Vrije Universiteit Amsterdam

Bachelor Thesis

Co-Evolution of Generalist Morphology and Control for Diverse Environments

Author: Kubilay Tarhan (2721178)

1st supervisor: Anil Yaman
2nd reader: supervisor name

A thesis submitted in fulfillment of the requirements for the VU Bachelor of Science degree in Computer Science

June 11, 2024

Abstract

This section should contain a brief summary of the key points of your paper, including the problem statement, methodology, results, and conclusion.

Keywords: keyword1, keyword2, keyword3

1 Introduction

Since Karl Sims' publication in 1994, 'Evolving Virtual Creatures' [1], demonstrated virtual creatures interacting within a simulated three-dimensional physical environment, numerous researchers have pursued simultaneous co-optimization of both morphology and control [2–5]. By co-optimizing both morphology and control, researchers can develop more effective AI agent systems for various tasks. This approach leverages the principle of embodied cognition, which suggests that intelligence arises not solely from the brain or an agent's control system, but from the dynamic interaction between the brain, body, and environment [6].

Unfortunately, an optimized morphology-controller pair will still specialize on the environment it has been trained in. This specialization occurs because the evolutionary process fine-tunes both the morphology and controller to the specific conditions of the training environment, resulting in high task performance in only the training environments but reduced task performance to new or marginally different environments. In the real-world, environments are inherently dynamic and unpredictable. Factors such as changing weather conditions, varying terrains, and unforeseen obstacles can significantly alter the operational context of an agent. This variability poses a substantial challenge to the robustness and generalizability of optimized morphology-controller pairs. For instance, an agent trained to navigate a smooth indoor surface may struggle when faced with outdoor terrains that include gravel, mud, or steep inclines. Similarly, agents designed for static environments might fail to adapt to environments with moving obstacles or varying lighting conditions.

To overcome this limitation, it is crucial to focus on developing robustness and generalizability in optimized agents. Robustness refers to the ability of an agent to maintain desirable behavior despite variations or perturbations in its input and output data [7–9]. Generalizability, on the other hand, refers to the ability of an agent to maintain desirable behavior under different conditions to those encountered during training [8, 9]. Co-optimized morphology-controller pairs have been shown to generalise well to environments similar to those encountered during training. However, comprehensive studies examining the robustness and generalizability of co-optimized morphology-controller pairs remain sparse in the existing literature.

In this paper, we perform an extensive analysis on ...BLABLABLA

2 Background Information

2.1 Co-optimizing of morphology and control

2.1.1 Embodied cognition

Despite the increase of computing power and knowledge, there has not been a significant leap forward since Sims' initial work on the co-optimization of morphology-controller pairs. Researchers have proposed various hypotheses causing this, ranging from deficiencies in optimization algorithms to the notion that the training environments are not complex enough to facilitate morphological evolution, as outlined by Cheney et al. [10]. However, they suggest a new hypotheses to explain the difficulty in co-optimizing morphology-controller pairs, which is the theory of embodied cognition.

(Put more elaboration on what embodied cognition is) Embodied cognition suggests that the cognitive processes arise not solely from the brain, but are a product of the dynamic interaction between the brain and morphology. This interplay means that even minor changes in morphology can disrupt the connection between the controller and morphology, requiring the controller to adapt itself again to the new morphology. This is also the reason why in [10] the morphologies generally converge before the 100th generation out of a total of 5000 generations, because at that point, further morphological evolution becomes less beneficial as it will only result in lower fitness scores. This phenomenon of premature convergence poses a challenge in co-optimizing morphology-controller pairs.

2.1.2 Premature convergence

Premature convergence occurs when the population quickly converges to a local optimum, resulting in a lack of genetic diversity and suboptimal solutions. In the co-evolution of robot morphology and controllers, this leads to early stagnation of morphology evolution, as morphological changes disrupt optimized controllers and are discarded by selection pressure, preventing the benefits of co-evolution. [5]

For the co-optimization of morphology-controller pairs, Lehman et al. [11] propose the Novelty Search with Local Competition (NSLC) algorithm, which utilizes a multi-objective search to optimize both diverse morphologies and fitness in conjunction with local competition, rewarding agents that outperform others with similar morphologies. While this approach did show an increase in the novelty of morphology, it did not yield higher fitness scores compared to using only a fitness objective function. Another method to address premature convergence is fitness sharing, which modifies the fitness function so that agents that are similar get penalized, thereby preserving and encouraging more morphological diversity [12]. Subsequent studies by Cheney et al. [2] propose explicitly protecting agents that have undergone a recent morphological mutation, which reduces the selection pressure and gives the controller more time to adapt to the new morphology. Using the open-source soft-body simulator VoxCad as the physics engine, their experiments showed that this method produced significantly higher fitness scores and increased morphological diversity. Additionally, it delayed premature convergence, as the best-performing agents emerged in later generations, who initially would have been discarded otherwise. Stensby et al. [3] extended upon this work by utilizing a more indirect approach to mitigate premature convergence. They discuss that increasing the agents' exploration of new morphologies can be facilitated by training the agents in a diverse range of environments. These environments were made incrementally more challenging using the Paired Open-Ended Trailblazer (POET) algorithm. Their findings demonstrated that environments generated by POET increased morphology diversity, indicating that POET, or other forms of curriculum training, could be effective in delaying convergence.

2.1.3 CPPN-NEAT

Neuroevolution, the process of evolving artificial neural networks using evolutionary algorithms, has been proven to be a great alternative to traditional reinforcement learning algorithms, particularly in continuous and high-dimensional input spaces. Conventionally, the topology of the neural network was established and fixed prior to training, which is a huge drawback, because the network's topology significantly influences the agent's performance. This drawback led to the development of Topology and Weight Evolving Artificial Neural Networks (TWEANNs), with the most prominent being the NeuroEvolution of Augmenting Topologies (NEAT) algorithm [13]. This algorithm evolves both the weights and the topology of neural networks. It initializes a population of minimal networks with only input and output nodes and incrementally adds nodes and connections through mutations while utilizing a genetic algorithm for optimization. An important technique NEAT uses is speciation, which protects mutated networks, ensuring that new structures are not discarded and have the chance to evolve.

A notable advantage of NEAT is its use in conjunction with Compositional Pattern-Producing Networks (CPPNs). Unlike conventional neural networks, CPPNs utilize a variety of activation functions to generate complex and regular patterns [14]. This can be employed for encoding morphology parameters by mapping morphology features to specific values. In a study by Cheney et al. [15], CPPN-NEAT was used to encode the morphology of a soft robot, demonstrating that generative encoding using CPPNs assigned tissue cells to the voxels in a logical way, ensuring global coordination, resulting in improved locomotion. In contrast, the direct encoding assigned the voxels more independently of their neighboring voxels, resulting in less coordination and worsening locomotion.

2.2 Robustness and generalizability

Robustness refers to the ability of an agent to maintain desirable behavior despite variations or perturbations in its input and output data [7–9]. Consequently, a robust agent is less susceptible to input and output perturbations. This makes robustness a critical attribute, as inputs and outputs from the training environment can differ significantly from those in the testing environment, especially in real-world scenarios. For instance, an input perturbation can be caused by a sensor defect, resulting in slight measurement errors. In reinforcement learning, this means that we need to build robustness against the uncertainty of state observations and the actual state. Similarly, an output perturbation can result from a motory issue of the agent. In reinforcement learning, this means that we need to build robustness against uncertain actions between the actions generated by the agent and the conducted actions [9].

On the other hand, generalizability refers to the ability of an agent to maintain desirable behavior under different conditions to those encountered during training. It assumes correct input and output data and is concerned on the actual differences between the training environment and the testing or production environment. Testing generalization is divided into two distinctive parts. The first part is called interpolation, which requires the agent to perform well in environments similar to those of training, with both the testing and training environment parameters drawn from the same distribution. The second part is called extrapolation, which requires the agent to perform well in environments different from those of training, with the testing and training environment parameters drawn from separate distributions [8, 9].

There are some ways to achieve more robust agents. In this example [16], they compared the robustness of the models DENSER and NSGA-Net on the CIFAR-10 image classification task. It was found that the DENSER model exhibited an higher robustness, which concludes that certain architectures inherently have a higher degree of exhibiting robustness. Furthermore, one of the most popular method is adverserial training, where the agent is trained on adverserially perturbated training data. One way of doing this is finding the worst case perturbation at each training episode and training the model using the dataset with this perturbation [17].

The two main approaches to achieving generalizable agents are either training a generalist agent, which involves a trade-off in performance under specific conditions compared to a specialist agent, as shown by Triebold et al. [18], or developing agents that can explicitly adapt to certain conditions [8]. In this paper, we will focus on the first approach. Recently, there has been a growing recognition of the importance of using variability to train better and more generalized agents. Raviv et al. [19] discusses the relationship between variability and learning outcomes, highlighting a universal principle that variability enhances learning. This principle also holds true for machine learning, where employing a more variable learning curriculum can improve an agent's generalizability across a wide range of morphologies. Similarly, Stensby et al. [3], applied the same principle of variability by training agents in a curriculum of environments generated by the Paired Open-Ended Trailblazer (POET) algorithm, where the generated environments were incrementally more difficult, enabling the agent to learn more efficient and complex behaviors. These studies demonstrate that using a variable learning curriculum increases both robustness and generalizability.

3 Experimental Setup

3.1 Environment

3.2 Algorithm

3.3 Training data

3.4 Testing and evaluation

4 Results

5 Discussion

6 Conclusion

References

[1] Karl Sims. "Evolving virtual creatures". In: *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '94. New York, NY, USA: Association for Computing Machinery, 1994, pp. 15–22. ISBN: 0897916670. DOI: 10.1145/192161.192167. URL: https://doi.org/10.1145/192161.192167.

[2] Nick Cheney et al. *Scalable Co-Optimization of Morphology and Control in Embodied Machines*. 2017. arXiv: 1706.06133 [cs.AI].

[3] Emma Hjellbrekke Stensby, Kai Olav Ellefsen, and Kyrre Glette. "Co-optimising robot morphology and controller in a simulated open-ended environment". In: *Applications of Evolutionary Computation* (2021), pp. 34–49. DOI: 10.1007/978-3-030-72699-7_3.

[4] Joshua E. Auerbach and Josh C. Bongard. "Environmental Influence on the Evolution of Morphological Complexity in Machines". In: *PLOS Computational Biology* 10 (2014), pp. 1–17. DOI: 10.1371/journal.pcbi.1003399. URL: https://doi.org/10.1371/journal.pcbi.1003399.

[5] Luis Eguiarte-Morett and Wendy Aguilar. "Premature convergence in morphology and control co-evolution: a study". In: *Adaptive Behavior* 32 (2024), pp. 137–165. DOI: 10.1177/10597123231198497. URL: https://doi.org/10.1177/10597123231198497.

[6] Josh C. Bongard. "Evolutionary robotics". In: *Commun. ACM* (2013), pp. 74–83. ISSN: 0001-0782. DOI: 10.1145/2493883. URL: https://doi.org/10.1145/2493883.

[7] R. Mangal, A. V. Nori, and A. Orso. "Robustness of Neural Networks: A Probabilistic and Practical Approach". In: *2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)*. 2019, pp. 93–96. DOI: 10.1109/ICSE-NIER.2019.00032. URL: https://doi.ieeecomputersociety.org/10.1109/ICSE-NIER.2019.00032.

[8] Charles Packer et al. *Assessing Generalization in Deep Reinforcement Learning*. 2019. arXiv: 1810.12282 [cs.LG].

[9] Mengdi Xu et al. *Trustworthy Reinforcement Learning Against Intrinsic Vulnerabilities: Robustness, Safety, and Generalizability*. 2022. arXiv: 2209.08025 [cs.LG].

[10] *On the Difficulty of Co-Optimizing Morphology and Control in Evolved Virtual Creatures*. Vol. ALIFE 2016, the Fifteenth International Conference on the Synthesis and Simulation of Living Systems. Artificial Life Conference Proceedings. 2016, pp. 226–233. DOI: 10.1162/978-0-262-33936-0-ch042. URL: https://doi.org/10.1162/978-0-262-33936-0-ch042.

[11] Joel Lehman and Kenneth Stanley. "Evolving a diversity of creatures through novelty search and local competition". In: 2011, pp. 211–218. DOI: 10.1145/2001576.2001606.

[12] Robert I McKay. "Fitness Sharing in Genetic Programming." In: *GECCO*. 2000, pp. 435–442.

[13] Kenneth O. Stanley and Risto Miikkulainen. "Evolving Neural Networks through Augmenting Topologies". In: *Evolutionary Computation* 10.2 (2002), pp. 99–127. DOI: 10.1162/106365602320169811.

[14] Kenneth O. Stanley. "Compositional pattern producing networks: A novel abstraction of development". In: *Genetic Programming and Evolvable Machines* 8.2 (2007), pp. 131–162. ISSN: 1389-2576. DOI: 10.1007/s10710-007-9028-8. URL: https://doi.org/10.1007/s10710-007-9028-8.

[15] Nick Cheney et al. "Unshackling evolution". In: *ACM SIGEVOlution* 7 (2014), pp. 11–23. DOI: 10.1145/2661735.2661737.

[16] Inês Valentim, Nuno Lourenço, and Nuno Antunes. *Adversarial Robustness Assessment of NeuroEvolution Approaches*. 2022. arXiv: 2207.05451 [cs.NE].

[17] Kai Liang Tan et al. *Robustifying Reinforcement Learning Agents via Action Space Adversarial Training*. 2020. arXiv: 2007.07176 [cs.LG].

[18] Corinna Triebold and Anil Yaman. *Evolving generalist controllers to handle a wide range of morphological variations*. Sept. 2023. DOI: 10.13140/RG.2.2.27217.71522.

[19] Limor Raviv, Gary Lupyan, and Shawn Green. "How variability shapes learning and generalization". In: *Trends in Cognitive Sciences* 26 (May 2022). DOI: 10.1016/j.tics.2022.03.007.