A-life Evolution with Human Proxies

Vadim Bulitko and Kacy Doucet and Daniel Evans
Hope Docking and Mac Walters and Marilene Oliver
Julian Chow and Shelby Carleton and Natali Kendal-Freedman

University of Alberta, Edmonton, AB, Canada
bulitko@ualberta.ca

Abstract

Recent successes in Artificial Intelligence (AI) use machine learning to produce AI agents with both hand-engineered and procedurally generated elements learned from large amounts of data. As the balance shifts toward procedural generation, how can we predict interactions between such agents and humans? We propose to use Artificial Life to study emergence of group behaviours between procedurally generated AI agents and humans. We simulate Darwinian evolution to procedurally generate agents in a simple environment where the agents interact with human-controlled avatars. To reduce human involvement time, we machine-learn another set of AI agents that mimic human avatar behaviours and run the evolution with such human proxies instead of actual humans. This paper is an update on the on-going project.

1 Introduction

Artificial Intelligence (AI) is increasingly present in our lives in the form of smartphone assistants, smart appliances, self-driving cars, non-player characters in video games, etc. Consequently, emergent interactions between AI agents and humans can significantly affect our lives and help us address various societal challenges. Artificial life (A-life) is a powerful setting in which we can study behaviours emerging from interactions among AI agents. By adding human-controlled avatars to an A-life environment we can additionally study interactions between A-life and humans to gain insight into the continual impact of AI on society.

Most recent advances in AI use machine learning wherein the AI agents become progressively smarter by learning from existing data and their own experiences at various temporal scales. Evolutionary algorithms can learn across generations while reinforcement learning can learn from the agent's experience within its lifetime. Seminal work by Ackley and Littman (1991) combined two techniques into Evolutionary Reinforcement Learning.

Models of the evolution of AI agents in society must be informed by human involvement. As evolution can take many generations and human time is limited, we proposed in our earlier paper (Bulitko et al., 2018) to derive AI-controlled proxy agents to mimic human-controlled avatars in an A-life environment. Such proxies would be machine learned from traces of actual human behaviour. These proxies could substitute for humans within the evolution for a set number of generations, with human-controlled avatars periodically re-introduced into the A-life environment where their behaviours are recorded to train the next set of proxies.

In this paper we present an update on this on-going project, including early proxy-creation results and the use of our approach in an interactive art installation.

2 Problem Formulation

The specific problem we focus on in this paper is to accelerate human-informed evolution of AI agents in an A-life setting. Our goal is to maximize the resulting reduction in evolution time while minimizing the loss of accuracy of the resulting AI behaviours relative to the baseline (i.e., using actual humans throughout the evolution).

To illustrate, consider evolving non-playable characters (NPCs) for a video game such as *Darwin's Demons* (Soule et al., 2017) or *No Man's Sky Next* (Hello Games, 2018). Suppose doing so with actual humans controlling avatars takes 100 hours for the NPCs to evolve to respond to the player in a desired fashion. Alternatively, we can substitute proxy agents for actual humans, reducing the evolution time down to, say, 10 hours (a 10x speed-up). However, since these NPCs were evolved against human proxies (i.e., *approximations* of humans), their evolved behaviour may be different from desired behaviour achieved in the former case. Our goal is thus to increase the speed-up while keeping both evolved response behaviours close.

3 Related Work

Surrogate models replace the actual, expensive fitness function with a more tractable approximation. Such models can be machine-learned (Kim et al., 2014; Rawal and Miikkulainen, 2018). Then evolution can be run with the surrogate model instead of the actual fitness function. This approach does not directly apply to our problem since the type of evolution we conduct in an A-life environment is asynchronous and does not have discrete generations on which a fitness function is used to sort the population towards forming the

Downloaded from http://direct.mit.edu/isal/proceedings-pdf/isal2019/31/465/1903467/isal_a_00204.pdf by guest on 25 June 2022

next generation. Instead, A-life agents reproduce at will and the fitness of an agent is implicit in the environment (Ackley and Littman, 1991; Bulitko et al., 2017).

4 Learning Human Proxies

We are implementing the approach proposed by Bulitko et al. (2018) in a simple Netlogo environment (Wilensky, 1999). The environment is a modification of our previous work (Bulitko et al., 2017; Soares et al., 2018) and involves predator and prey AI agents on a two-dimensional 48×48 rectangular grid (Figure 1, left).

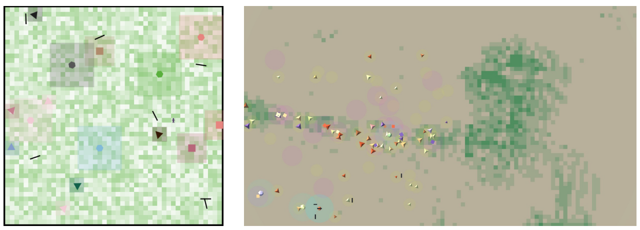

Figure 1: Our A-life environments: in Netlogo and in Unity.

Herbivores vary in shape and colour. They consume grass, shown as green cells. Predator agents consume herbivores and are visualized as short black line segments. A single human avatar is controlled by a player via keyboard. On each tick, the avatar movement along with a description of the 3×3-cell neighbourhood centred on the avatar are recorded. Specifically, we record 99 numbers which represent the presence of obstacles, amounts of grass, and the numbers of each type of AI-controlled agents in each of the 9 cells (11 numbers for each cell). We then record the avatar's action (i.e., moving to one of the 8 neighbours or staying put).

A run of N time ticks yields two matrices: the avatar's inputs ($N \times 99$) and the avatar's actions ($N \times 1$), which constitute the training data from which we learn a proxy to mimic the avatar's actions on a previously unseen run of M ticks. Such a run yields test data with $M \times 99$ inputs and $M \times 1$ actions. We then calculate proxy accuracy as the percentage of rows in the test input matrix for which the proxy gives the actual action taken by the human avatar.

As a proxy, we use a memory-free linear model that computes a scalar utility of each of the 9 cells as a dot product of the 11 numbers describing the cell and 11 proxy parameters (i.e., a multichannel convolution with a $11 \times 1 \times 1$ kernel). The proxy agent then moves to the cell with the highest utility. The 11 proxy parameters are found via a simple genetic search with the fitness computed on the training data. Preliminary results suggest a proxy test accuracy of 78 to 91% after training on matrices containing 600 to 4000 rows.

5 Future Work & Conclusions

We are concurrently implementing a similar A-life environment using Unity game engine (Unity Technologies, 2018) for an interactive art installation *Dyscorpia: Human in the Loop* to run in Edmonton, Alberta, Canada in April 2019: `https://www.dyscorpia.com`. Here, instead of a single keyboard-controlled avatar, multiple people are tracked with a camera as they walk around the installation. Their movements are painted into the A-life environment as grass. The right panel in Figure 1 shows a profile of a single human pointing to the left with his arm. We plan to record human movement data in the day time, use it to construct human proxies in the evening, run evolution of AI agents with the proxies overnight, and release the evolved agents in the installation next morning.

In conclusion, this paper presents an update on our ongoing project (Bulitko et al., 2018). Preliminary results with even rudimentary proxy learning are promising.

6 Acknowledgments

We appreciate support by The Kule Institute for Advanced Research, Social Sciences and Humanities Research Council and Edmonton Arts Council.

References

Ackley, D. and Littman, M. (1991). Interactions Between Learning and Evolution. *Artificial Life II*, 10:487–509.

Bulitko, V., Carleton, S., Cormier, D., Sigurdson, D., and Simpson, J. (2017). Towards positively surprising non-player characters in video games. In *Proceedings of the Experimental AI in Games (EXAG) Workshop at the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, pages 34–40.

Bulitko, V., Walters, M., Cselinacz, M., and Brown, M. R. (2018). Evolving NPC behaviours in A-life with player proxies. In *Proceedings of the Experimental AI in Games (EXAG) Workshop at the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*.

Hello Games (2018). No Man's Sky Next.

Kim, Y.-H., Moraglio, A., Kattan, A., and Yoon, Y. (2014). Geometric generalisation of surrogate model-based optimisation to combinatorial and program spaces. *Mathematical Problems in Engineering*, 2014.

Rawal, A. and Miikkulainen, R. (2018). From nodes to networks: Evolving recurrent neural networks. *CoRR*, abs/1803.04439.

Soares, E. S., Bulitko, V., Doucet, K., Cselinacz, M., Soule, T., Heck, S., and Wright, L. (2018). Learning to recognize A-Life behaviours. In *Poster collection: The Annual Conference on Advances in Cognitive Systems (ACS)*.

Soule, T., Heck, S., Haynes, T. E., Wood, N., and Robison, B. D. (2017). Darwin's Demons: Does evolution improve the game? In *Proceedings of European Conference on the Applications of Evolutionary Computation*, pages 435 – 451.

Unity Technologies (2018). Unity Engine.

Wilensky, U. (1999). Netlogo.

Downloaded from http://direct.mit.edu/isal/proceedings-pdf/isal2019/31/465/1903467/isal_a_00204.pdf by guest on 25 June 2022