Performative 3D Agents leveraging Reinforcement Learning in *the Fold*

Alex M. Lee
Arizona State University
Phoenix, Arizona USA
Alex.M.Lee@asu.edu

Abstract

The Fold: Episode II involves several innovative implementations of tools for a subset of machine learning called reinforcement learning applied within video game engines. Through goal-based training using positive and negative rewards, the artist demonstrates the potential of leveraging this training-based method of creating computer animations towards 3D character performances within a virtual-reality based art video game that are unique, metaphorically tied to the processes that underlie the core of machine learning, and addresses issues surrounding automation, labor, and craft.

Keywords

Machine Learning, Reinforcement Learning, Virtual Reality, Artificial Intelligence, Artifice, Human-Machine Collaboration

Introduction to Reinforcement Learning

Figure 1. Reinforcement learning training (left) with its corresponding training environment (right) ©2022.

Reinforcement learning is a form of machine learning involving agents that take actions in an environment based on limited perceptions. Their actions result in a new state as well as a reward or punishment fed back into the agents. Through reinforcement learning, the agents adapt to their environment by learning how to optimize their actions in order to maximize their reward over time [1].

The development of reinforcement learning frameworks connecting Tensorflow (an end-to-end machine learning platform) to video game engines such as Unity and Unreal has presented the opportunity for artists to leverage the potential of reinforcement learning to create synthetic 3D agent-based performances. These performances, depending on the training involved, allows the agent to exhibit some level of autonomous behavior by dynamically responding to changes in their environment and the user.

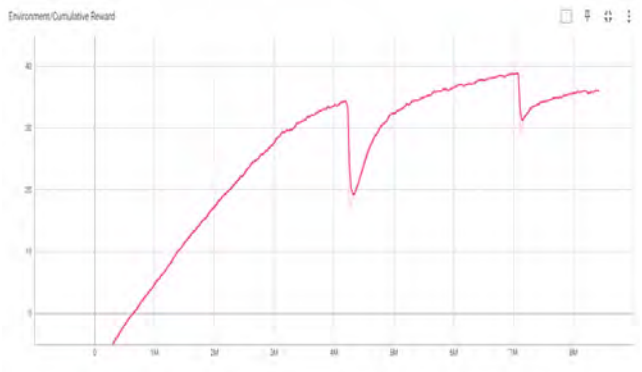

Figure 2. Tensorboard results of reinforcement learning training session. Note that time is counted as millions of steps (X-axis). The dips in cumulative reward (Y-axis) are gradient descent steps which happens when an agent tries something new (by taking a small risk) in the training process to optimize the learning process even further which results in a dip in its reward state ©2022.

This process of having custom 3D characters learn to animate on their own was a compelling prospect and in my research found that it could provide innovative and compelling animations that are expressively tied to the processes inherent to machine learning. To me, these animations are uncannily machine-like with behavior a traditional animator would not have crafted through classical keyframe animation. Additionally, having multiple agents exhibit optimized autonomous and cooperative group behavior for XR devices was a compelling possibility to exploit.

Training 3D agents has a historical precedent set by Karl Kims' *Evolved Virtual Creatures*, 1994 [2]. These first virtual creatures were trained using custom made physics engines (nothing like the video game engines we have access to today) and took months of training. Current day RL frameworks cuts training time exponentially as well as leverage the existing physics and high-end look development of video game engines. Additionally, more complex characters are afforded by advancements in computer power.

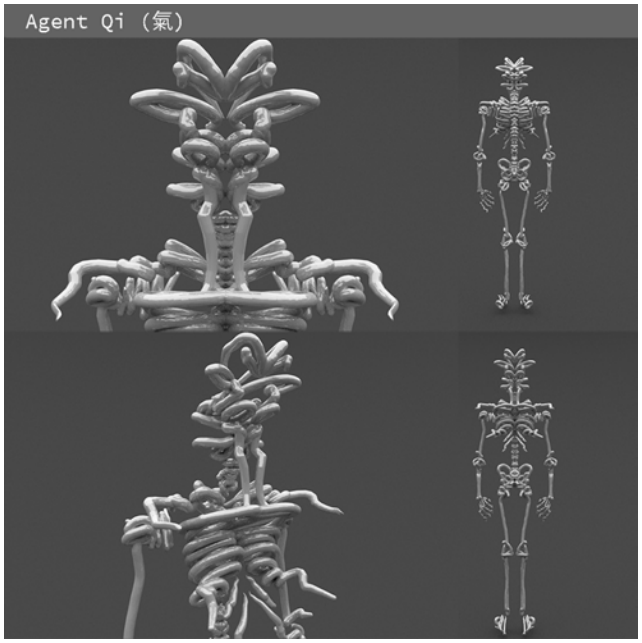

Figure 4. ***Agent Qi/氣*** character design ©2022

Figure 3. ***Agent άλφα*** character design ©2022

During my research process I had the opportunity to view the various rock gardens in Kyoto, Japan. I realized during my site visits that these various rock garden designers throughout history were doing physical level designs over a long period of time. Formal rock garden designs date back all the way to the Heian Period (794 – 1185) [3]. These gardens were meditative spaces which allowed viewers to enter an alternate state of mind to reposition your Qi or energy. I make a connotative link between the synthetic agent training in virtual rock garden environments in a subsequent section of this paper. For now, I will explain that I was interested in presenting rock garden environments that were impossible to achieve in the physical world.

Training a 3D agent on a 3D environment within a video game engine typically requires sensors to observe its environment. A sensor could be something like synthetic vision (a rudimentary method of sight using raycasting). Each agent has a given number of sensors which total the amount of vector observations taking place within a scenario. Some vector observations could be done by the agent while others are taking place by the environment sensors. The total amount of vector observations determines the complexity of the training. The more vector observations, the more complex the training sessions becomes.

Once training starts, the agent performs an action using the Multi-Layer Perception (MLP) structure, essentially learning its environment by trial and error and remembering its actions with the goal of achieving a positive reward during each step [4].

The image in figure 1 shows a training session which involves an agent, *Agent **άλφα***, attempting to caress the synthetic representation of a pipe with its body within a 3D simulated low-gravity environment. The training session is within Anaconda and each line represents an epoch which contains 10,000 training iterations. At each line, the mean reward state is reported which shows essentially how much the agent has learned up until that point.

Figure 2 represents a diagrammatical visualization of the reward state in Tensorboard which shows the number of steps involved before learning exponentially increases. Depending on the complexity of the environment, the agent, and the task, it could take anywhere from a few hours to several days to train on a scenario.

It was fascinating to see the Tensorboard graphs evolve as I watched these agents learn. The agents go from not knowing anything about their digital bodies to fully navigating their 3D environment over several hours (which can be seen in real-time)! Some of the training footage I have included in subsequent links next to each learning agent as written descriptions does not do justice to the actual footage!

The Reinforcement Learners

Digital Readymade Agents

Representing the first characters to be utilized by the artist within the reinforcement learning (RL) framework, there are 3D digital agents have the ability to jump, move along the Vector 3 space (x, y, z) by rotating and translating forward and back. There are six readymade agents altogether

comprising of the following characters*: the pipe, the urinal, the bottle rack, the teapot, the bicycle stool,* and *the pawn*. Each agent is trained individually and have unique constraints to their movement. Video footage of these agents exhibiting autonomous and cooperative behavior can be found in this link: https://vimeo.com/658880804

Agent *άλφα*

Agent άλφα is a humanoid RL character comprised of elements from the 'readymade agents'. The six agents in the previous description are the elements that comprise agent Alpha (see figure 3). Agent *άλφα* contains represents Eurocentric knowledge the details of which reveals itself either symbolically, metaphorically, or in narrative form throughout the experience. Representing 212 vector observations, this character is one of the most sophisticated characters in the art game and takes several aggregated hours to train to walk using reinforcement learning. Video footage of this training process can be found in this link: https://vimeo.com/697348740

Agent Qi/氣

Agent Qi/氣 is a humanoid RL character sculpted in a VR-based sculpting software called Oculus Medium. The character is made to mimic the Chinese character Qi (氣) in its form. The term Qi comes as close as possible to constituting a generic designation equivalent to the Western word 'energy'. When Eastern thinkers are unwilling or unable to fix the quality of an energetic phenomenon, the character Qi inevitably comes to mind [5]. This character also uses 212 vector observations. Video footage of this agent's process of learning to walk can be found in this link: https://vimeo.com/766984704

Agent Marc/Gritte

Agent Marc/Gritte is the only quadruped RL character in the game and is also comprised of elements from the 'readymade agents'. Marc/Grittes have been trained to catch the player in various levels of the game. Agent *Marc/Gritte* has 126 vector observations. Video footage of this agent adaptively responding to obstacles in their environment can be found in this link: https://vimeo.com/760071134

Mudra Agents

In one level of the game, there are agents who have learned to balance a golden orb using their hands. Originating from Buddhist religion, each agent represents a mudra for emotional wellness including: concentration, energy, stability, and acceptance. Footage of one of these agent's learning can be found in the following link. It requires multiple instances of the same agent to maximize the learning process within a given training session: https://vimeo.com/766999820

For me, it is quite interesting to see these 3D agents responding to you as the viewer and their environment with a heightened sense of presence VR affords. Seeing them at eye-level as if they are in front of you creates a different experience compared to through a screen or projection.

The Game: *the Fold: episode II*

The VR-based art game '*the Fold: episode II*' involves a starting scene involving five gates. Inspired by the Great Buddha of Kamakura, each gate corresponds to an element of Zen referred to in the five elements (in English, Chinese, Korean, and Japanese respectively):

Earth 地 (지,ち): collectiveness, stability, physicality, and gravity.

Water 水 (수, すい): the fluid, flowing and the formless things in the world.

Fire 火 (화,か): the energetic, forceful, moving things in the world.

Wind 風 (풍,ふう): things that grow, expand, and enjoy freedom of movement.

Void 空 (공,くう): those things beyond and within our everyday comprehension.

Figure 5. Lobby scene involving five gates labeled with an element of Zen ©2022

Each level metaphorically corresponds to each of the five elements at the starting scene (figure 5). All of these levels are presented as white gates in the starting scene. Several of them will be described subsequently. A trailer for the game can be found in this link: https://vimeo.com/791490196

The Levels / Environments

Rock Gardens

There are various rock gardens within the game which can be found through the 'water' and 'wind' gate as they involve agents that are fluid and flowing as well as enjoying freedom of movement. The connotation of the rock garden being one of artifice and a stand-in for nature within a construct. A user is placed within each rock garden competing with the agents to get to the next level. Upon observing the agents, the viewer will notice that agents will adaptively respond to the

viewer's movements and also exhibit collective cooperation to achieve tasks.

Simulated Low Gravity Environments

There are a number of simulated low gravity environments within the game. This presents the agents a level of difficulty in achieving tasks. Two levels involve suspended objects that the agent must caress with its digital body. The viewer is also placed within this environment. Training on this type of scenario would be very difficult to do in real life, which is why it is ideal to do this type of scenario training digitally within a 3D environment.

Stairs

There are levels where agents must climb stairs which presents a unique challenge to agents with limbs such as *Agent Qi*/氣, *άλφα*, and *Marc/gritte*. Certain agents such as Qi/氣 have figured out that jumping is more efficient than walking. This leads to a key point about the training process. After the agent has learned about its body, it finds unique solutions to challenges that has consistently surprised me. An agent figuring out something new like jumping rather than walking to be more efficient was a solution that I would not have thought would happen, but pleasantly surprised when it does. Its method of jumping is unique to a machine solution to a problem and aggregated footage of the training process is reminiscent of Duchamp's 'Nude Descending a Staircase' (a visual metaphor about the human body in the machine age) but in the opposite direction. See figure 6 below. Video footage of this process can be found in this link: https://vimeo.com/768162797

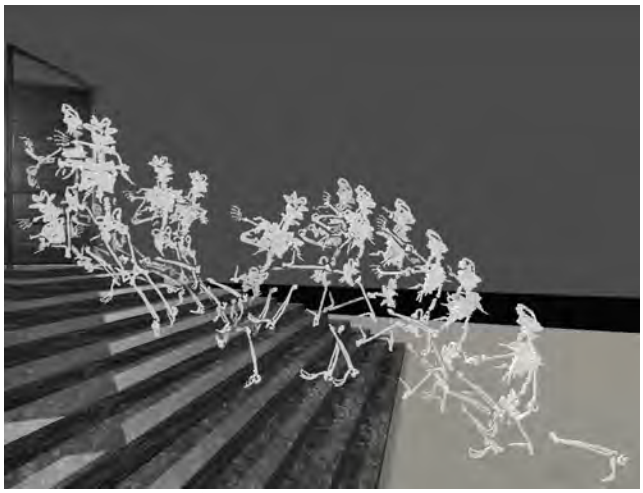

Figure 6. ***Agent* Qi/氣** climbing a staircase (aggregated footage) ©2022

Balancing

In certain levels, balancing in VR is involved. There is one level in particular where balancing is the main theme. Titled the *Temple of Balancers*, in this level there are the previously mentioned mudra agents whose task it is to balance golden orbs. You start on a platform within a shrine and you must get to the other end to get to the next level. This tests your sense of balance within the VR environment, but don't worry. If you fall off the platform you start again at the beginning of the shrine.

Sentience and Zen

Figure 7 shows the back-end of a given training session for a readymade agent. The white lines are sensors that represent the agents' field of view. This represents its synthetic vision. Note that there are white lines at the 'head' of the agent as well as at its base. This allows the agent to tell whether there are obstacles that are its height or taller to jump over. These white lines represent the agent's field of view as well as its line of sight. The terminal point of an agent line of sight ends where the white line ends. When an agent sees a task or goal it line of sight turns to red. The red sphere indicates that it has seen the target, in this case is a door that the agent must go through. At first the agent has no idea about its environment and falls off the T-shaped platform repeatedly until it figures out the bounds of the platform. Over time, it learns where to go to get to through the door. Usually, this requires several hours of training before it can reliably do so.

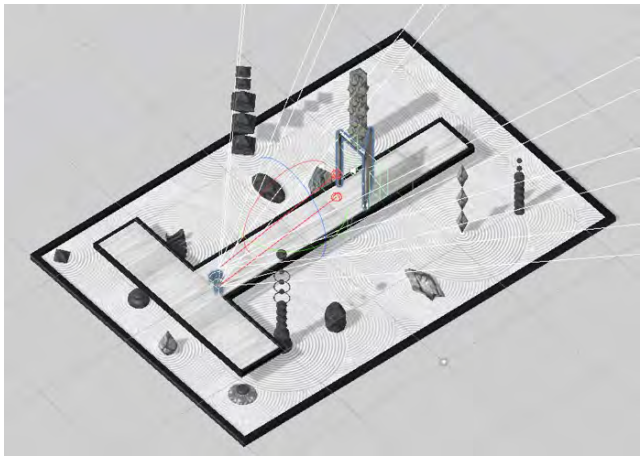

Figure 7. Rock garden training still involving a readymade agent ©2022

It is worth noting that Ryoanji Garden in Kyoto Japan (built sometime during the Muromachi Period in the 15th Century) is considered to be a quintessential example of a rock garden and represents the highest degree of intellectual refinement that was possible to attain for a rock garden. There have been many attempts at interpreting its meaning. These usually involved diagrams highlighting line of sight to objects within a given arrangement (see figure 8). These attempts privilege a Western Cartesian interpretation of rock gardens and attempts at an analysis imposing Cartesian rationality upon a system meant to imitate the essence of nature and reposition your Qi (氣) [6].

Figure 7 represents a 3D agent using Cartesian space to observe the synthetic representation of a rock garden. Two

fundamental philosophical views about sentience in friction playing itself out within a VR simulation. The environment of the garden itself evoking artifice and immateriality while VR experientially evokes the idea of letting go of one's corporeal 'ego' and centering your Qi (a very Eastern and Zen idea).

The agent training can suggest a rhizomatic approach to perception, a reference to Gilles Deleuze [7], by suggesting infinite vantage points and pathways which become more and more concrete as the training progresses. The rock garden itself suggesting letting go of all thoughts, distractions, and ego (agents have no ego other than the will to receive a reward). Eventually, a consensus of the T-shaped platform is numerically refined over time by training within the reinforcement learning framework.

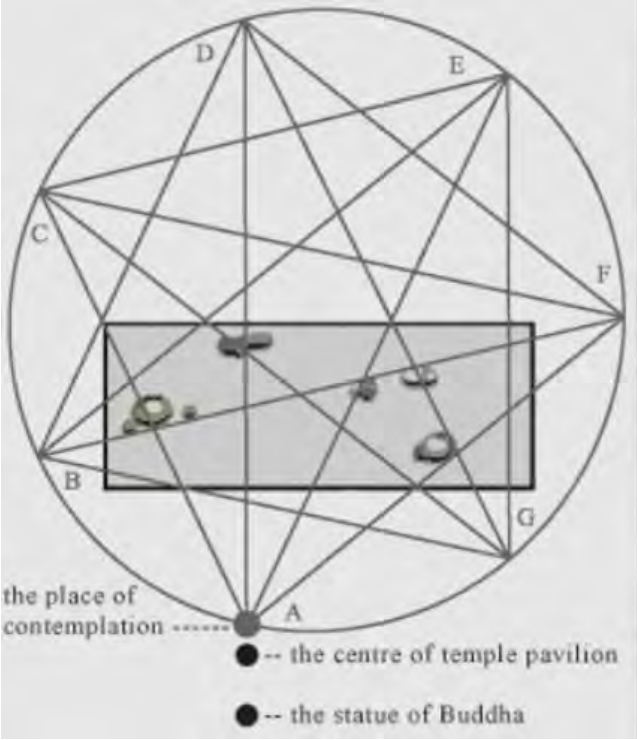

Figure 8. A diagrammatical attempt at analysis of Ryoanji Garden. Unknown author ©2022.

Labor and Craft

This method of producing animation represents a paradigm shift in the craft of creating animations for video games as well as the real-time method of making animated sequences. The potential here is that dynamic performances can be achieved by utilizing reinforcement learning with 3D characters. These agents can interpolate and respond to changes to its environment. The animations do not need to be hard coded and can rely on the neural model to interpolate the in between poses. One does not have to manually animate and can delegate these tasks to the computer. Additionally, depending on the scenario, agents can perform basic reasoning tasks which I have yet to fully exploit (but plant to in subsequent episodes). With all the potential, there are several challenges in this workflow that makes working with this framework difficult and labor intensive in of itself:
First, there is the requirements of computing which is the first barrier to entry. This process requires a Windows computer with enough compute strength to handle machine learning.

Second, there is the challenge of setting up the framework within Unity. Tensorflow needs to be installed through a Python environment such as Anaconda. Both of these need to be installed with the correct versions which match in order for both to communicate with each other. Unity is used to train the 3D environment while Tensorflow is used to train the neural modeling.

Third, getting the agent to learn requires trial and error. This usually requires scenario testing and training to see whether the agent is actually learning. These test training sessions can take a considerable amount of time depending on the complexity of the scenario which amounts to several hours or even several days. This initial-end labor means that the whole computing process of AI and the human worker never fully leaves the loop. The assembly line of labor involving animation shifts towards preparation and decision making at the front of the process.

Fourth, at the moment of publishing this paper, this workflow is still Beta and there are various bugs with operating the framework and its examples. Support is only up to 2020 version of Unity at the moment. Because of this, it will be difficult to use the latest render pipelines for high definition production.

Despite these challenges, the potential here is tremendous in regards to creating 3D character performances. I believe we will see some of these challenges resolve themselves in subsequent versions of this workflow or as new frameworks for reinforcement learning for 3D characters and environments become released that are more streamlined and does not require multiple software platforms communicating with each other. Due to the potential of finding new forms of animation utilizing this workflow which involve human-machine collaboration, these challenges are worth overcoming.

Historians of computation have already stressed the early steps of machine intelligence in the 19th century project of mechanizing the division of mental labor, specifically the task of hand calculation. The enterprise of computation has since then been a combination of surveillance and disciplining of labor with the optimal calculation of surplus-value and planning of collective behaviors [8][9].

Thus, as a process of automation, reinforcement learning will have an impact on the job market just as much as supervised and unsupervised deep learning. The impact of AI on labor is predicted by industry leaders such as Dreamworks

Figure 9. Larger version of Figure 1 from the first section of paper ©2022.

co-founder Jeffrey Ketzenberg's comments about the Animation Guild possibly striking in 2024. His support for AI creating feature length animated movies in the future, cutting 90% of labor. He stated, "In the good old days, when I made an animated movie, it took 500 artists five years to make a world-class animated movie. I think it won't take 10% of that. Literally, I don't think it will take 10% of that three years from now." [10]

Acknowledgements

The artist would like to acknowledge the feedback, guidance, and encouragement of Dr. Sofian Audry at the University of Quebec in Montreal whose expertise in reinforcement learning provided valuable insight into the development (and performance) of these 3D agents throughout to development process of *the Fold: episode II*.

References

[1] Richard Sutton, Andrew Barto, *Reinforcement Learning*, MIT Press, 1998
[2] Karl Sims, *Evolved Virtual Creatures*, https://www.karlsims.com/evolved-virtual-creatures.html 1994
[3] Mumonkan Heikiganroku, *Two Zen Classics*, trans. Katsuki Sekida, Blue Cliff Records, 1977
[4] Koki Mitsunami, 'Unity ML-Agents on Arm and How We Created Game AI', *Arm Community Blogs*, https://community.arm.com/arm-community-blogs/b/graphics-gaming-and-vr-blog/posts/2-unity-ml-agents-on-arm-how-we-created-game-ai#:~:text=By%20default%2C%20ML%2DAgents%20uses,as%20shown%20in%20figure%205, July 29, 2022
[5] Manfred Porkert, *The Theoretical Foundation of Chinese Medicine: Systems of Correspondence, 1974*
[6] Manfred Porkert, *The Theoretical Foundation of Chinese Medicine: Systems of Correspondence, 1974*
[7] Deleuze, Gilles, *The Fold: Leibniz and the Baroque*, trans. Tom Conley, p. 18, 1988
[8] Matteo Pasquinelli, 'On the Origins of Marx's General Intellect', *Radical Philosophy* 2.06, 2019
[9] Simon Schaffer, 'Babbage's Intelligence: Calculating Engines and the Factory System', *Critical Inquiry* 21, 1994. Lorraine Daston, 'Enlightenment calculations'. *Critical Inquiry* 21, 1994. Matthew L. Jones, *Reckoning with Matter: Calculating Machines, Innovation, and Thinking about Thinking from Pascal to Babbage*. Chicago: University of Chicago Press, 2016.
[10] Hope Mullinax, 'Dreamworks Co-Founder Thinks 90% of Animation Labor Could be Cut with AI', *Collider*, Nov 10th, 2023