Comparing lifetime learning methods for morphologically evolving robots

Fuda van Diggelen
Vrije Universiteit Amsterdam
fuda.van.diggelen@vu.nl

E. Ferrante
Vrije Universiteit Amsterdam
e.ferrante@vu.nl

A.E. Eiben
Vrije Universiteit Amsterdam
a.e.eiben@vu.nl

ABSTRACT

The joint evolution of morphologies and controllers of robots leads to a problem: Even if the parents have well-matching bodies and brains, the stochastic recombination can break this match and cause a body-brain mismatch in their offspring. This can be mitigated by having newborn robots perform a learning process that optimizes their inherited brain quickly after birth. An adequate learning method should work on all possible robot morphologies and be efficient. In this paper we apply Bayesian Optimization and Differential Evolution as learning algorithms and compare them on a test suite of different robot bodies.

CCS CONCEPTS

• **Computer systems organization** → **Evolutionary robotics**;
• **Computing methodologies** → *Machine learning*;

KEYWORDS

Evolutionary robotics, morphological evolution, lifetime learning

ACM Reference Format:
Fuda van Diggelen, E. Ferrante, and A.E. Eiben. 2021. Comparing lifetime learning methods for morphologically evolving robots. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459530

1 INTRODUCTION

Evolutionary Robotics (ER) is concerned with using evolutionary methods to automate the design process of robots [5]. In general, robots consist of two major components, a body (morphology, hardware) and a brain (controller, software). Evolution can be used to optimize both of them simultaneously. In such a system bodies and brains design are inheritable, hence the body as well as the brain of a 'child robot' is a combination of the bodies and the brains of its parents. As noted long ago [3], the stochastic nature of reproduction can lead to a *body-brain mismatch* problem: Even though parents have well-matching bodies and brains, recombination and mutation can shuffle the parental genotypes in such a way that the resulting body and brain do not fit well. This implies inferior behaviour in the offspring and can lead to suboptimal solutions in the long run. The remedy offered by the Triangle of Life framework [3] is the addition of a learning stage, where newborn robots improve their inherited brain to optimally control their inherited body.

The main goal of this paper is to investigate algorithms that could be used as learners on potentially evolving robot morphologies. Given our application, there are two important requirements to this end. First, an evolutionary process will produce a large variety of morphologies. The shapes, sizes, and complexity of the evolved robots can be very different and unpredictable. Consequently, a suitable learning algorithm needs to work well on all possible morphologies. Second, the number of trials in a learning algorithm multiplies the total computational effort required for the evolutionary process. This implies a strong preference for learning with a very low budget. Based on these requirements we selected two algorithms that performed well in recent studies:

(1) Bayesian Optimization (BO) [4]
(2) Reversible Differential Evolution (RevDE) [6]

2 RELATED WORK

Other papers have suggested slightly different approaches to tackle the body-brain mismatch problem. Cheney *et al.* [1] implemented a form of novelty protection in which 'younger' robot designs were protected from removal for several generations while evolving their brains further. This allowed them to adapt their controllers properly for the given body, which corresponds with implementing a single lifetime learning iteration. Similarly, De Carlo *et al.* [2] implemented protection in the form of speciation within their NEAT algorithm. The preservation of diversity in the population allowed new morphologies to survive, thus reducing the effects of body-brain mismatch.

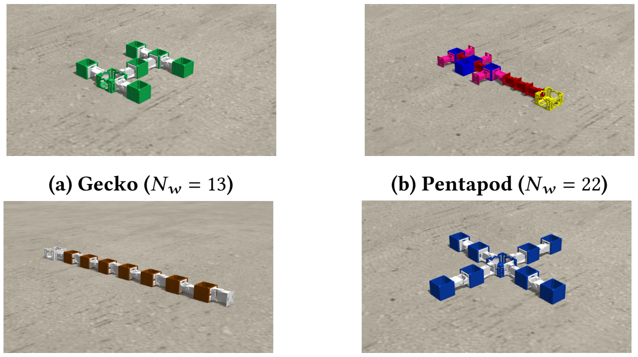
3 METHODS

Forced by time limitations we decided to use a test suite of n robots with different bodies, instead of running evolution combined with learning (that would take a very long time), where we chose $n = 4$. All robots (Figure 1) are driven by a network of interconnected CPGs and a learning method that optimizes the corresponding weights inside. The two algorithms, BO and RevDE, are run 30 times on each robot with a budget of 300 learning trials. One trial is a test period of 60 seconds in simulation and the task performance is defined as the average speed (cm/s) during a trial (displacement in centimetres divided by 60 seconds). Thus, we are addressing the task of gait learning (similar to [7]).

Bayesian Optimization is a state-of-the-art framework that had successful implementations in machine-learning, engineering, and science [4]. In short, the BO algorithm contains two main ingredients. 1) A function approximator that tries to model the fitness as a function of the search space parameters using Gaussian Processes

GECCO '21 Companion, July 10–14, 2021, Lille, France

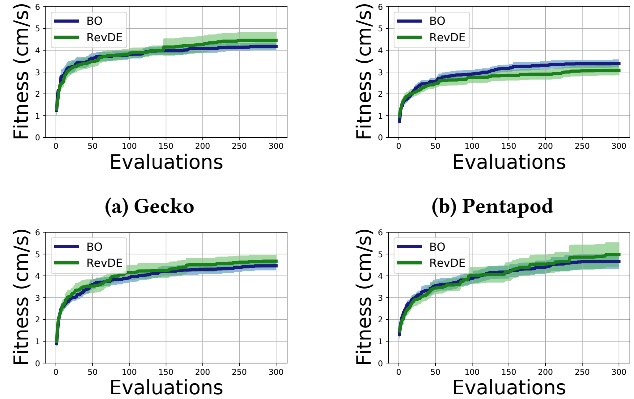
ACM ISBN 978-1-4503-8351-6/21/07.
https://doi.org/10.1145/3449726.3459530



(a) Gecko ($N_w = 13$) **(b) Pentapod ($N_w = 22$)**

(c) Snake ($N_w = 15$) **(d) Spider ($N_w = 18$)**

Figure 1: Test suite of the 4 robots used. N_w indicate the number of weights to be optimized.



(a) Gecko **(b) Pentapod**

(c) Snake **(d) Spider**

Figure 2: Average speed as a function of the number of learning trials. BO is shown in blue and RevDE in green.

($\mathcal{GP}$); 2) an acquisition function that selects the next sample using $\mathcal{GP}$ to investigate regions of high 'predicted fitness' or a high 'degree of uncertainty' (i.e. exploitation vs. exploration).

Reversible Differential Evolution (RevDE) is an Evolutionary Algorithm (EA) that can maintain a high diversity with a low population size by perturbing the current population through a special set of linearly reversible operations [6]. These perturbations impose stable exploration and exploitation properties that depend on a scaling factor. The parameters of our lifetime learning algorithms are presented in Table 1.

Table 1: Hyperparameters of both algorithms

BO	Value	Description
Initial sampling	LHS	Sampling method
Initial samples	50	Number of samples
Learning iterations	250	Number of evaluations
Kernel type	Matérn 5/2	Approximation kernel
Kernel variance	1.0	
Kernel length	0.2	
UCB alpha	3.0	Acquisition function weight
RevDE	Value	Description
λ	30	Population size
μ	10	Top-samples size
F	0.5	Scaling factor
CR	0.9	Crossover probability

4 RESULTS AND DISCUSSION

Figure 2 shows the mean fitness curves $\pm 1.96 \times SE$ ($N = 30$), with BO in blue, and RevDE in green. The corresponding areas indicate their respective 95% confidence intervals with non-overlapping regions indicating statistically significant differences.

The final speeds at the end of the learning period show differences between the robot morphologies. The Spider can learn brains –that is, weights for the CGP network– that move it with about 5 cm/s, while the Pentapod will not become faster than about 3 cm/s. Considering the top speeds obtained over all runs, the overall fastest gait was found in the Spider (10.19 cm/s), with the Gecko coming in second (8.89 cm/s).

Regarding the learning algorithms, the differences are less pronounced. The blue and green curves in the plots shown in Figure 2 overlap indicating that RevDE can perform just as sample efficient as BO. Another notable effect is the rather flat curves in the second half of the learning process. Specifically, after 150 trials the curves are about at 80 to 90 percent of final speed. This is good news from a practical perspective, suggesting that the learning budgets could be halved without losing too much performance.

The most important caveat in our results is the relatively limited test suite. Our approach to compare learning algorithms on a given set of fixed robot bodies is supported by the generally applied methodology in Machine Learning, where algorithms are compared on a number of data sets. For a more solid comparison, further research should be conducted with more than four robots. Furthermore, other possible learning methods can be tested and compared to BO and RevDE. To this end, we are inclined to investigate other EA, for instance Evolution Strategies. Using an EA for lifetime learning will result in an interesting system of nested evolution. In this system, we will have an outer evolutionary loop that optimizes both body and brain and an inner evolutionary loop to optimize the controller in 'newborn' bodies. Finally, we will investigate more complex tasks and environments including underwater.

REFERENCES

[1] Nick Cheney, Josh Bongard, Vytas SunSpiral, and Hod Lipson. 2018. Scalable co-optimization of morphology and control in embodied machines. *Journal of The Royal Society Interface* 15, 143 (2018), 20170937.

[2] Matteo De Carlo, Daan Zeeuwe, Eliseo Ferrante, Gerben Meynen, Jacintha Ellers, and A.E. Eiben. 2020. Influences of Artificial Speciation on Morphological Robot Evolution. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, IEEE Computer Society, Washington, DC, USA, 2272–2279.

[3] A E Eiben, Nicolas Bredeche, Mark Hoogendoorn, Jürgen Stradner, Jon Timmis, Andy M Tyrrell, and A Winfield. 2013. The triangle of life: Evolving robots in real-time and real-space. In *Artificial Life Conference Proceedings 13*. MIT Press, MIT Press, Cambridge, MA, 1056–1063.

[4] Gongjin Lan, Matteo De Carlo, Fuda van Diggelen, Jakub M Tomczak, Diederik M Roijers, and A E Eiben. 2020. Learning Directed Locomotion in Modular Robots with Evolvable Morphologies. *arXiv preprint arXiv:2001.07804* (2020).

[5] Stefano Nolfi, Dario Floreano, and Director Dario Floreano. 2000. *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines.* MIT press.

[6] Jakub M Tomczak, Ewelina Weglarz-Tomczak, and A.E. Eiben. 2020. Differential Evolution with Reversible Linear Transformations. *arXiv preprint arXiv:2002.02869* (2020).

[7] Fuda van Diggelen, Robert Babuska, and A.E. Eiben. 2020. The Effects of Adaptive Control on Learning Directed Locomotion. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, IEEE Computer Society, Washington, DC, USA, 2117–2124.