Internship project, master level, 2024

Evolving large populations of adaptive neural agents in ecologically plausible environments

Supervision: Clément Moulin-Frier and Gautier Hamon

Research group: [Flowers team](), Inria Bordeaux, France

Duration: 6 months, around February 2024 (with flexibility).

Level: Master 2 research internship

Keywords: Multi-agent systems, meta reinforcement learning, recurrent neural networks, simulation environments, eco-evolutionary dynamics, scientific programming on GPUs with Python.

How to apply: contact [clement.moulin-frier@inria.fr]() ; [gautier.hamon@inria.fr]() ; with a CV and letter of motivation, ideally by the end of November 2023. We also recommend sending documents or reports describing previous projects you have been working on (even if they are not directly related to the topic), as well as your grades and links to code repositories.

Requirements: We are looking for highly motivated MSc students (Master II). Programming skills and prior experience with Python and deep learning frameworks (e.g. Pytorch, JAX) are expected.
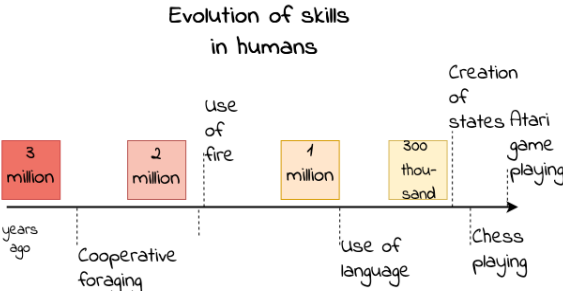
In open-ended environments, like Minecraft, agents can acquire skills in an ecological context that resembles the one encountered by humans throughout their evolution

Context

There are striking differences in how adaptation operates in biological versus artificial systems. Reinforcement learning optimizes an action policy in order to maximize reward provided by the environment. In contrast, the natural world does not contain any explicit reward whatsoever: reward systems have instead evolved physiologically in biological organisms. Moreover, agent's training in RL relies on an episodic paradigm where the environment is reset to initial conditions at the start of each new episode. In contrast, natural environments are never reset: they are instead continuously modified by their inhabitants over many generations.

Those differences arise from the central assumption in AI that intelligence must be implemented in a structured cognitive architecture (integrating e.g. control, learning and memory) which is optimized (using machine learning methods) through pre-defined objective functions (Chollet, 2019; Silver et al., 2021). The performance of these architectures are then evaluated in benchmarks that quantitatively capture various aspects of intelligence (e.g. Chollet, 2019). In contrast, biological adaptation seems to be better characterized by the notion of open-endedness (the continual generation of increasingly diverse organisms and behaviors) than by performance (Stanley et al., 2017). An important paradigm shift is taking increasing importance in evolutionary biology, recognizing the crucial role of eco-evolutionary feedbacks as a main driver of evolution: Developing organisms are not solely products of evolution but, by modifying their niche and therefore its associated fitness landscape, are also causes of their own evolution and of others (Laland et al., 2015). Following a similar paradigm shift, a recent trend in AI is increasingly recognizing the importance of reciprocal influence between adaptation and environmental dynamics (Clune, 2020; Leibo et al., 2019; Moulin-Frier, 2022).

Project

In a recent paper (Hamon et al., 2023), we proposed to study the eco-evolutionary dynamics of non-episodic neuroevolution in large multi-agent environments, based on the following principles. *Non-episodic learning:* We prevent any environment or population reset during a simulation, which leads to continuous environmental and population dynamics. *Bi-level adaptation mimicking the interplay between evolution and development*: Agents' behavior is controlled by recurrent neural networks optimized through neuroevolution (Lehman & Miikkulainen, 2013), potentially enabling adaptation within the agent's lifetime in the absence of weight updates (in the spirit of meta reinforcement learning, Duan et al., 2016). *Physiology-driven death and reproduction*: There is no notion of rewards, agents are instead equipped with a physiological system modulating their energy level according to the resources they consume, surviving and reproducing as long as they are able to maintain this level within a reasonable range (i.e. no explicit notion of "survival of the fittest", Brant & Stanley, 2020). *Ecologically valid environment with complex intrinsic dynamics:* We model our environment after common-pool resource (CPR) appropriation problems (Pérolat et al., 2017), where a group of agents competes for finite resources in the presence of multiple niches.

The objective of the proposed internship is to extend these preliminary experiments along several possible directions of research (with flexibility depending on the student's background and interest). These include:

- The use of indirect encoding of macro properties of the agents' neural networks using recent methods from neuroevolution (e.g. Najarro et al., 2023; Zuo et al., 2023), meta reinforcement learning (e.g. Duan et al., 2016) or transformer-based architectures.
- The implementation of a more complex ecologically plausible simulated environment with realistic physics, compositional dynamics (where resources can be combined together to create new ones, in the spirit of Minecraft, see also Bornemann et al., 2023 from our team), spatio-temporal variability (e.g. seasonal cycles) and multiple co-evolving species (e.g. prey-predator dynamics).
- Thorough qualitative and quantitative evaluation of the system dynamics, both at the macro level (characterizing eco-evolutionary dynamics in the system) and of the agent's evolved behavior, in particular in their ability to adapt during their lifetime (e.g. implementing a battery of test lab environments characterizing their generalization abilities in unknown situations).

References

Bornemann, R., Hamon, G., Nisioti, E., & Moulin-Frier, C. (2023). *Emergence of Collective Open-Ended Exploration from Decentralized Meta-Reinforcement Learning*. NeurIPS 2023 - Conference on Neural Information Processing Systems / ALOE Workshop. https://doi.org/10.48550/arXiv.2311.00651

Brant, J. C., & Stanley, K. O. (2020). Diversity preservation in minimal criterion coevolution through resource limitation. *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, 58–66. https://doi.org/10.1145/3377930.3389809

Chollet, F. (2019). *On the Measure of Intelligence* (arXiv:1911.01547). arXiv. https://doi.org/10.48550/arXiv.1911.01547

Clune, J. (2020). AI-GAs: AI-generating algorithms, an alternate paradigm for producing general artificial intelligence. *arXiv:1905.10985 [Cs]*. http://arxiv.org/abs/1905.10985

Duan, Y., Schulman, J., Chen, X., Bartlett, P. L., Sutskever, I., & Abbeel, P. (2016). RL^2: Fast Reinforcement Learning via Slow Reinforcement Learning. *arXiv:1611.02779 [Cs, Stat]*. http://arxiv.org/abs/1611.02779

Hamon, G., Nisioti, E., & Moulin-Frier, C. (2023). *Eco-evolutionary Dynamics of Non-episodic Neuroevolution in Large Multi-agent Environments* (arXiv:2302.09334). arXiv. https://doi.org/10.48550/arXiv.2302.09334

Laland, K. N., Uller, T., Feldman, M. W., Sterelny, K., Müller, G. B., Moczek, A., Jablonka, E., & Odling-Smee, J. (2015). The extended evolutionary synthesis: Its structure, assumptions and predictions. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1813), 20151019. https://doi.org/10.1098/rspb.2015.1019

Lehman, J., & Miikkulainen, R. (2013). Neuroevolution. *Scholarpedia*, *8*(6), 30977. https://doi.org/10.4249/scholarpedia.30977

Leibo, J. Z., Hughes, E., Lanctot, M., & Graepel, T. (2019). Autocurricula and the emergence of innovation from social interaction: A manifesto for multi-agent intelligence research. *arXiv Preprint arXiv:1903.00742*.

Moulin-Frier, C. (2022). *The Ecology of Open-Ended Skill Acquisition* [Habilitation thesis (HDR), Université de Bordeaux (UB)]. https://hal.inria.fr/tel-03875448

Najarro, E., Sudhakaran, S., & Risi, S. (2023, July 24). *Towards Self-Assembling Artificial Neural Networks through Neural Developmental Programs*. ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference. https://doi.org/10.1162/isal_a_00697

Pérolat, J., Leibo, J. Z., Zambaldi, V., Beattie, C., Tuyls, K., & Graepel, T. (2017). A multi-agent reinforcement learning model of common-pool resource appropriation. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 3643–3652. http://papers.nips.cc/paper/6955-a-multi-agent-reinforcement-learning-model-of-common-pool-resource-appropriation

Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. *Artificial Intelligence*, *299*, 103535. https://doi.org/10.1016/j.artint.2021.103535

Stanley, K., Lehman, J., & Soros, L. (2017). *Open-endedness: The last grand challenge you've never heard of*. https://www.oreilly.com/radar/open-endedness-the-last-grand-challenge-youve-never-heard-of/

Zuo, W., Pedersen, J., & Risi, S. (2023). Evolution of an Internal Reward Function for Reinforcement Learning. *Proceedings of the Companion Conference on Genetic and Evolutionary Computation*, 351–354. https://doi.org/10.1145/3583133.3590610