Arbitrary Order Meta-Learning with Simple Population-Based Evolution

Chris Lu, Sebastian Towers and Jakob Foerster

Department of Engineering Sciences, University of Oxford, Oxford, United Kingdom
christopher.lu@exeter.ox.ac.uk

Abstract

Meta-learning, the notion of learning to learn, enables learning systems to quickly and flexibly solve new tasks. This usually involves defining a set of outer-loop *meta-parameters* that are then used to update a set of inner-loop parameters. Most meta-learning approaches use complicated and computationally expensive bi-level optimisation schemes to update these meta-parameters. Ideally, systems should perform multiple orders of meta-learning, i.e. to learn to learn to learn and so on, to accelerate their own learning. Unfortunately, standard meta-learning techniques are often inappropriate for these higher-order meta-parameters because the meta-optimisation procedure becomes too complicated or unstable. Inspired by the higher-order meta-learning we observe in real-world evolution, we show that using simple population-based evolution *implicitly* optimises for arbitrarily-high order meta-parameters. First, we theoretically prove and empirically show that population-based evolution implicitly optimises meta-parameters of arbitrarily-high order in a simple setting. We then introduce a minimal *self-referential* parameterisation, which in principle enables arbitrary-order meta-learning. Finally, we show that higher-order meta-learning improves performance on time series forecasting tasks.

Introduction

The natural world contains multiple orders of meta-evolution and adaptation (Vanchurin et al., 2022). For example, DNA has not just evolved to produce an organism, but has also evolved to be *evolvable* (Zheng et al., 2020; Woods et al., 2011; Metzgar and Wills, 2000). In other words, DNA has evolved such that random mutations in a genotype frequently result in useful or adaptive changes in the resulting organism's phenotype. Furthermore, the evolution of DNA has created organisms that have the ability to *adapt* within their lifetime, one form of which is organisms that perform *reinforcement learning* (Bateson, 1984). These learning organisms further influence their own learning through social interactions and culture (Henrich, 2015; Heyes, 2018).

However, most existing works only investigate single-order meta-learning, for example for evolving reinforcement learning algorithms (Lu et al., 2022a). These approaches commonly use computationally expensive bi-level optimisation schemes that quickly becomes unstable or computationally intractable when applied to higher orders of meta-learning (Metz et al., 2021b).

Past work has empirically shown that population-based evolution implicitly selects for single order meta-learning, usually by simultaneously evolving mutation rates (Frans and Witkowski, 2021; Bäck et al., 1992; Smith, 1998). Other work has investigated multiple orders of meta-learning, but in the context of gradient-based optimisation (Chandra et al., 2019) and multi-agent learning (Willi et al., 2022). Finally, Kirsch and Schmidhuber (2022); Lange et al. (2022); Metz et al. (2021a) empirically investigate using evolution-like algorithms on *self-referential* systems to perform self-referential meta-learning, an idea first articulated in Schmidhuber (1987). However, these works do not theoretically prove that they perform higher-order meta-learning. We connect these works by *theoretically proving* and empirically showing that under some circumstances simple population-based evolution selects for *arbitrarily-high* orders of meta-learning, which in principle allows for arbitrary orders of self-improvement in self-referential systems.

Numeric Fitness World

We perform population-based evolution by selecting and mutating the top k most fit individuals at each generation. Unlike past work, we do this on genomes with *multiple orders* of meta-parameters. In particular, we represent a genome x at generation t with n orders of meta-parameters as a vector of n parameters, $x_t = \{x_t^0, x_t^1, \cdots, x_t^n\}$, where x^i represents the ith-order meta-parameter. We consider the setting of "Numeric Fitness World" in which *fitness*$(x_t) = x_t^0$, proposed in Frans and Witkowski (2021).

We *mutate* x_t using the following update rule:

$$x_{t+1}^i = x_t^i + x_t^{i+1} + B_t^i, 0 \leq i < n, 0 < t \quad (1)$$

$$x_{t+1}^n = x_t^n + B_t^n, 0 < t \quad (2)$$

$$B_t^i \sim \mathcal{N}(0, \beta), i.i.d, \forall t, i \quad (3)$$

In other words, we update the ith-order meta-parameter by adding the $(i+1)$th parameter and noise sampled from a normal distribution. We update the last meta-parameter

(x_t^n) with just the noise. To instead create a *self-referential* parameterisation, we update the last meta-parameter with *itself* and the noise. This exact form of self-reference is likely inappropriate in most settings, but may be sensible in other parameterisations, such as neural networks (Irie et al., 2022).

Theoretical Results

We prove that top-k selection selects for the fitness of higher-order meta-parameters in this setting if and only if $k > 1$.

Let P define a population of individual members as defined above. Let x be a specific member of P. Let $\bar{P}$ define a population identical to P except in the n-th parameter of x. More specifically, $\bar{x}^n - x^n = \delta, \delta > 0$.

Let $F(P, B, t)$ and $F^{-1}(P, B, t)$ represent the set of fitnesses of the descendants and non-descendants of x respectively in population P after t generations of selection and vector of mutations B. Note that $|F(P, B, t + 1)|$ would be the number of descendants of x after generation t.

First, we show that top-1 (single-genome) selection does not select for higher-order meta-parameters.

Theorem 1. $\mathbb{E}[|F(\bar{P}, B, n + 1)|] = \mathbb{E}[|F(P, B, n + 1)|]$ *under top-1 selection for* $n > 1$.

Proof. The number of descendants at generation $t > 1$ is entirely determined by the first selection. Either *all* members of the population at generation n are descendants of x, or *none* are. As x^n $(n > 1)$ does not affect the first selection, it is independent to the number of descendants. □

Next, we show that top-k selection selects for higher-order meta-parameters for $k > 1$.

Lemma 2. $|F(\bar{P}, B, n+1|B = b)| \geq |F(P, B, n+1|B = b)|$ *for any vector of mutations b. for* $k > 1$

Proof. Note that for $t < n$, $F(\bar{P}, B, t|B = b) = F(P, B, t|B = b)$, as none of the fitnesses are influenced by x^n, the only value in which the two populations differ.

$F(\bar{P}, B, n|B = b) = F(P, B, n|B = b) \oplus \delta$ where $\oplus$ represents a distributed addition.

$F^{-1}(\bar{P}, B, n|B = b) = F^{-1}(P, B, n|B = b)$ because x^n can not influence the selection or fitness of other members before generation $n + 1$.

Thus, there can be no fewer descendants of $\bar{x}$ than descendants of x in the top-k of the next generation . □

Theorem 3. $\mathbb{E}[|F(\bar{P}, B, n + 1)|] > \mathbb{E}[|F(P, B, n + 1)|]$

Proof. By Lemma 2, $|F(\bar{P}, B, n + 1)| \geq |F(P, B, n + 1)|$. Hence, showing $\mathbb{P}(|F(\bar{P}, B, n + 1)| > |F(P, B, n + 1)|) > 0$ is sufficient for our result. In particular, we show $\mathbb{P}(|F(P, B, n + 1)| = 0 \cap |F(\bar{P}, B, n + 1)| = 1) > 0$. There is a set of intervals over B such that exactly k members of $F^{-1}(P, B, n)$ lie in the range $[\max F(P, B, n), \max F(P, B, n) + \delta]$ and the rest are less than $\max F(P, B, n)$. Thus, after selection there is exactly one descendant of $\bar{x}$, and none of x. □

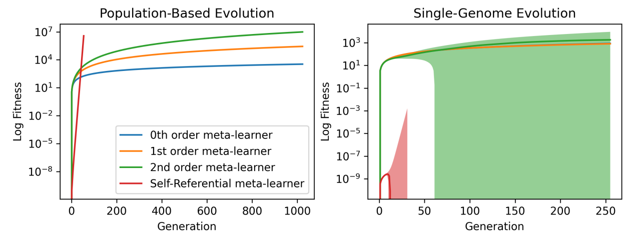

Figure 1: Population-based evolution (Top-2) and single-genome evolution with varying orders of meta-learning with a population size of 2048. The shaded region refers to the standard error of the mean across 1024 seeds.

	Meta-Learning Order			
$f(t)$	0th	1st	2nd	3rd
t	1.0	**3.7e-4**	9.4e-3	5.0e-2
t^2	1.3e7	7.5	0.77	**0.56**
$\sin(t)$	6.6e-2	1.0e-3	**9.4e-4**	1.2e-2
$\sin(t \sin(t))$	2.4	0.67	0.31	**0.16**

Table 1: The average prediction error across 4096 generations of evolution with population size 16384 and top-1024 selection with 64 seeds. For each entry we selected the best performing $\beta \in \{1.0, 0.5, 0.1, 0.05, 0.01\}$
.

Empirical Results

We simulate the evolution using Jax (Bradbury et al., 2018) and show the results in Figure 1. We observed that the asymptotic growth in fitness is approximately of the order x^n where n is the number of meta-parameters. The fitness of the self-referential meta-learner grows *exponentially*. Thus, in our population-based setting higher orders of meta-parameters improve fitness. In contrast, in single-genome selection, the expected value of the fitness is independent of the number of meta-parameters, demonstrating that single-genome selection does not perform meta-optimisation.

Time Series Forecasting

Next, we consider a time series forecasting task where the goal is to predict the next value of some function $f(t)$. The fitness of an individual x_t is determined by *fitness*$(x_t) = -|f(t/100) - x_t^0|$. We report the results on a number of functions in Table 1. Higher orders of meta-parameters improve performance in many of these settings.

Future Work

One could investigate the emergence of higher-order meta-learning in multi-agent systems (Lu et al., 2022b,c) or open-ended simulations (Chan, 2018). Future work would also involve theoretically analysing the long-term properties of these systems, alongside evaluating other parameterisation and selection schemes on more practical time series forecasting tasks.

References

Bäck, T. et al. (1992). Self-adaptation in genetic algorithms. In *Proceedings of the first european conference on artificial life*, pages 263–271. MIT press Cambridge.

Bateson, P. (1984). Genes, evolution, and learning. In *The Biology of Learning: Report of the Dahlem Workshop on the Biology of Learning Berlin, 1983, October 23–28*, pages 75–88. Springer.

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., and Zhang, Q. (2018). JAX: composable transformations of Python+NumPy programs.

Chan, B. W.-C. (2018). Lenia-biology of artificial life. *arXiv preprint arXiv:1812.05433*.

Chandra, K., Meijer, E., Andow, S., Arroyo-Fang, E., Dea, I., George, J., Grueter, M., Hosmer, B., Stumpos, S., Tempest, A., et al. (2019). Gradient descent: The ultimate optimizer. *arXiv preprint arXiv:1909.13371*.

Frans, K. and Witkowski, O. (2021). Population-based evolution optimizes a meta-learning objective. *arXiv preprint arXiv:2103.06435*.

Henrich, J. (2015). The secret of our success. In *The Secret of Our Success*. princeton University press.

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.

Irie, K., Schlag, I., Csordás, R., and Schmidhuber, J. (2022). A modern self-referential weight matrix that learns to modify itself. In *International Conference on Machine Learning*, pages 9660–9677. PMLR.

Kirsch, L. and Schmidhuber, J. (2022). Eliminating meta optimization through self-referential meta learning. *arXiv preprint arXiv:2212.14392*.

Lange, R. T., Schaul, T., Chen, Y., Zahavy, T., Dallibard, V., Lu, C., Singh, S., and Flennerhag, S. (2022). Discovering evolution strategies via meta-black-box optimization. *arXiv preprint arXiv:2211.11260*.

Lu, C., Kuba, J. G., Letcher, A., Metz, L., de Witt, C. S., and Foerster, J. (2022a). Discovered policy optimisation. *arXiv preprint arXiv:2210.05639*.

Lu, C., Willi, T., De Witt, C. A. S., and Foerster, J. (2022b). Model-free opponent shaping. In *International Conference on Machine Learning*, pages 14398–14411. PMLR.

Lu, C., Willi, T., Letcher, A., and Foerster, J. (2022c). Adversarial cheap talk. *arXiv preprint arXiv:2211.11030*.

Metz, L., Freeman, C. D., Maheswaranathan, N., and Sohl-Dickstein, J. (2021a). Training learned optimizers with randomly initialized learned optimizers. *arXiv preprint arXiv:2101.07367*.

Metz, L., Freeman, C. D., Schoenholz, S. S., and Kachman, T. (2021b). Gradients are not all you need. *arXiv preprint arXiv:2111.05803*.

Metzgar, D. and Wills, C. (2000). Evidence for the adaptive evolution of mutation rates. *Cell*, 101(6):581–584.

Schmidhuber, J. (1987). *Evolutionary principles in self-referential learning, or on learning how to learn: the meta-meta-... hook*. PhD thesis, Technische Universität München.

Smith, J. (1998). *Self adaptation in evolutionary algorithms*. PhD thesis, Citeseer.

Vanchurin, V., Wolf, Y. I., Katsnelson, M. I., and Koonin, E. V. (2022). Toward a theory of evolution as multilevel learning. *Proceedings of the National Academy of Sciences*, 119(6):e2120037119.

Willi, T., Letcher, A. H., Treutlein, J., and Foerster, J. (2022). Cola: consistent learning with opponent-learning awareness. In *International Conference on Machine Learning*, pages 23804–23831. PMLR.

Woods, R. J., Barrick, J. E., Cooper, T. F., Shrestha, U., Kauth, M. R., and Lenski, R. E. (2011). Second-order selection for evolvability in a large escherichia coli population. *Science*, 331(6023):1433–1436.

Zheng, J., Guo, N., and Wagner, A. (2020). Selection enhances protein evolvability by increasing mutational robustness and foldability. *Science*, 370(6521):eabb5962.