Reinforcement learning to solve Agar.io

Vishruit, Suhas

May 2, 2016

Abstract

Agar.io is a blob-eat-blob 2D world where the player controls a circular blob whose primary objective is to accumulate the largest amount of volume. This objective boils down to the amount of food it retains, supply of which consists of both static randomly dropped food and other real-time players. The available high-level decisions and actions for the blob include eating static food, actively avoiding bigger blobs, actively pursuing them. In addition, the blob has the additional possible action of splitting its mass to project a percentage of itself with a higher velocity in the intended direction.
The current report focuses on an exercise to solve a simpler version of the problem[1], a single player world with limited observability and continuous state space.[1]

1 Problem statement

The problem space consists of a large grid-world where the agent has a limited field of view, and their only primitive actions are forward movements or rotations. The objective is to keep moving to grids which have an green(positive reward) object, and avoid red(negative reward) objects. New objects are placed at random locations, keeping the total number of each kind same. The game isn't episodic, and the objective is to maximize reward/step averaged over a thousand steps.

The agent has sensors which output the (continuous dimension) distance to green, red and wall objects, along the single direction they sense.

2 Formulation

We used parametrized state spaces and function approximation to store and represent, as action values, the policy to follow if greedy.

Exploration was epsilon greedy.

[1]Project repository, including latest report and source code at https://github.com/snugghash/verbose-spork

Figure 1: Average reward over 100000 steps

3 Results

References

[1] http://cs.stanford.edu/people/karpathy/convnetjs/demo/rldemo.html