  | 685.9000  | 673.9000  |
| Mut 4%  | 582.4000 | 578.4000  | 571.0000  | 582.8000  | 581.5000  |
| 50 Trials |         |           |           |           |           |
| Mut 1%  |          |           |           |           |           |
| Mut 2%  |          |           |           |           |           |
| Mut 4%  |          |           |           |           |           |
Table 3. Performance of the best agents evolved in the intermediate environmental condition obtained by systematically varying the number of evaluation trials, the mutation rate, and the level of stochasticity. Each number indicates the average results of 10 replications of each experiment. Data obtained by post-evaluating evolved agents for 1000 trials.

|          | Trials | Stoch 0% | Stoch 10% | Stoch 20% | Stoch 30% | Stoch 40% |
|----------|--------|----------|-----------|-----------|-----------|-----------|
| Mut 1%   | 100    | 712.5000 | 714.8000  | 718.9000  | 686.4000  | 682.4000  |
|          | 150    | 716.4000 | 694.2000  | 687.6000  | 694.8000  | 691.5000  |
|          | 200    | 667.4000 | 676.2000  | 678.6000  | 663.8000  | 652.5000  |
| Mut 2%   | 100    | 636.0000 | 643.5000  | 624.4000  | 658.9000  | 650.8000  |
|          | 150    | 627.0000 | 634.7000  | 602.6000  | 654.8000  | 652.5000  |
|          | 200    | 480.5000 | 471.8000  | 435.8000  | 458.3000  | 455.5000  |
| Mut 4%   | 100    | 546.7000 | 543.0000  | 547.0000  | 525.9000  | 527.7000  |
|          | 150    | 520.1000 | 495.1000  | 495.1000  | 464.3000  | 458.1000  |
|          | 200    | 460.5000 | 451.8000  | 425.8000  | 448.2000  | 435.6000  |

6. Evaluating agents in systematically varied environmental conditions

To analyze how the behavior of the agents changes in systematically varied environmental conditions that covered the entire spectrum of possible variations we evaluated the agents for $3^6 = 729$ trials during which the state of the six variables that encode the characteristics of the plane and the most important initial characteristics of the cart were varied systematically. More specifically in each trial we selected one of the possible combination of the following six variables that assumed one of the three states indicated in square brackets: $\alpha [-0.1385, 0.0, 0.1385]$, $\mu_c [-0.15, 0.0, 0.15]$, $x [-0.75, 0.0, 0.75]$, $\dot{x} [-0.6, 0.0, 0.6]$, $\theta_1 [-0.05235, 0.0, 0.05235]$, $\dot{\theta}_1 [-0.0675, 0.0, 0.0675]$. The angular position and velocity of the second pole ($\theta_2 \land \dot{\theta}_2$) were always set to 0.0.

Acknowledgment

Work partially funded by CAPES through the Brazilian program science without borders.

References

Anderson C. J. R. (2013). The role of standing genetic variation in adaptation of digital organisms to a new environment. Artificial Life 13: 3–10.

Bongard J. (2011). Morphological change in machines accelerates the evolution of robust behavior. Proceedings of the National Academy of Sciences (108) 4: 1234-1239.
Florian R. V. (2007). Correct equations for the dynamics of the cart-pole system. Technical Report. Center for Cognitive and Neural Studies. Con neural, Romania.

Hare M. and Elman J.L. (1995). Learning and morphological change. Cognition (56) 1: 61-98.

Harvey I. (2001). Artificial evolution: A continuing SAGA. Evolutionary Robotics, in T. Gomi (Ed.), Lecture Notes on Complex Systems 2217: 94-109. Berlin Heidelberg: Springer Verlag.

Jacoby N. (1997). Evolutionary robotics and the radical envelope-of-noise hypothesis. Adaptive Behavior, 6 (2): 325-368.

Janssen R., Nolfi S., Haselager P and Sprinkhuizen-Kuyper I. (2016). Cyclic incrementality in competitive coevolution: Evolvability through pseudo-baldwinian switching genes. Artificial Life, (22) 3:319-351.

Jin Y. (2005). Evolutionary optimization in uncertain environments: A survey. IEEE Transactions on Evolutionary Computation, 9 (3): 303-317.

Kashtan N., Noor E., and Alon U. (2007). Varying environments can speed up evolution. PNAS, (104) 34: 13711-13716.

Kawecki T.J. (1994). Accumulation of deleterious mutations and the evolutionary cost of being a generalist. American Naturalist 144: 833–838.

Masser G., Ferrauto T., Gigliotta O. and Nolfi S. (2014). Designing adaptive humanoid robots through the FARSA open-source framework. Adaptive Behavior (22) 4: 255-265.

Mikkulainen R. and Gomez F. (1997). Incremental evolution of complex general behavior. Adaptive Behavior (5) 3-4: 317-342.

Milano N., Carvalho J. and Nolfi S. (in press). Environmental variation promotes adaptation in artificial evolution. Proceedings of the IEEE Symposium Series on Computational Intelligence (IEEE SSCI). IEEE Press.

Nolfi S. & Floreano D. (1999). Co-evolving predator and prey robots: Do 'arm races' arise in artificial evolution? Artificial Life, 4 (4), 311-335.

Nolfi S. and Floreano D. (1999). Learning and Evolution, Autonomous Robots, 7(1): 89-113

Nolfi S., Bongard J., Husband P, and Floreano D (2016). Evolutionary Robotics, in Siciliano B., Oussama Khatib (eds.), Handbook of Robotics II Edition, Berlin: Springer Verlag.

Nolfi S., Floreano D., Miglino O. & Mondada F. (1994). How to evolve autonomous robots: different approaches in evolutionary robotics. In R. Brooks and P. Maes (Eds.), Proceedings of the International Conference Artificial Life IV. Cambridge Mass: MIT Press.

O'Donnell D. R., Parigi A., Fish J.A., Dworkin I, and Wagner A.P. (2014). The roles of standing genetic variation and evolutionary history in determining the evolvability of anti-predator strategies. PloS ONE, 9 (6): e100163.

Palmer M.E. and Feldman M.W. (2011). Spatial environmental variation can select for evolvability. Evolution, 65 (8): 2345-2356.

Rechenberg I. (1973). Evolutionstrategie—Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. Stuggart: Frommann-Holzboog.

Sgro C.M., and Hoffmann A.A. (2004). Genetic correlations, tradeoffs and environmental variation. Heredity 93: 241-248.

Szafraniec K., Borts R., Korona R. (2001). Environmental stress and mutational load in diploid strains of
the yeast Saccharomyces cerevisiae. Proceedings of the National Academy of Science USA 98:1107–1112.
Vieira C., Pasyukova E.G., Zeng Z.B., Hackett J.B., Lyman R.F., and Mackay T.F.C. (2000). Genotype–
environment interaction for quantitative trait loci affecting life span in Drosophila melanogaster. Genetics
154: 213–227.
West-Eberhard M. J. (2003). Developmental Plasticity and Evolution. New York, U.S.A: Oxford University
Press
Whitley D. (2001). An overview of evolutionary algorithms: practical issues and common pitfalls. Information and Software Technology, 43: 817-831.
Wieland, A. (1991). Evolving controls for unstable systems. In Proceedings of the International Joint
Conference on Neural Networks, Volume II, Seattle, WA, USA, pp. 667–673. IEEE Press.
Zhou Z-H. and Liu X-Y. (2006). Training cost-sensitive neural networks with methods addressing the class
imbalance problem. IEEE Transactions on Knowledge and Data Engineering, (18) 1: 63-77.