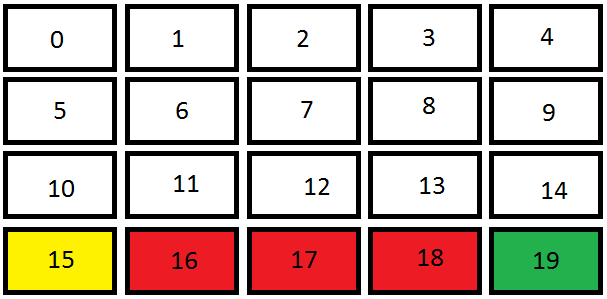
Report of implementing Q-Learning in simple game: cliff climber

Problem: control the player walk from 1 point passing the cliff and go to the destination point, avoiding to drop to the cliff.

Solution: e-greedy algorithm with Q-learning method is used as RL method, with epsilon value of 0.1, decay factor of 0.9 and learning rate of 0.01.

The reward is that, -100 if it falls to cliff, -1 for each movement, 100 if it gets destination.

Intuitive model is this:



Yellow: player start position

Red: cliff

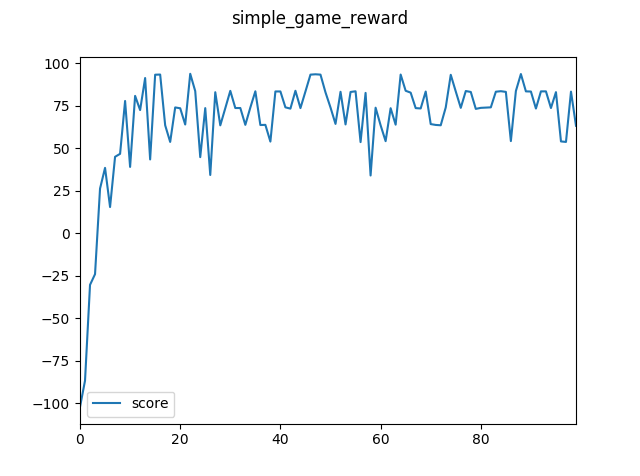
Green: destination

Possible actions are go up, down, left and right.

States are represented by the number of players’ position, 100 stands for terminal state.

The training was done in 2,000 games.

Result:



Scores are rewards averaged every 20 games.

