
Behaviour Aesthetics of Reinforcement Learning in a Robotic Art Installation

Sofian Audry
Université du Québec à Montréal
Montréal, QC (Canada)

Rosalie Dumont Gagné
Cégep régional de Lanaudière à Joliette
Joliette, QC (Canada)

Hugo Scurto
EnsAD—PSL Research University
Paris (France)

Abstract

This paper presents our practice with reinforcement learning (RL) in the context of a robotic art project. We adopted a design-oriented approach to RL, privileging embodied robot behaviours generated by interactive learning processes over offline evaluations of optimal agent policies. We describe aesthetic and technical issues surrounding our experiments with a spheroid robot, and share our reflections on the potential of using RL for behaviour aesthetics.

Reinforcement learning (RL) traditionally adopts an engineering-oriented approach to creativity and design [1,2,3]. Such an approach typically applies RL to a pre-defined creative task, optimising parameters of a complex model—*e.g.*, deep reinforcement learning—over large amounts of GPU-simulated training data [4,5]. However, recent work showed that design-oriented approaches, focusing on interactive learning in “small data” use cases [6], could foster alternative creative uses of RL [7]. This paper presents our practice with RL and reflects on such design-oriented approaches in the use case of a robotic art project, called *Morphosis*. In this project, we are exploring the use of the *learning process itself* to generate robotic behaviours in an art installation [8]. In so doing, we are inverting the traditional mode of operation of RL—where the learning behaviour of RL agents is often accessory to the evaluation of their optimal policy.

Our work is an attempt to use RL as an approach to *behaviour aesthetics*, a growing field within new media arts that focuses on the behaviours of artificial systems as an artistic material [9]. As an interdisciplinary group working across art and engineering, we are interested in the perceptual, sensitive and emotional potentials of adaptive behaviours as produced by robotic agents as they are learning in real time. What are the specific challenges artists face when designing robotic behaviours with RL? What kind of aesthetic experiences can be created using such behaviours?

The *Morphosis* project rests on three similar spheroid robots (see Figure 1). Robots are able to freely move using a pair of servo-motors (one for traction and one for steering) and can sense their own spatial location and rotation relying on an inertial measurement unit and a real-time location system. Custom-designed silicon “skin” wrapped around each robot gives them different shapes.¹

Initial experiments with RL followed an engineering-oriented approach. We created a 3D simulation of the robot to train a deep-Q network over large amounts of high-dimensional, continuous synthetic data, without explicitly designing a state space. Yet, it quickly became clear that the material irregularities of the robot would eclipse the optimal learning performance of the RL agent; and that

¹One of the end goals of the project is to explore how a common RL model would generate different behaviours in the three robots due to the different shape of their “skins”. However, this article focuses on only one robot.

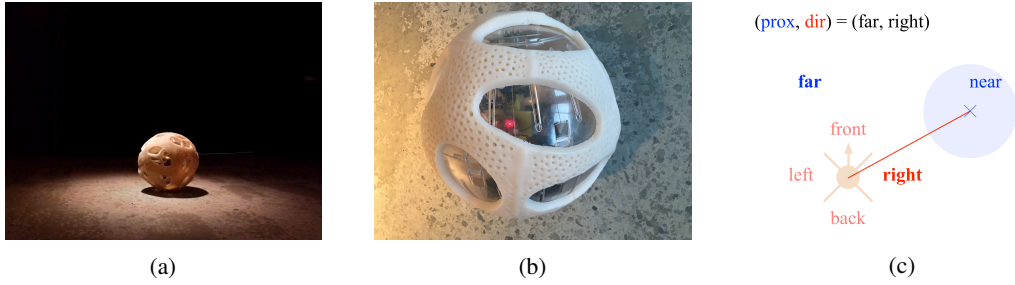


Figure 1: A spheroid robot’s (a) material, (b) technical, and (c) learning configuration in *Morphosis*.

the simulation did not give an accurate sense of the actual embodied experience of witnessing a robot moving in the physical world. We thus adopted a design-oriented approach to RL, tackling real-time learning from small amounts of material data, and focusing on learning behaviours embodied and revealed by the actual robot, as motivated by the artwork.

In order to make the robot’s learning process perceptible by viewers, we attempted to generate behaviours that could evoke the struggles of a living being in its attempt to reach intelligible goals such as remaining motionless, moving towards a position, or fleeing away from it. We considered all affordances provided by RL models along our creative process. We conceived low-dimensional *state spaces* in close link to the behavioural modalities we wanted to focus on, crafted *reward functions* based on our embodied intuitions of how the robot would learn, and employed *action policies* to control the agent’s exploration-exploitation trade-off during its learning process. We also experimented with *curiosity mechanisms* to provide the agent with motivational behavioural traits [10]. Finally, we relied on tabular *learning algorithms* to account for interactive learning in small state spaces. Interestingly, we quickly noticed that the choice of a particular learning algorithm did not have much impact on the robot’s behaviour. Rather, state space and reward function design were crucial to experiment with a diversity of embodied experiences related to learning processes.

One of our RL models, designed to progressively learn to attain a target position in the space, demonstrates this approach. Here, the agent state consisted in the robot’s discretized polar coordinates relative to the target (*prox, dir*), with *prox* being *near* when the robot is in a target’s neighborhood, and *far* elsewhere; and *dir* being front, back, left, or right (see Figure 1c). Agent actions were interspersed with short stabilising pauses to reduce noise from physical robot movement data. The reward function punished each action that moved the robot away from the goal.

Despite the model’s apparent simplicity, we observed rich behaviour morphologies [8], typically emerging over five to ten minutes of learning. Crucially, these were not directly linked to the RL task: for example, moving back and forth, standing still, approaching the target, eventually staying close to it, then slowly oscillating, stepping back, coming close again, etc. Complementary to aesthetics, the target-centric model also presented technical advantages, supporting real-time learning from a single standard CPU, and transferable by design to any target position.

Reflecting on our artistic practice, we found that designing RL models provided us with more control over behaviour generation than training deep neural networks, while at the same time fostering more machine autonomy than direct programming. Conversely, it also brought us closer to original use cases of artificial intelligence relying on tabular models and small data [11]. The novelty of our approach lies in considering *reinforcement learning processes as behaviour generators on their own* in the context of robotic art. While engineering-oriented approaches have so far considered RL errors as a bottleneck for real-world applications, robotic art practice allows to consider the phenomenology of RL as a vector of expressive behaviours that echoes our own errors and hesitations as humans, thus opening space for projection, affects and empathy in the viewer. Contrary to a behaviour entirely programmed to mimic the living, our approach allows the emergence of unpredictable and awkward movements that change through time.

We will soon exhibit *Morphosis* to study audiences’ phenomenological perception and experience of RL behaviours. By then, we hope that our reflections on design-oriented approaches to RL will resonate with other researchers and practitioners in the area of machine learning for creativity and design.

Acknowledgments and Disclosure of Funding

Funding in direct support of this work: Canada Council for the Arts, University of Maine, Clement and Linda McGillicuddy Humanities Center. Additional revenues related to this work: Titan Xp GPU donated through the NVIDIA Academic Grant Program. The authors would like to thank Pierre Gaudet, Martin Peach and Drake Perkins for their assistance on this project.

References

- [1] Guss, W. H., Houghton, B., Topin, N., Wang, P., Codel, C., Veloso, M., & Salakhutdinov, R. (2019). MineRL: A large-scale dataset of Minecraft demonstrations. *arXiv preprint arXiv:1907.13440*.
- [2] Chan, M. T., Gorbet, R., Beesley, P., & Kulič, D. (2015). Curiosity-based learning algorithm for distributed interactive sculptural systems. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 3435-3441). IEEE.
- [3] Schmidhuber, J. (2006). Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connection Science*, 18(2), 173-187.
- [4] Maekawa, A., Kume, A., Yoshida, H., Hatori, J., Naradowsky, J., & Saito, S. (2019). Improvised Robotic Design with Found Objects. In *NeurIPS Workshop on Machine Learning for Creativity and Design*.
- [5] Rojas, J., Coros, S., & Kavan, L. (2019). Deep reinforcement learning for 2D soft body locomotion. In *NeurIPS Workshop on Machine Learning for Creativity and Design*.
- [6] Fiebrink, R., & Sonami, L. (2020). Reflections on Eight Years of Instrument Creation with Machine Learning. In *Proceedings of the 20th International Conference on New Interfaces for Musical Expression (NIME)*.
- [7] Scurto, H., Van Kerrebroeck, B., Caramiaux, B., & Bevilacqua, F. (2019). Designing deep reinforcement learning for human parameter exploration. *arXiv preprint arXiv:1907.00824*.
- [8] Audry, S. (2019). Behavior Morphologies of Machine Learning Agents in Media Artworks. *Leonardo*, 1-10.
- [9] Penny, S. (2017). *Making sense: Cognition, computing, art, and embodiment*. MIT Press.
- [10] Gemeinboeck, P., & Saunders, R. (2013). Creative Machine Performance: Computational Creativity and Robotic Art. In *Proceedings of the Fourth International Conference on Computational Creativity (ICCC)*(p. 215).
- [11] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.