Shaastra 2022
RL Games

Team Beast:

Shashwat Garg &
Aviral Batra

(Registration ID: WGK6451)

Understanding of Problem Statement

This is a Classic Snake game Problem in which we need to create our own bot with the help of reinforcement learning although increased amount of food spawns and the presence of another agent increases the complexity of the state space by a lot. We not only have to improve(reduce) the amount of state spaces but also consider the actions of other agent.

The Goal of the Agent is to maximise the score during the whole process given the reward system and to interact with the other agent given both their scores accordingly.

Approach

We will Use Deep Q-Learning to Train our Agent to learn to Play Snake Game.

WHY ?

It is Because it seeks to find the best action to take given the current state that are outside the current policy

This means he will choose the action that maximizes the reward, based on his experience.

Approach (State Spaces)

- Obstacle Directly Up
- Obstacle Directly Down
- Obstacle Directly Left
- Obstacle Directly Right
- Nearest object food or opponent
- Opponent Above
- Opponent Below
- Opponent Left
- Opponent Right
- Nearest Food Above
- Nearest Food Below
- Nearest Food Left
- Nearest Food Right
- Is our Score more
- Is head North
- Is head South
- Is head East
- Is head West

All values are **0 or 1**

Approach (Why these State Spaces)

We have Chosen the above State Space as providing Coordinates would lead to **slow learning rate**.
A reason for the slow learning might be the number of possible states: $9*8^4*2*2*4$ = 5,89,824 different states are possible (there are 9 options for obstacles, the canvas is 8*8 steps, 2 for closest is food or opponent, 2 for if our score is more than opponents and 4 options for the current direction). For the our state space the number of possible states is equal to: $9*3^2*3^2*2*2*4$ = 11664 (9 options for obstacle each, 3 each for up/down and left/right, 2 for closest is food or opponent, 2 for if our score is more than opponents and 4 options for the current direction). 11664 is nearly 50 times smaller than previous one. This influences the learning process.