Breakout Bot: Initial Project Proposal

Angela Jiang
Northwestern University
2317 Ridge Avenue
Evanston, IL
(616)635-9604
angelajiang2014
@u.northwestern.edu

Motoki Mizoguchi
Northwestern University
1116 Garnett Place
Evanston, IL
(859)230-0662
motokimizoguchi2013
@u.northwestern.edu

Max New
Northwestern University
1930 Ridge Avenue
Evanston, IL
(985)397-1770
maxnew2013
@u.northwestern.edu

1. BACKGROUND

We will create an AI bot for the arcade game Breakout, that learns how to beat a level as quickly as possible. We want to further explore reinforcement learning techniques outside of the scope of the class. Previous works have already created AI bots for real time games using reinforcement learning. However, we want to create to create a bot that not only can win a level of Breakout but that can win it quickly. This seems like an appropriate challenge to try reinforcement learning.

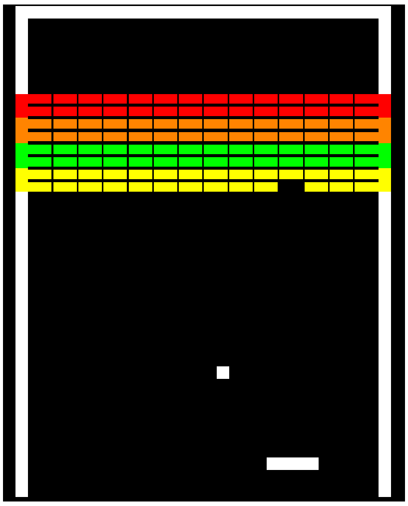

Figure 1: Breakout game

2. APPROACH

2.1 Idea

We will construct a live learner that will use its own game sessions as training data. We will be using a javascript breakout game to learn the ways to play. The 3 components of the game are bricks, the ball and the player. We will extract these components directly from the game runtime. The state of the game will be the number of bricks remaining in each section of the board and the location and velocity of the player and the ball. Given the state of these components, the bot will predict the next move. The bot will have the following inputs and outputs:
Input:

- Number of bricks left in section of board
- Ball location
- Ball velocity
- Player location
- Player velocity

Output:

- Movement (move left, move right, stay still)

Every epoch, the bot will move based on the current state and policy. After the ball enters the plane of the paddle, we will update the policy we used to decide the move. There are two stages to our project: learning with and without bricks. First, our bot will learn to hit the ball back, without the goal of beating a level. Then, starting with this resulting policy, we will introduce bricks into the game state. In the first stage, we will reward the bot for hitting the paddle, and penalize it for losing the game. In the next stage, there is no longer a reward for hitting the paddle. Instead, we will give a discounted cumulative reward for hitting bricks. This will motivate the AI to maximize its reward in the smallest number of moves.

2.2 Evaluation

Instead of comparing to another breakout bot, we will be evaluating the evolved bot over time starting at a random set of weights. Therefore our baseline approach will be the performance of an initial policy. Our first goal is to create a learner that will never lose the game, i.e., it always keeps the ball in play. For our second stage, we want to optimize the number of moves it takes to complete a level. We will compare the number of moves to complete a level using an initial and future policy.

3. WORK PLAN

3.1 Division of Work

We have divided the project into components that we estimate to take roughly equal effort. To maintain accountability one person will be designated owner of each component.

We also assign each component an assistant that will work with the project owner so that each group member gets to learn from multiple parts of the project.

Angela will implement the core reinforcement learning algorithm itself. She will also be responsible for aggregating learning data and Max will assist in both of these tasks. Motoki will work on random level generation and hooking our learner into an existing implementation of the Breakout game, with Angela as assistant. Finally, Max will design and implement a website with which users can interactively run our learner and watch it play and improve. The website will also feature our results and describe the project. Motoki will assist Max on the website.

3.2 Milestones

Milestone	Group Member(s)	Date
Website Skeleton	Max, Motoki	November 23
Implement Hand-written Bot	Motoki, Angela	November 23
Running Pong Learner	Angela, Max	November 23
Website Demonstrates Learner	Max	November 30
Running Breakout Learner	Angela, Motoki	November 30
Collect Evaluation Data	Angela	December 7
Complete Website	Max	December 7
Print Poster	Motoki	December 7

4. REFERENCES

[1] Bellemare, M. G., Naddaf, Y., Veness, J., Bowling, M. (2012). "The arcade learning environment: An evaluation platform for general agents." *arXiv preprint arXiv:1207.4708.*

[2] Togelius, J., Karakovskiy, S., KoutnŠk, J., Schmidhuber, J. (2009, September). "Super mario evolution." *In Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on* (pp. 156-161). IEEE.

[3] T. G. Tan, J. Teo, P. Anthony, J. H. Ong. "Neural network ensembles for video game AI using evolutionary multi-objective optimization." HIS, page 605-610. IEEE, (2011)

[4] L. Fausett, "Fundamentals of Neural Networks Architectures, Algorithms, and Applications", Prentice-Hall, Inc, 1994.

[5] R. Graham, H. McCabe, S. Sheridan, "Neural networks for real-time path-finding in computer games" `http://www.gamesitb.com/nnpathgraham.pdf`

[6] R. Graham, H. McCabe, S. Sheridan, "Neural Pathways for real-time dynamic computer games." `http://gamesitb.com/Graham_EGIreland.pdf`