Embodiment can combat catastrophic forgetting.

Joshua P. Powers
University of Vermont
Burlington, VT, USA
joshua.p.powers@uvm.edu

Sam Kriegman
University of Vermont
Burlington, VT, USA
sam.kriegman@uvm.edu

Josh C. Bongard
University of Vermont
Burlington, VT, USA
josh.bongard@uvm.edu

ABSTRACT

We use an evolutionary robotics approach to demonstrate how the choice of robot morphology can affect one specific aspect of neural networks: their ability to resist catastrophic forgetting.

ACM Reference Format:

Joshua P. Powers, Sam Kriegman, and Josh C. Bongard. 2018. Embodiment can combat catastrophic forgetting.. In *GECCO '18 Companion: Genetic and Evolutionary Computation Conference Companion, July 15–19, 2018, Kyoto, Japan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3205651.3205781

1 INTRODUCTION

It has been shown elsewhere that embodiment can influence the positive or negative aspects of neural networks. For instance, work in morphological computation has shown that a good choice of morphology can allow for simplified neural networks (e.g. [3]). Morphology may also render a robot more robust to external environmental perturbation [1] or internal changes to the neural controller [4]. In this work we introduce a heretofore unexplored aspect of the interaction between body plan and neural control of embodied agents: how the choice of body plan may render the neural controller more or less resistant to catastrophic forgetting.

2 METHODS

Three types of robots were used in this study (Fig. 1). All are variations on a standard radially symmetric quadrupedal form, differing in their use of legs and/or wheels. The robots contain a touch sensor in each leg that detect contact with the ground, and an on-board light sensor that detects light according to the inverse square law. Robots are controlled by a fully-connected neural net with 5 input neurons (one for each sensor), 8 output neurons (one for each motor), and no hidden layers.

Robots are evolved to perform phototaxis in multiple training environments, with different light source locations. The performance of a robot in a single environment is the value of the robot's light sensor at the end of an evaluation period. We combine performance across multiple environments in three different ways:

$$\text{Sum} \doteq \sum_{i=1}^{n} p_i, \qquad \text{Product} \doteq \prod_{i=1}^{n} p_i, \qquad \text{Min} \doteq \min_{i=1}^{n} p_i. \qquad (1)$$

where p_i is the individual's performance in environment $i \in (1, n)$.

GECCO '18 Companion, July 15–19, 2018, Kyoto, Japan

ACM ISBN 978-1-4503-5764-7/18/07.
https://doi.org/10.1145/3205651.3205781

2.1 Measuring catastrophic forgetting.

Catastrophic forgetting, in its simplest formulation, occurs when an improvement in one environment incurs reduced performance in one or more other environments [2]. In an evolutionary setting, catastrophic forgetting can be measured at the highest temporal resolution by considering mutations: the change in performance between a parent and child for each environment experienced by both agents (Fig. 2).

We define a function D on an individual such that it returns a vector $[d_1, d_2, \ldots, d_n]$, where d_i is the distance from the light source at the end of simulation in environment i. As shown in Fig. 2, fitness and distance are inversely correlated therefore we record change in distance after every *successful* mutation as:

$$\Delta D = -\left[\, D(\text{child}) - D(\text{parent})\,\right]. \qquad (2)$$

We negate this difference, so that a positive increase in fitness in an environment causes an increase in the corresponding component of ΔD. We did not record deleterious or stagnant mutations.

3 RESULTS

In Fig. 3, as we change the morphology from legged to whegged, the robots demonstrate increased evolvability. Thus the fitness landscape allows for larger jumps towards the optima. This includes those jumps that avoid catastrophic forgetting altogether: mutations visualized by points in Fig. 3 that lie in the upper right quadrant.

In conjunction, as we change the fitness function from sum to min, we see the spread of points in Fig. 3 condense toward the origin. When combined with the whegged robot, we see a reduction catastrophic forgetting. It appears that it is the combination of correct fitness function (min) with the correct morphology (whegged) that resists catastrophic forgetting: changes in morphology or fitness alone are not sufficient. We hypothesize that this greater resistance to catastrophic forgetting is what enables the whegged robot, under the min fitness function, to achieve higher fitness within environments and consistent fitness across environments.

One objection to this hypothesis could attribute the performance of the whegged robot to the increased speed allowed for by wheels. We do not feel that this is valid for two reasons: the wheeled robot also has wheels and does not achieve the same level of performance, and the evaluation time of a simulation was set such that all morphologies are able to reach the light source before the end of simulation. Indeed we observed that all robots reached and waited at the light source when trained against a single environment.

In observing the behavior of the robots we noticed a pattern among whegged robots that could account for their resistance to forgetting. Whegged robots move very rapidly in a circular pattern during the initial time steps of a simulation which may allow them to 'sidestep' catastrophic forgetting by rapidly turning unfamiliar environments into familiar ones. An example is shown in Fig. 4:

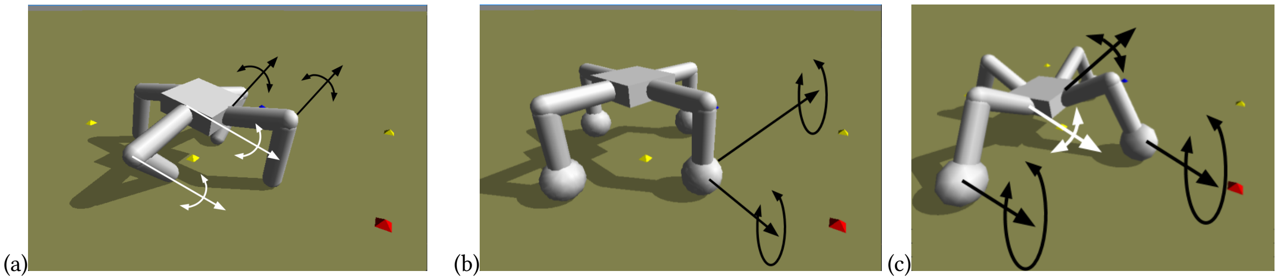

Figure 1: Three classes of phototaxic robots—legged (a), wheeled (b), and whegged (c)—and their environments were simulated using Pyrosim (ccappelle.github.io/pyrosim). Each robot has eight degrees of freedom, as depicted by the black and white arrows which indicate the axis (straight) and direction (curved) of rotation for a particular hinge-joint (a, c) or wheel (b, c). Video of all three robot types can be seen at youtu.be/yY7Vi7fw7Ik .

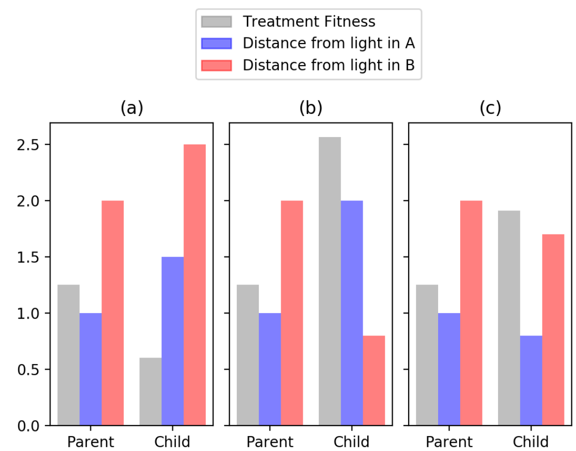

Figure 2: Example mutations that (a) are deleterious, (b) result in catastrophic forgetting, and (c) avoid catastrophic forgetting. The smaller the distance from the light source (blue, red), the higher the fitness.

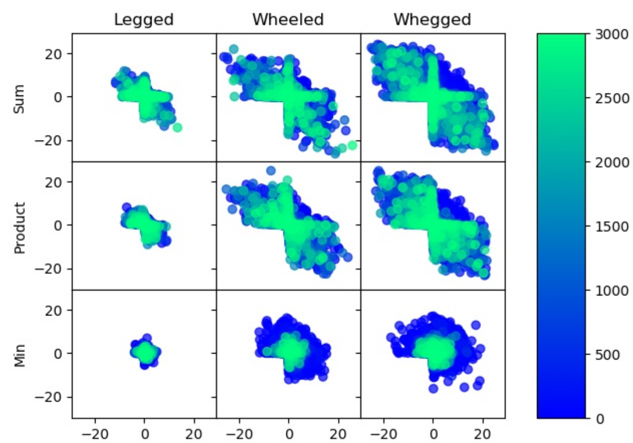

Figure 3: Change in performance (ΔD, as defined in Eq.2), in two environments, for the three robots (Fig. 1) and three fitness metrics (Eq. 1), colored by the generation of the mutation. Dots in the upper-right quadrants of each robot-fitness cell represent beneficial changes in both environments; these mutations avoided catastrophic forgetting. Dots in the upper-left and lower-right quadrants of each cell contain mutations that were beneficial in one environment but deleterious in the other; these changes caused catastrophic forgetting. We did not record mutations that were deleterious in both environments (lower-left quadrants).

the rotationally symmetric trajectories of the blue whegged robot indicates it has recognized two versions of the same environment. The red legged robot does not: its two trajectories are different, and take longer to diverge. The wheeled and legged robot both seem to have much more difficulty in turning.

This implies that the very phenomenon of catastrophic forgetting itself may be to some degree a false problem arising from studies using non-embodied systems: Since such systems do not have control over their input, they cannot align objects of interest in different training instances and thus reduce catastrophic forgetting.

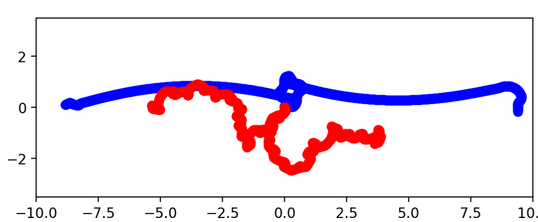

Figure 4: A tracing of a typical whegged robot (blue) and legged robot (red) trained in two environments. The light source is first placed at (9, 0), and then at (-9, 0). Video is available at youtu.be/uWy33A5HZGM.

REFERENCES

[1] Josh Bongard. 2011. Morphological change in machines accelerates the evolution of robust behavior. *Proceedings of the National Academy of Sciences* 108, 4 (2011), 1234–1239.

[2] Robert M French. 1999. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences* 3, 4 (1999), 128–135.

[3] Helmut Hauser, Auke J Ijspeert, Rudolf M Füchslin, Rolf Pfeifer, and Wolfgang Maass. 2011. Towards a theoretical foundation for morphological computation with compliant bodies. *Biological cybernetics* 105, 5-6 (2011), 355–370.

[4] Sam Kriegman, Nick Cheney, and Josh Bongard. 2017. How morphological development can guide evolution. *arXiv preprint arXiv:1711.07387* (2017).