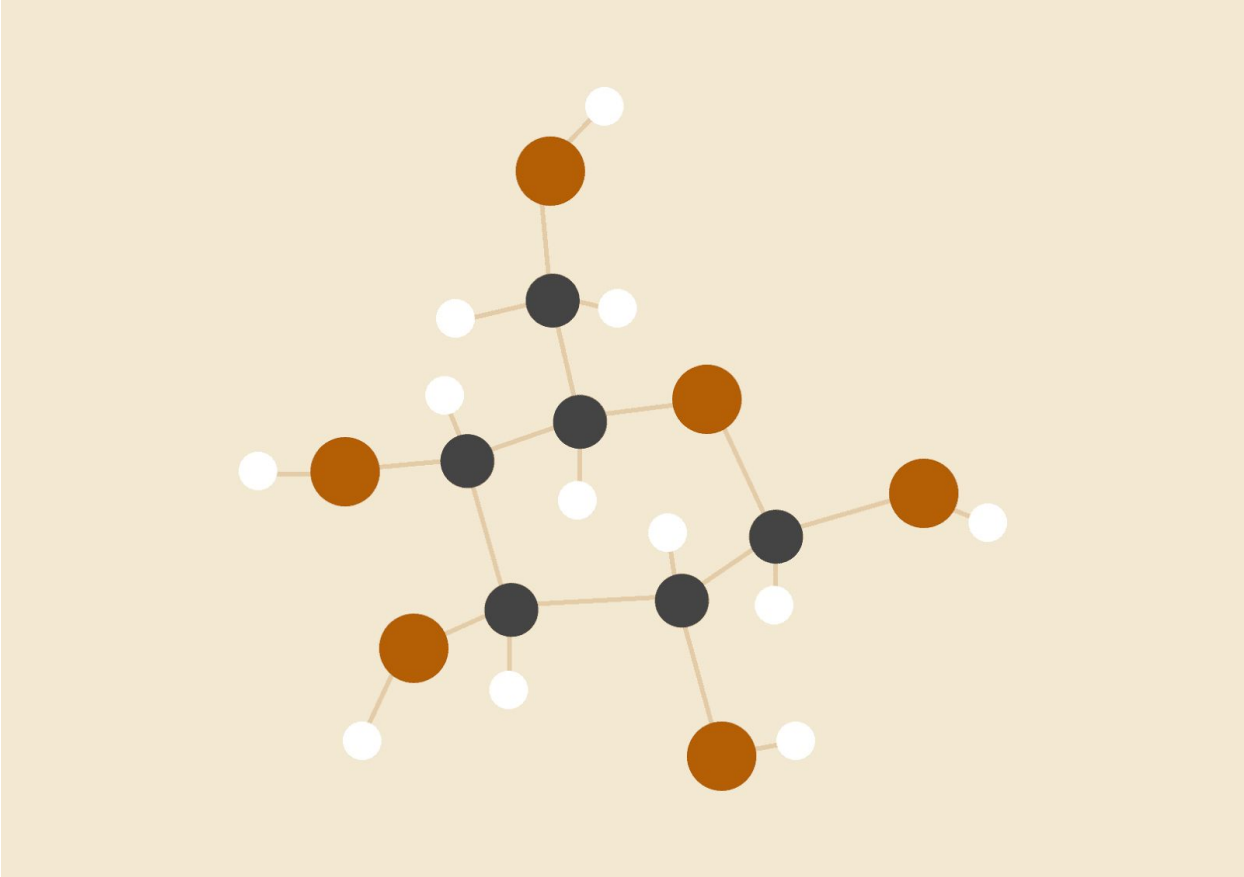
DEEP LEARNING

Q Learning Randomly Spawned Rewards



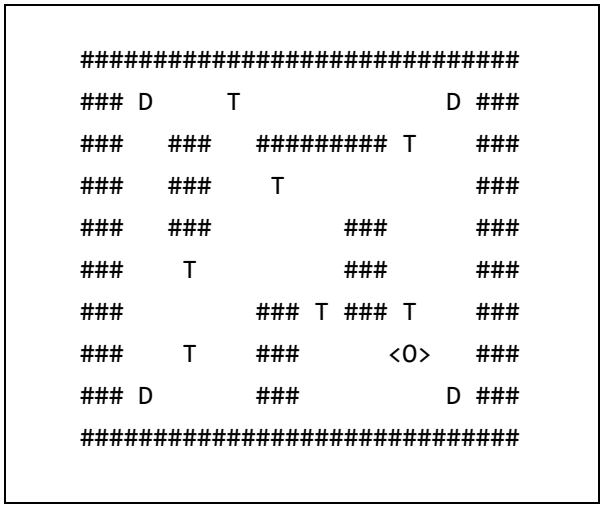
Michael Harrington

12.13.2017

CS 500

INTRODUCTION

This project is an investigation and implementation of using Q Learning to teach an agent to maximize its reward in a static environment with randomly spawning rewards. This investigation simulates a fictional character's, Tinny Tim's, basement and the randomly spawning donuts. The environment is setup such that there can only be one spawned donut at a time, if no donut there is 25% chance of a donut spawning, there is a 50% chance of a falling tile hazard on hazard locations, walls that prevent movement, and only an 82% chance of taking the desired Q Learning action. Donuts are a reward of 10, falling tiles a punishment of 10, and walking into walls a punishment of 1. Below is a depiction of the basement using the ascii printout of the developed program.



CONFIGURATION

Located in the configs/ folder is a file called config.json. It specifies the following Q Learning configuration parameters.

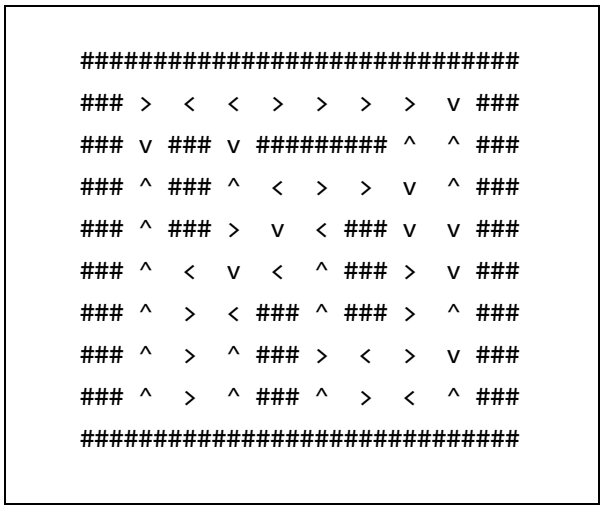
1. Simulated steps: 100,000
2. Learning rate (alpha): 0.15
3. Discount factor (gamma): 0.75

DATA

A step by step history is located in logs/basement_log.txt. A printed output is located in logs/output_capture.txt.

RESULTS

Below is the final policy network visualized.



The bottom left of the above visualization is coordinate 0, 0. Below is the reward value in the policy of the q learning agent.

x	y	value
0	0	4.1354805154725624e-05
0	1	4.0959544483241485e-05
0	2	4.037316443191692e-05
0	3	4.14941685873353e-05
0	4	4.1569327267904306e-05
0	5	4.481815011323156e-05
0	6	4.3456761408030645e-05
0	7	4.26509410695165e-05
1	0	4.2220866129246746e-05
1	1	0.0004654107615251269
1	2	3.8605439638716424e-05
1	3	8.637649263389756e-05
1	7	4.2339979801979645e-05

2 0 4.125620106569655e-05
2 1 4.10599516720775e-05
2 2 3.9926424088985646e-05
2 3 3.93714250290497e-05
2 4 1.4641567200749225e-05
2 5 1.4336577884296952e-05
2 6 1.3993416579082065e-05
2 7 0.01323301457958631
3 3 2.7033067052107685e-05
3 4 1.291560265372888e-05
3 5 2.9534161796910795e-05
3 7 0.00015717348296013874
4 0 0.0008925669022524375
4 1 0.001316495759132578
4 2 0.004375020240332637
4 3 1.3285153854613913e-05
4 4 1.2427691278332719e-05
4 5 -0.16275052715668678
4 7 0.009779107145152793
5 0 0.000990361704553123
5 1 0.0011695089420100772
5 5 0.000129741801592963
5 7 0.011889656110697951
6 0 0.0010328178017260299
6 1 0.0010684489195049859
6 2 0.003203804976903856
6 3 0.001320877276684433
6 4 0.0003073082235910531
6 5 0.00014969480921402732
6 6 0.07270129816424241
6 7 0.01217461138997413
7 0 0.0010608853792152654
7 1 0.0010694215948439077
7 2 0.0010145319283879507
7 3 0.0010021061763223402
7 4 0.001169043115331519
7 5 0.00017621213428245934
7 6 0.012278290568657138
7 7 0.012357447140404815