Playing Jedi Academy with Curious A3C

Dmitri Tkatch

Abstract

A curious artificial intelligence learns to explore free-for-all maps and defend itself with a saber against built-in bots in duel mode in mere hours of learning on a modest 8GB GPU. It is based on six deep learning neural networks – a future predictor, an action predictor, a state embedder, an image embedder, a snapshot embedder and an A3C actor. The AI can use either raw game data, raw pixels or both as input. It uses a novel form of curiosity where unpredictable futures are positively rewarded while unpredictable actions are negatively rewarded. It defeats Cultist 3-1, Jedi Trainer 1-0 while starting at a disadvantage and the final boss of the game Desann 20-11. We note the importance of curiosity for maintaining a variety of play and our novel form of curiosity for maintaining smooth actions. Complete program is available upon request.

Videos – https://www.youtube.com/@eternalyze0.

Program Structure – https://github.com/Eternalyze0/jediknight.

Correspondence – eternalyze0@gmail.com

1 Introduction

Star Wars Jedi Knight Jedi Academy is a 3D competitive saber-combat game where players may choose from 3 sabers (one saber, dual sabers, staff), 3 saber styles (fast, medium, strong), 7 saber swings (left-to-right, right-to-left, bottom-left-to-top-right, bottom-right-to-top-left, top-to-bottom, top-right-to-bottom-left, top-left-to-bottom-right in each saber style), and 3 katas (one in each saber style). Players are able to move freely in 3D with the WASD keys and look freely with the mouse, even during a saber-swing. Players can also jump with space and switch saber styles with L. Swings are chosen with the direction keys. For instance a bottom-right-to-top-left swing corresponds to pressing the AS keys. Although published in 2003 Jedi Knight Jedi Academy still sees competitive play thanks to its deep combat system. An experienced player may be victorious 10-0 against a novice player. A professional player may be victorious 10-0 against an experienced player.

2 Curiosity

An artificial intellgience is curious if part of its rewards correspond to errors in the future prediction model. Although there are many papers for 2D or toy 3D environments, there are relatively few papers for realistic (3D, realtime, competitive) games. Part of the reason may be because of the exploration challenge introduced by an additional dimension. In particular, this is the first time curiosity has been applied to a realtime 3D game that sees competitive play. Moreover, the artificial intelligence learns to defend itself against built-in bots in mere hours while maintaining a variety of play. In particular, while artificial intelligences without curiosity tend to stagnate and get stuck in a mode of behavior, a curious artificial intelligence maintains a variety of play.

3 Program

Our inspiration and implementation is based on Deepak Pathak et al's Curiosity-driven Exploration by Self-supervised Prediction [1], Tim Pearce et al's Counter-Strike Deathmatch with Large-Scale Behavioural Cloning [2], and Seungeun Rho's implementation of the A3C [3]. In particular, we use the intrinsic curiosity module from [1], the interfacing and mouse movement bucketing ingenuity of [2] and the clean program of [3]. We perform experiments without curiosity, with the intrinsic curiosity module and with a modified smooth version of the intrinsic curiosity module. Without curiosity the artificial intelligence easily gets stuck in some corner of the map and in a repetitive pattern of behavior. With curiosity the artificial intelligence maintains a variety of play and exploratory behavior, preventing it from ever getting stuck. With smooth curiosity the artificial intelligence also never gets stuck but also behaves in more human-like, less jittery way.

4 Motivation

Intrinsic motivation is dedicated to unsupervised A3C learning. That is, instead of rewards coming from the environemnt the artificial intellignece is free to choose which goals it would like to undertake. The intrinsic curiosity module not only rewards error in the future prediction model (which predicts the future state given the present state and action) but also employs an coordination model (which predicts behavior given a state transition). Moreover neither the future nor coordination models operate states directly but rather utilize a neural state compression module. The coordination model backpropagates errors to the neural state compression module.

5 Results

We name our artificial intelligence Sapphire in reference to the dragon Sapphira from Eragon by Christopher Paolini, the Phi neural module from Pathak et al's paper, and the gemstone skin of the blue shadowtrooper from the Jedi Knight game. See videos,

- https://www.youtube.com/watch?v=folA8Rqyi20

- https://www.youtube.com/watch?v=luNo1oCSoHw
- https://www.youtube.com/watch?v=9vLv0ws3hoU

Sapphire runs on one 8GB graphics processing unit, explores half of Tattooine in half an hour, adequately defends iteslf against built-in bots, and learns to do all this in mere hours.

6 References

We thank the authors of these references for their curious dedication as well as friends and coworkers who helped with interfacing with the game and editing the paper,

[1] Deepak Pathak, Pulkit Agrawal, Alexei A. Efros and Trevor Darrell. Curiosity-driven Exploration by Self-supervised Prediction. In ICML 2017.

[2] Tim Pearce, Jun Zhu. Counter-Strike Deathmatch with Large-Scale Behavioural Cloning, In proc. of IEEE Conference on Games (CoG), Beijing, China, 2022.

[3] Seungeun Rho. A3C. 2019.