

Combating catastrophic forgetting with developmental compression.

Shawn L.E. Beaulieu
University of Vermont
Burlington, VT, USA
shawn.beaulieu@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

Generally intelligent agents exhibit successful behavior across problems in several settings. Endemic in approaches to realize such intelligence in machines is catastrophic forgetting: sequential learning corrupts knowledge obtained earlier in the sequence, or tasks antagonistically compete for system resources. Methods for obviating catastrophic forgetting have sought to identify and preserve features of the system necessary to solve one problem when learning to solve another, or to enforce modularity such that minimally overlapping sub-functions contain task specific knowledge. While successful, both approaches scale poorly because they require larger architectures as the number of training instances grows, causing different parts of the system to specialize for separate subsets of the data. Here we present a method for addressing catastrophic forgetting called developmental compression. It exploits the mild impacts of developmental mutations to lessen adverse changes to previously-evolved capabilities and ‘compresses’ specialized neural networks into a generalized one. In the absence of domain knowledge, developmental compression produces systems that avoid overt specialization, alleviating the need to engineer a bespoke system for every task permutation and suggesting better scalability than existing approaches. We validate this method on a robot control problem and hope to extend this approach to other machine learning domains in the future.

CCS CONCEPTS

• **Computing methodologies** → **Mobile agents; Neural networks;**

KEYWORDS

Catastrophic forgetting.

ACM Reference Format:

Shawn L.E. Beaulieu, Sam Kriegman, and Josh C. Bongard. 2018. Combating catastrophic forgetting with developmental compression.. In *GECCO '18: Genetic and Evolutionary Computation Conference, July 15–19, 2018, Kyoto, Japan*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3205455.3205615>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

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

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5618-3/18/07...\$15.00

<https://doi.org/10.1145/3205455.3205615>

1 INTRODUCTION

The method introduced here resists catastrophic forgetting on a simple task by exploiting the evolution of development. We thus discuss below relevant work from these two domains.

1.1 Catastrophic forgetting.

It has been known since the early days of machine learning research that catastrophic interference [28], now more commonly known as catastrophic forgetting [15, 16], is a major challenge to training neural networks effectively. Even for the most common forms of network training such as the backpropagation of error, there is no guarantee that reducing the network’s error on the current training sample does not increase error on the other samples.

For these reasons, much effort has been expended to address this problem. One family of solutions involves constructing modular networks [4, 10, 13, 19, 27, 32] in which different modules deal with different subsets of the training set. In such networks, changes to one module may result in improved performance for the training subset associated with that module without disrupting performance on other subsets. Such modularity has indeed been demonstrated to minimize catastrophic forgetting [11, 14, 25, 25, 31, 31]. Related to this concept of modularity are networks in which some subsets of the network that have a large impact on the current training set are made resistant to change during subsequent training [20, 39]. The remaining parts of the network, which are rendered more adaptive, are able to fit to new training instances without disrupting behavior on previous instances.

The drawback of these approaches however is that network size tends to increase with the amount of training data, because new modules must be implicitly or explicitly added for new training data. In the work presented here, we introduce a method for combating catastrophic forgetting by drawing on the idea of compression: the size of the learner expands when new training instances are encountered, but the learner is then compressed back to its original size in a gradual manner through evolutionary canalization.

Besides modularity, another guard against catastrophic forgetting is to reduce the magnitude of behavioral impact after some change is made during training. The intuition here is that small changes to network behavior may increase the likelihood of local improvements for new training instances while minimizing or nullifying performance decreases on the previous training set.

In evolutionary methods, one way of reducing behavioral impacts is to dynamically tune, during evolution, mutation rates [6] and/or crossover events [37]. A recent approach demonstrated for neuroevolution is to dynamically tune individual synaptic weights proportionally to their impact on the network’s behavior [26]. Similarly, in the genetic programming community, semantic variation

operators have been reported [3, 30, 36, 38]. These operators take into account the semantics of subtrees or individual tree nodes, and attempt to replace them with new genetic material that exhibits similar semantics.

Here, we introduce a new method that exploits development—the fact that some agents may change their internal structure over their lifetimes—to reduce the behavioral impact of evolutionary perturbations. This reduced behavioral impact, which results in late onset behavioral change, then facilitates the compression of networks specialized in different training subsets back down to one network. Compression is mediated by conditions that allow for the Baldwin effect, in which beneficial behaviors that manifest late in life occur increasingly earlier in successive generations.

1.2 The evolution of development.

The evolution of development, or ‘evo-devo’ has received increasing scrutiny in AI and robotics research. [17] showed that combining learning and evolution can smooth the fitness landscape and thus facilitate evolution. Since learning can be considered a form of development [21], that experiment can thus be considered the initial investigation into how combining evolutionary change with lifetime change can facilitate evolution. In evolutionary robotics, several evo-devo approaches were investigated [8, 29], some of which involved considering morphological change in the robot over its lifetime [1, 2, 9]. This work was followed by the introduction of compositional pattern producing networks (CPPNs; [34]), which are described as capturing one important aspect of development: its ability to bias evolutionary search toward solutions with regular internal structure.

Despite their clear ability to increase evolvability in many systems, CPPNs do not model one important aspect of biological development: change over the lifetime of the learner. The very fact that agents may change over time provides a unique opportunity to increase their evolvability. If a mutation alters the developmental program of an agent, the behavioral impact of that mutation may not manifest until later in the descendant agent’s lifetime. If the fitness of an agent is integrated over its lifetime, it follows that the overall behavioral impact of a developmental mutation is likely to be less than that of a non-developmental mutation. This follows because a non-developmental mutation takes immediate effect at the beginning of an agent’s lifetime.

Last, the likelihood that a mutation is beneficial is inversely proportional to its magnitude of behavioral change. Thus, a developmental mutation which has a late onset effect, and so less behavioral impact, is more likely to be beneficial than a non-developmental mutation. In effect, one can increase the likelihood of beneficial mutations, and thus overall evolvability, simply by including developmental change in populations of evolving agents. That development unlocks broader and more nuanced ranges of behavioral change compared to non-developmental programs has recently been documented [23, 24].

In the next section we describe how evo-devo methods can be used to guard against catastrophic forgetting by gently mutating learners in such a way that improves their ability on one training instance without damaging performance on other training instances.

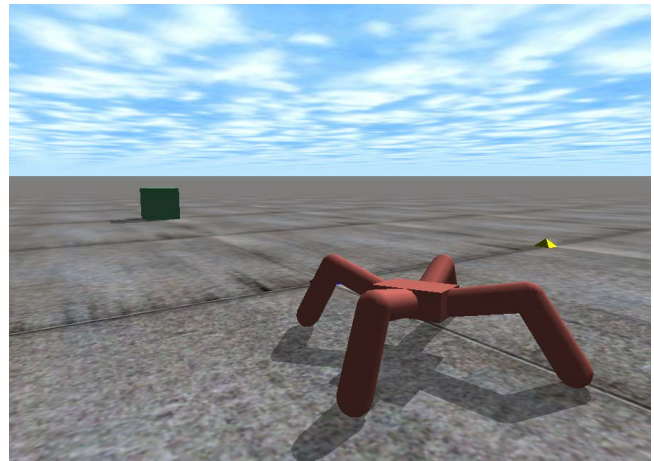


Figure 1: A screen shot of environment *A*, in which a light-emitting block is placed 30 body lengths *in front of* the robot’s starting position. In environment *B*, the block is placed 30 body lengths *behind* the robot. The robot is to perform phototaxis with an on-board sensor that detects light intensity according to the inverse square law.

2 METHODS

We chose to test developmental compression on a robot control problem for three reasons. First, the continuous control of legged robots is a notoriously difficult machine learning problem. Second, by dint of the problem’s inherent difficulty, if a robot is exposed to multiple disparate environments, catastrophic forgetting is likely to occur. That is, improvements to the robot’s ability in one training environment will likely disrupt previously-evolved capabilities in other environments. Third, a robot’s behavior at one time step has a cascading effect on its behavior in future time steps; so, if fitness is integrated over multiple time steps, and developmental change in control policy is allowed, developmental mutations that manifest late in the life of individuals will impact overall behavior less. This is in contrast with non-developmental mutations which, by definition, take effect at the first time step of a robot’s evaluation. This difference is exploited by developmental compression, as explained in 2.5.

2.1 The robot.

The robot and its environment were simulated using Pyrosim¹ [22], an open source Python wrapper built atop the Open Dynamics Engine (ODE) physics engine [33] (Fig. 1).

The architecture, or morphology, of the robot is generic by design. It was not selected to optimally solve the task it was given and, indeed, was chosen before the task environment was fully specified. It is characterized by a rectangular abdomen, attached to which are four legs, each composed of an upper and lower cylindrical object. The knee and the hip joint of each leg contain a rotational hinge joint with one-degree-of-freedom. Each hinge joint can flex inward or extend outward by up to 90 degrees away from its initial angle. The orientations of the hip and knee joints are set such that

¹ccappelle.github.io/pyrosim

each leg moves within the plane defined by its upper and lower leg components.

Inside each lower leg is a touch sensor neuron, which at every time step detects when the lower leg to which it belongs makes contact with the environment. It takes on either -1 (no contact) or $+1$ (contact). Apart from the four touch sensors, a light sensor (whose output is a floating-point number) is embedded in the abdomen of the robot. Due to the inverse square law of light propagation, the light sensor is set to $\ell = 1/d^2$, where d is the Euclidean distance between the light sensor and the light source. Motor neurons innervate each of the eight joints of the robot and enable it to move.

2.2 The controller.

In total, there are five sensor neurons fully connected to eight motor neurons such that each sensor neuron feeds into every motor neuron. No hidden neurons were employed.

How the set of sensor neurons connects to, and communicates with, the motor neurons is determined by the robot's neural controller. The edges of the network are represented by a weighted adjacency matrix that is optimized by a direct encoding evolutionary algorithm. Synaptic weights are constrained to $[-1, +1]$.

Motor neurons are updated according to

$$m_i^{(t)} = \tanh \left[m_i^{(t-1)} + \tau_i \sum_{j=1}^{J=5} w_{ji} s_j^{(t)} \right] \quad (1)$$

where $m_i^{(t)}$ denotes the value of the i th motor neuron at the current time step, $m_i^{(t-1)}$ is a momentum term that guards against 'jitter' (high-speed and continuous reversals in the angular velocity of a joint), τ_i is a time constant that can strengthen or weaken the influence of sensation on the i th motor neuron relative to its momentum, and w_{ji} is the weight of the synapse connecting the j th sensor neuron to the i th motor neuron. In order to ensure that random controllers produce diverse yet not overly-energetic motion, all τ_i were set 0.3 via empirical investigation.

2.3 The task environment.

When the simulation begins the robot is placed at a location that is identically equal to the origin of a smooth two-dimensional plane. Two tasks are presented sequentially to the robot. For the first, a light-emitting box is placed 30 body lengths away from the origin in the *positive* y-direction (that is, *into* the screen); for the second, the box is placed 30 body lengths from the origin in the *negative* y-direction. Success is defined as the ability to walk toward the box in both environments in the allotted time (1000 ts).

2.4 The fitness function.

As in the real world, the decay of a light signal's strength is proportional to the inverse squared distance from the source location. Consequently, fitness is taken to be the mean light sensor value experienced by the robot during the simulation. Integrating over time imposes a soft constraint on long peregrinations that obtain good performance only at the end of the evaluation. This selects for behavior of high proficiency throughout the life of an individual. Fitness increases both with proximity to the light source and with increasingly economical behavior that results in quicker travel.

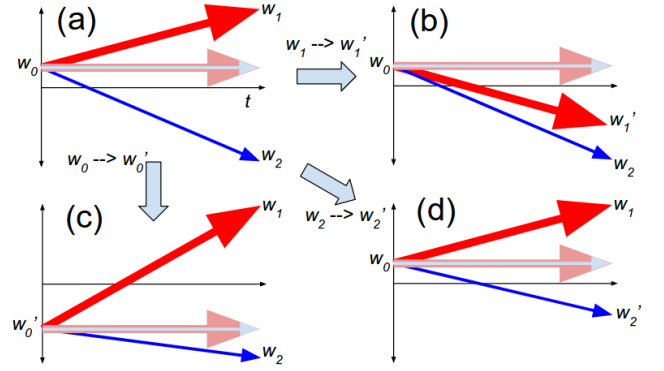


Figure 2: A conceptual visualization of the developmental compression algorithm on a single synaptic weight. (a) Evolution tries to locate a compressed representation of target weights w_1 and w_2 . (b) Mutation affects w_1 . (c) Mutation affects base w_0 . (d) Mutation affects w_2 . Horizontal axes represent developmental time.

Formally, fitness is defined as:

$$f = \sum_{e=1}^E \frac{1}{T} \sum_{t=0}^T p_t^{(e)} \quad (2)$$

where $p_t^{(e)}$ denotes light sensor's value in environment e at time t

In the experimental treatment each individual is evaluated 4 times, whereas the control treatment has only two evaluations per individual. To ensure equality of resources, the number of generations in the experimental treatment is half the number used in the control treatment.

2.5 The compression algorithm.

The intuition for developmental compression is that a given neural controller contains a number of sub-controllers which evolve to specialized, successful behavior in separate training environments. Then, mild developmental mutations gradually 'compress' the specialized sub-controllers into a single base controller, W_0 .

More specifically, in two task environments, the networks we seek to compress are those that comprise an individual's genetic tensor ($5 \times 8 \times 3$), represented by three distinct weight matrices (5×8). One such matrix serves as the base controller, while the other two are the target specialist controllers toward which the base linearly develops in the corresponding task environment. Concretely, the zeroth sheet of the tensor, $W_D^{(0)}$, is the base controller while $W_D^{(e)}$ is the target controller for the e^{th} environment.

Development is the process by which compression is enforced. It works by moving the weights of the base state toward a given target state as the simulation proceeds. Under the developmental treatment for the first task, the robot begins the simulation with controller $W_D^{(0)}$ and ends the simulation with controller $W_D^{(1)}$ under a development schedule that's linear in time.

Generically, for a target state belonging to task environment e , every element of $W_D^{(0)}$ moves toward $W_D^{(e)}$ by:

$$w_{ji}^{(0)} \rightarrow w_{ji}^{(e)} = \alpha_t w_{ji}^{(0)} + (1 - \alpha_t) w_{ji}^{(e)} \quad (3)$$

where α is a weight ranging from 1 to 0 that linearly decays with t that moves from 0 to $T - 1$ over T time steps in the evaluation. Thus, the weight of synapse w_{ji} at time t becomes

$$w_{jit} = \frac{(T - 1 - t)w_{ji}^{(0)} + t w_{ji}^{(e)}}{T - 1} \quad (4)$$

As the evaluation proceeds, α will dampen the base element and amplify the target element (line 32 in Algorithm 1).

For every task environment, a single agent is evaluated twice: once developmentally and once non-developmentally (lines 23 and 24 in Algorithm 1, respectively). The fitness they obtain in each treatment is then added to their total fitness.

For mutation, a single synaptic weight is chosen at random within the $S \times M \times (e + 1)$ weight matrix W : w_{jik} . The new weight of w_{jik} is sampled from a Gaussian distribution whose mean and standard deviation depend on the selected synapse's prior value:

$$w_{jik}^{(new)} \sim \mathcal{N}\left(\mu = w_{jik}^{(old)}, \sigma = |w_{jik}^{(old)}|\right) \quad (5)$$

If a mutation carries a weight above 1, it is set to 1; if a mutation carries a weight below -1, it is set to -1 (Fig.2).

2.6 The control algorithm.

The control algorithm optimizes performance by summing fitness scores of non-developmental agents across all task environments. That is, individuals are evaluated once in each environment and the fitness they obtain is added to their respective total fitness, F_p , which is reported once all environments in E have been encountered.

Mutation here mirrors Equation 5, except that one synaptic weight is chosen at random from the 5×8 weight matrix, not the genetic tensor.

2.7 Random search algorithm.

As a sanity check, the compression and control algorithms were compared with a random search algorithm, in which heredity is discarded. In other words, random search is the control treatment with reproduction removed. However, the ways in which individuals are mutated and evaluated are unchanged.

The total number of robot evaluations across all three algorithms is equalized to ensure a fair comparison.

3 RESULTS

We performed 100 independent runs of the experimental treatment (750 generations per run, Sect. 2.5), the control treatment (1500 generations, Sect. 2.6), and random search (1500 generations, Sect. 2.7). For each run, the fittest individual across all generations was extracted: that is, the overall run champion, as defined by Equation 2.

The results presented in Figures 4 and 3 (generational champions) and 1 (overall champions) strictly pertain to non-developmental performance. In the DC treatment, run champions are those whose fitness is highest across all $2E$ evaluations, but we report only the

Algorithm 1 Developmental Compression vs. Control

```

1:  $E \leftarrow$  The Task Environments ( $1 \times$  number of envs)
2:  $W_S \leftarrow$  Static Genome ( $5 \times 8 \times 1$ )
3:  $W_D \leftarrow$  Developmental Genome ( $5 \times 8 \times$  number of envs)
4:  $G \leftarrow$  Number of Generations
5:
6: if controlTreatment then
7:   for  $g=1:G$  do
8:     CONTROL( $W_S, E$ )
9: if experimentalTreatment then
10:  for  $g=1:G/2$  do
11:    DEVELOPMENTAL_COMPRESSION( $W_D, E$ )
12:
13: procedure CONTROL( $W_S, E$ )
14:   for each individual in the population do
15:     Fitness = 0
16:     for  $e$  in  $E$  do
17:       Fitness += SIM( $e$ , base =  $W_S$ , target =  $W_S$ )
18:
19: procedure DEVELOPMENTAL_COMPRESSION( $W_D, E$ )
20:  for each individual in the population do
21:    Fitness = 0
22:    for  $e$  in  $E$  do
23:      Fitness += SIM( $e$ , base =  $W_D^{(0)}$ , target =  $W_D^{(e)}$ )
24:      Fitness += SIM( $e$ , base =  $W_D^{(0)}$ , target =  $W_D^{(0)}$ )
25:
26: procedure SIM( $e$ , base, target)
27:   $p_e = 0$  ▷ performance (light intensity)
28:  for  $t=1:T$  do ▷ time steps
29:    for  $s=1:S$  do ▷ sensor neurons
30:      for  $m=1:M$  do ▷ motor neurons
31:
32:         $w_{smt} = ((T - 1 - t)w_{sm}^{(base)} + t w_{sm}^{(target)}) / (T - 1)$ 
33:        Perform in env  $e$  for time step  $t$  using  $W_t$ .
34:         $p^{(e)} += p_t^{(e)}$ 
35:  return  $p_t^{(e)}$ 

```

non-developmental performance of those run champions (line 24 in Algorithm 1) for fairness of comparison.

Fig. 3a reports the median fitness for the generational champions' best environment for all three treatments. This figure suggests that the control treatment is significantly outperforming both the experimental treatment and random search. However, Fig. 4a demonstrates that this conclusion is premature. This figure shows the median fitness of generational run champions in their worst environment. Catastrophic interference is responsible for the discrepancy between the quantity being optimized by the evolutionary algorithm (mean fitness) and the quantity we implicitly seek to optimize (general proficiency).

The ability of robots in the experimental treatment to continue improving in their worst environment is reported in Table 1. For the case of two environments ($E = 2$), we assessed statistical significance for the median minimum fitness of each treatment's overall run champion with the Mann-Whitney U test and Bonferroni

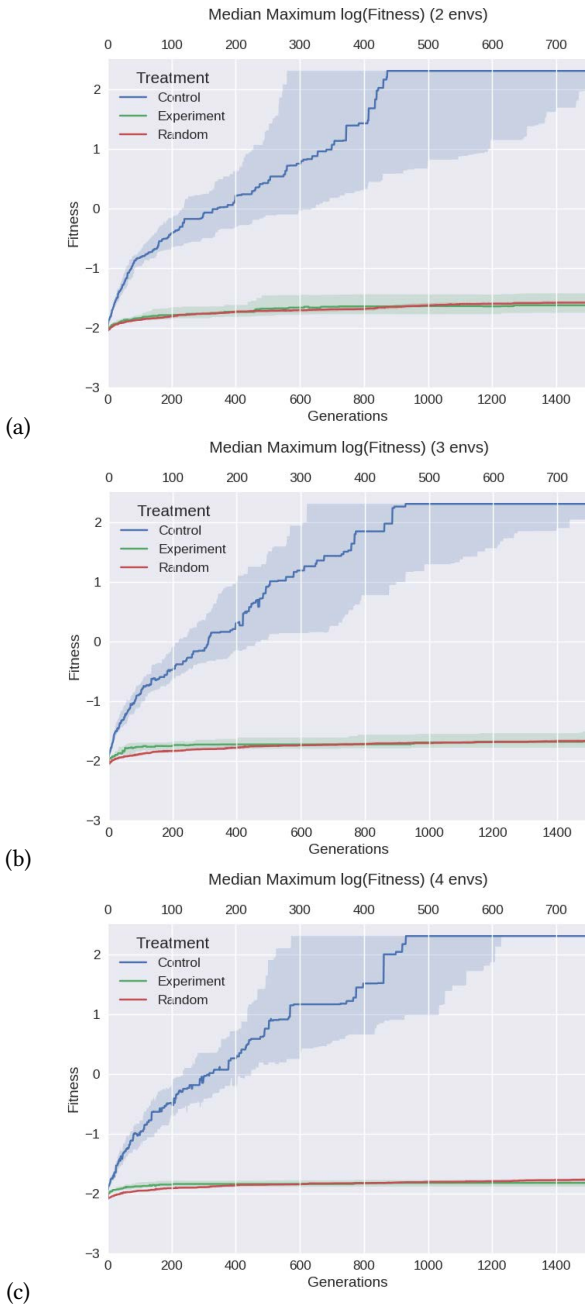


Figure 3: Median Maximum Fitness. The median score across all runs for the maximum fitness value of generational run champions (95 percent confidence intervals). (A) two task environments (E=2); (B) three task environments (E=3); (C) four task environments (E=4). Y-axis is log-scale. For reference, with two environments the experimental treatment has median=0.19, std=0.202. The top axis indicates the generations seen by developmental compression

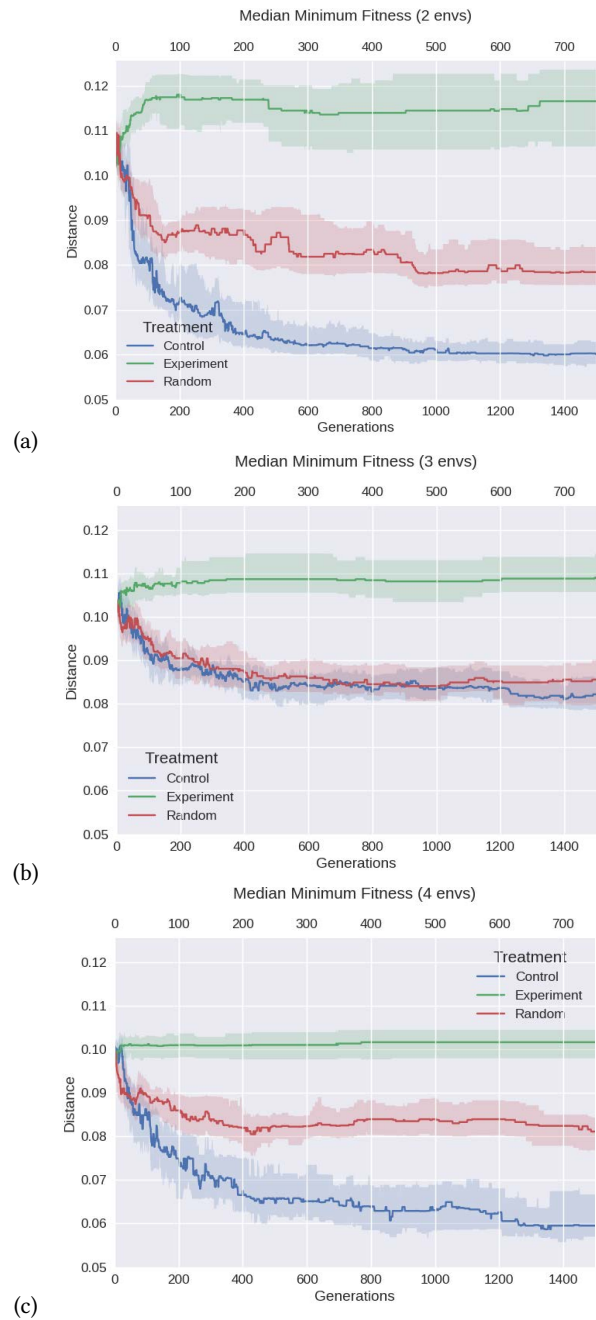


Figure 4: Median Minimum Fitness. The median score across all runs for the minimum fitness value of generational run champions (95 percent confidence intervals). (A) two task environments (E=2); (B) three task environments (E=3); (C) four task environments (E=4). This figure highlights catastrophic forgetting in the random and control methods. Control is more adversely affected by forgetting due to overspecialization. The experimental method avoids the negative impact of task antagonism.

Table 1: Median minimum fitness of run champions.

	Random	Control	Compression
E=2	0.0826 ± 0.0186	0.0781 ± 0.0372	0.1103 ± 0.0240
E=3	0.0870 ± 0.0188	0.0847 ± 0.0229	0.1094 ± 0.0188
E=4	0.0823 ± 0.0142	0.0704 ± 0.023	0.1016 ± 0.014

correction across three possible comparisons (compression vs. control, compression vs. random, random vs. control). We found that developmental compression significantly outperformed both the control ($P < 0.0001$) and random treatments ($P < 0.0001$). Results for the median minimum fitnesses of overall run champions are shown in table 1.

Minimum fitness is an important metric in this context because if it decreases over evolutionary time it reveals the degree to which an individual narrowly specializes to one task environment. This trend can also arise in the developmental treatment if the base genome evolves to be overly dependent on their scaffolds, sacrificing performance without them. A similar problem can afflict the non-developmental treatments in increasingly many environments, where a good strategy is to avoid costs to overall fitness by doing the same thing in every environment: stand still.

The extent to which development compresses the weight matrices in an agent’s controller is shown in 5. For every environment, the mean Euclidean distance between the base state and the target state for the generational run champions is computed over the course of evolution. The mean distance is then averaged over the number of environments and reported with a 95 percent confidence interval. At a glance, compressibility appears to correlate with the median minimum fitness plots in Figure 4, which suggests that DC’s ability to avoid catastrophic forgetting depends on the quality of compression.

By observing where and how individuals fail to achieve optimal performance where they otherwise should, we can glimpse into the logic under which they operate. For example, in the control, a high maximum fitness (Fig.3) does not seem to correlate with the ability of phototaxis (Fig.4), but rather with the ability to walk remarkably well in one direction. These results suggest that developmental compression was partly able to overcome tempting local optima like specialization and inaction, while the control and random treatments were not.

4 DISCUSSION

Developmental compression appears to limit specialization, which afflicts both random search and the naive approach of averaging non-developmental performance across all environments (Fig. 4). These latter two treatments likely specialize because they suffer from catastrophic forgetting: seemingly redundant improvements in one’s best environment recur because such improvements outweigh the penalty to overall fitness incurred by improving one’s worst environments. Only robots within the developmental compression paradigm exhibit continued improvement in all environments over evolutionary time (Figs. 3a and 4a). This is may be ascribed to (at least) two factors.

First, mutations within developmental compression tend to have more mild behavioral effects. This is illustrated in Fig. 2. A mutation

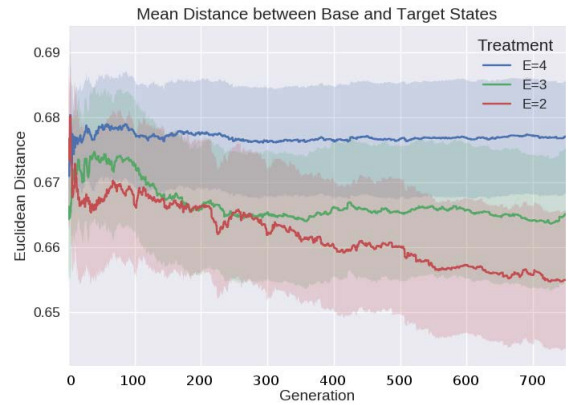


Figure 5: Compression over time. Mean distance during evolution between the base state and target states in the corresponding environments. Values computed using generational run champions across all runs for each treatment (E=2:4) with 95 percent confidence intervals.

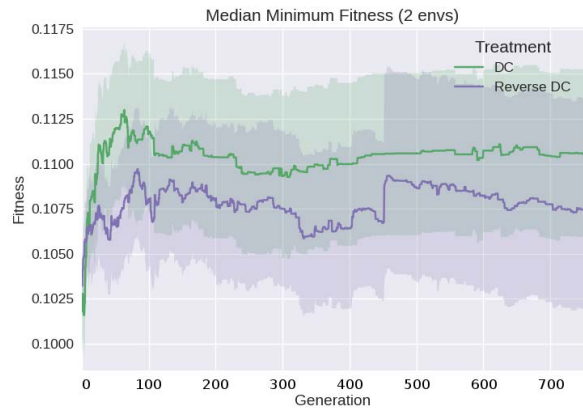


Figure 6: DC vs reverse DC. Comparing the experimental treatment (green) and a reversed version (purple) with 95 percent confidence intervals. See Discussion for details.

that strikes a feature in the base matrix, w_0 , (Fig. 2c) impinges on behavior in all E environments, but this effect is attenuated as the base matrix develops toward its scaffolding. A mutation that strikes a target matrix (Fig. 2b) only modulates behavior in one of the E environments. This is likely to increase the likelihood that the controller may improve in one environment without adversely impacting its performance in the other $E - 1$ environments.

The second reason that developmentally compressed controllers outperform those from the random and control treatments at local task generalization is that evolution gradually ignores target scaffolding over evolutionary time. This can be seen in Fig. 5 where the distance between the base and target matrices decreases with time. This means that the three matrices within a controller are gradually

approaching one another, and the path traversed by development becomes shorter. This is canalization.

The success of developmental compression in this domain is even more surprising given that its fitness landscape is twice as large as the control treatment. However, work on extra-dimensional bypasses [5] has shown that at high dimensions, if an optimization procedure can smooth the fitness space by bridging local optima in lower dimensional spaces, evolvability can be increased. Thus developmental compression can be viewed as a way to safely introduce extra-dimensional bypasses, which may militate against catastrophic forgetting.

4.1 Scalability.

Despite its success in two environments, developmental compression weakened in its ability to obtain generalists as the problem was scaled up to three and four environments (Figs. 3b,c and 4b,c). Indeed, as Fig. 5 illustrates, DC also fails to compress increasingly many weight matrices as more problems are introduced. At present it is difficult to say why this occurs. However, the most likely culprit is a weak evolutionary algorithm and sparse neural controller. An extremely simple genetic algorithm was employed here—one that doesn't attend to diversity, or other salient factors, and thus fails to obtain good and compressible scaffolds on complex problems. In future work we plan to investigate how DC can be strengthened with better evolutionary constraints, and be made to synergize with related algorithms.

Interestingly, the behavioral impacts of mutations become increasingly mild in DC as the number of task environments increases. For two environments ($E = 2$), a mutation to a given target matrix impacts behavior in only one of environment; for three environments ($E = 3$), a mutation to a target matrix impacts behavior in only one of three environments; and so on, with mutation becoming increasingly diluted. While DC might dull the behavioral consequences of any single mutation, there's no guarantee that this will result in good generalists. Instead, this may slow the pace of progress. In future work we will investigate how to address this by using adaptive mutation rates, rather than one that is fixed.

4.2 Different forms of development.

In this work, a simple process of development was employed: robots begin with a single controller and move toward E controllers in the corresponding environments. As development proceeds, agents exhibit increasingly unique behaviors. However, different developmental trajectories—ones that aren't strictly linear—might be even better at avoiding catastrophic forgetting.

In this vein, we conducted another set of 100 independent runs of developmental compression in each of two, three, and four environments. But in this case developmental interpolation was reversed: for developmental evaluations, E base states developed linearly toward a *single* target state (line 23 of Algorithm 1). Curiously, this reversed variation did not appear to avoid catastrophic forgetting as well as the original experimental method (Fig. 6). The suggests that the order of compression is important for success in this domain. How compression varies with different developmental programs will be the subject of future investigations. In Fig. 6 we only report performance for two environments, but the gap

between DC and reverse DC widened more strikingly in three and four environments.

Another potential avenue for improvement is to change the entire process by which development occurs. In its current form, development is uniformly distributed across network parameters and is linear in time. But this need not be true. One could include time as an input to compositional pattern producing networks [34] which would 'paint' different, possibly nonlinear, developmental trajectories onto different synapses. One could also envision employing NEAT [35] to compress base and target neural controllers with different cognitive architectures into a single, non-developing yet generalist architecture. Indeed, prior work has sought to marry network complexification and deep learning [18], which may also be applicable to developmental compression.

4.3 Synergies with other methods.

There exist myriad ways to combat catastrophic forgetting, each motivated by specific problems and the obstacles inherent to them. We do not purport to show that developmental compression succeeds where prior methods fail, or that it is preferable in situations where catastrophic forgetting has already been reasonably tempered [20]. Rather, we believe the results reported in this paper suggest developmental compression may be a potential mechanism for overcoming catastrophic forgetting—either independently or in concert with existing methods.

Although other techniques could be used to resist specialization, they often incur additional costs. For example, an objective that rewards generalists could be included in a multi-objective evolutionary algorithm. However, doing so would increase the dimensionality of the Pareto front, which is known to weaken optimization and requires bulwarking countermeasures [7], thus increasing algorithmic complexity. Another approach would be to formulate a better fitness function. However, this would put us on a frictionless path toward fitness function engineering, which may be appropriate for narrow problems but is likely to produce systems that break outside of their training regime. One could also take the minimum fitness across E to guard against specialization, but this will collapse gradients in the fitness landscape, because mutations that do not affect the worst fitness component aren't seen by search.

Unlike existing methods for combating catastrophic forgetting, developmental compression is expressly designed to minimize the need for supervised intervention in the form of domain knowledge or explicit constraints on behavior. Rather than decide what features of the system are necessary for generalist capabilities, developmental compression lets evolution discover the best way to preserve knowledge through a development program that integrates information obtained over the life of the agent. Given the generality of this method, it is likely to synergize well with existing techniques.

Future implementations will seek to improve developmental compression and to reconcile the results reported here with the work being done in deep neuroevolution [26]. The Atari-57 game suite and DMLab-30 environments are problems for which developmental compression could be scaled up, modified, and rigorously tested. Whether DC can be parallelized and used to compress populations of specialists across dissimilar tasks into a single master network remains to be seen. If such an ability can be realized, it

would represent an evolutionary alternative to the multi-task learning paradigm currently being explored in reinforcement learning [12].

5 CONCLUSIONS

Developmental compression is a novel technique for avoiding catastrophic forgetting. It works by relying on the smoothing effect of development, and how specialists can thus be gradually and developmentally 'compressed' into a generalist by canalization. It could also work in domains beyond robotics provided two conditions are met. First, each training instance must be extended in time. That is, several forward passes are required to complete one evaluation. Such a condition is usually met by reinforcement learning tasks. Second, performance must be integrated over all time steps of the evaluation. In future work we plan to study the efficacy of developmental compression in machine learning domains outside robotics and in tandem with other advances in the artificial intelligence literature.

6 ACKNOWLEDGMENTS

This work was supported by DARPA contract HR0011-18-2-0022. The computational resources provided by the UVM's Vermont Advanced Computing Core (VACC) are gratefully acknowledged.

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] J. Bongard and R. Pfeifer. 2003. Evolving complete agents using artificial ontogeny. *Morpho-functional Machines: The New Species (Designing Embodied Intelligence)* (2003), 237–258.
- [3] M. Castelli, L. Vanneschi, and S. Silva. 2014. Semantic Search-Based Genetic Programming and the Effect of Intron Deletion. *IEEE Transactions on Cybernetics* 44, 1 (2014), 103–113.
- [4] Jeff Clune, Jean-Baptiste Mouret, and Hod Lipson. 2013. The evolutionary origins of modularity. In *Proc. R. Soc. B*, Vol. 280. The Royal Society, 20122863.
- [5] Michael Conrad. 1990. The geometry of evolution. *BioSystems* 24, 1 (1990), 61–81.
- [6] Duc-Cuong Dang and Per Kristian Lehre. 2016. Self-adaptation of mutation rates in non-elitist populations. In *International Conference on Parallel Problem Solving from Nature*. Springer, 803–813.
- [7] Kalyanmoy Deb and Himanshu Jain. 2014. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints. *IEEE Trans. Evolutionary Computation* 18, 4 (2014), 577–601.
- [8] F. Dellaert and R.D. Beer. 1994. Toward an evolvable model of development for autonomous agent synthesis. *Artificial Life IV, Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems* (1994).
- [9] Peter Eggenberger. 1997. Creation of neural networks based on developmental and evolutionary principles. In *International Conference on Artificial Neural Networks*. Springer, 337–342.
- [10] KO Ellefsen, JB Mouret, J Clune, and Josh C Bongard. 2015. Neural Modularity Helps Organisms Evolve to Learn New Skills without Forgetting Old Skills. *PLoS Comput Biol* 11, 4 (2015), e1004128.
- [11] Kai Olav Ellefsen, Jean-Baptiste Mouret, and Jeff Clune. 2014. Neural Modularity Reduces Catastrophic Forgetting. *The Evolution of Learning: Balancing Adaptivity and Stability in Artificial Agents* (2014), 111.
- [12] Lasse Espelthol, Hubert Soyer, Remi Munos, Karen Simonyan, Volodymyr Mnih, Tom Ward, Yotam Doron, Vlad Firoiu, Tim Harley, Iain Dunning, Shane Legg, and Koray Kavukcuoglu. 2018. IMPALA: Scalable Distributed Deep-RL with Importance Weighted Actor-Learner Architectures. *arXiv preprint arXiv:1802.01561* (2018).
- [13] Carlos Espinosa-Soto and Andreas Wagner. 2010. Specialization can drive the evolution of modularity. *PLoS Comput Biol* 6, 3 (2010), e1000719.
- [14] Chrisantha Fernando, Dylan Banarse, Charles Blundell, Yori Zwols, David Ha, Andrei A Rusu, Alexander Pritzel, and Daan Wierstra. 2017. Pathnet: Evolution channels gradient descent in super neural networks. *arXiv preprint arXiv:1701.08734* (2017).
- [15] Robert M French. 1999. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences* 3, 4 (1999), 128–135.
- [16] Ian J Goodfellow, Mehdi Mirza, Da Xiao, Aaron Courville, and Yoshua Bengio. 2013. An empirical investigation of catastrophic forgetting in gradient-based neural networks. *arXiv preprint arXiv:1312.6211* (2013).
- [17] Geoffrey E Hinton and Steven J Nowlan. 1987. How learning can guide evolution. *Complex systems* 1, 3 (1987), 495–502.
- [18] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. 2017. Progressive Growing of GANs for Improved Quality, Stability, and Variation. *arXiv preprint arXiv:1710.10196* (2017).
- [19] Nadav Kashtan and Uri Alon. 2005. Spontaneous evolution of modularity and network motifs. *Proceedings of the National Academy of Sciences of the United States of America* 102, 39 (2005), 13773–13778.
- [20] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences* (2017), 201611835.
- [21] Kostas Kouvaris, Jeff Clune, Loizos Kounios, Markus Brede, and Richard A Watson. 2017. How evolution learns to generalise: Using the principles of learning theory to understand the evolution of developmental organisation. *PLoS computational biology* 13, 4 (2017), e1005358.
- [22] Sam Kriegman, Collin Cappelle, Francesco Corucci, Anton Bernatskiy, Nick Cheney, and Josh C Bongard. 2017. Simulating the evolution of soft and rigid-body robots. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. ACM, 1117–1120.
- [23] Sam Kriegman, Nick Cheney, and Josh Bongard. 2017. How morphological development can guide evolution. *arXiv preprint arXiv:1711.07387* (2017).
- [24] Sam Kriegman, Nick Cheney, Francesco Corucci, and Josh C Bongard. 2017. A minimal developmental model can increase evolvability in soft robots. In *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 131–138.
- [25] Sang-Woo Lee, Chung-Yeon Lee, Dong-Hyun Kwak, Jiwon Kim, Jeonghee Kim, and Byoung-Tak Zhang. 2016. Dual-Memory Deep Learning Architectures for Lifelong Learning of Everyday Human Behaviors.. In *IJCAI*. 1669–1675.
- [26] Joel Lehman, Jay Chen, Jeff Clune, and Kenneth O Stanley. 2017. Safe Mutations for Deep and Recurrent Neural Networks through Output Gradients. *arXiv preprint arXiv:1712.06563* (2017).
- [27] Hod Lipson, Jordan B Pollack, Nam P Suh, and P Wainwright. 2002. On the origin of modular variation. *Evolution* 56, 8 (2002), 1549–1556.
- [28] Michael McCloskey and Neal J Cohen. 1989. Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of learning and motivation* 24 (1989), 109–165.
- [29] Julian F Miller. 2003. Evolving developmental programs for adaptation, morphogenesis, and self-repair. In *European Conference on Artificial Life*. Springer, 256–265.
- [30] Tomasz P Pawlak, Bartosz Wieloch, and Krzysztof Krawiec. 2015. Semantic backpropagation for designing search operators in genetic programming. *IEEE Transactions on Evolutionary Computation* 19, 3 (2015), 326–340.
- [31] Andrei A Rusu, Neil C Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. 2016. Progressive neural networks. *arXiv preprint arXiv:1606.04671* (2016).
- [32] Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. 2017. Dynamic routing between capsules. In *Advances in Neural Information Processing Systems*. 3859–3869.
- [33] Russell Smith. 2008. Open dynamics engine v0.5 user guide. URL <http://www.ode.org/ode-docs.html> (2008).
- [34] Kenneth O Stanley. 2007. Compositional pattern producing networks: A novel abstraction of development. *Genetic programming and evolvable machines* 8, 2 (2007), 131–162.
- [35] Kenneth O Stanley and Risto Miikkilainen. 2002. Evolving neural networks through augmenting topologies. *Evolutionary computation* 10, 2 (2002), 99–127.
- [36] Marcin Szubert, Anuradha Kodali, Sangram Ganguly, Kamalika Das, and Josh C Bongard. 2016. Semantic Forward Propagation for Symbolic Regression. In *International Conference on Parallel Problem Solving from Nature*. Springer, 364–374.
- [37] Jason Teo, Asni Tahir, Norhayati Daut, Nordaliela Mohd Rusli, and Norazlina Khamis. 2016. Fixed vs. Self-Adaptive Crossover-First Differential Evolution. *Applied Mathematical Sciences* 10, 32 (2016), 1603–1610.
- [38] Leonardo Vanneschi, Mauro Castelli, and Sara Silva. 2014. A Survey of Semantic Methods in Genetic Programming. *Genetic Programming and Evolvable Machines* 15, 2 (2014), 195–214.
- [39] Roby Velez and Jeff Clune. 2017. Diffusion-based neuromodulation can eliminate catastrophic forgetting in simple neural networks. *arXiv preprint arXiv:1705.07241* (2017).