Reinforcement Learning - Assignment - Policy Gradient

In this assignment, you will implement the Monte-Carlo REINFORCE policy-gradient algorithm on the Basketball-v2 domain. You will also investigate the impact of different parameters on the performance of the reinforcement learning algorithm.

Due Date: 23:59 on Monday, 17th January 2022

Important Updates

Page opened Monday, 3rd January 2022

Environment

This environment is based on the FIRA Basketball competition. Implement the following environment as Open AI gym environments.

Basketball-v2

The basketball environment is a discretized environment with WIDTH and HEIGHT cells. At position (0,HEIGHT/2) is a ball. At position (WIDTH, HEIGHT/2) is a basket.

A robot can move, dribble, or shoot the ball.

The robot starts in an empty cell with uniform distribution.

Movement commands

The movement commands are UP, LEFT, DOWN, RIGHT. These commands will move the robot and if the cell is either empty or includes only a ball, then the robot will move into the target cell.

If the robot executes a movement or ball-handling command, then the robot will move 1 cell with a probability of 70% and 2 cells with a probability of 30% in the indicated direction. If the cell is not empty, then the movement command fails and the robot will remain in the start cell.

Dribble commands

The robot can dribble the ball into one of the four directions (DRIBBLE-UP, DRIBBLE-LEFT, DRIBBLE-DOWN, DRIBBLE-RIGHT). If the robot is in the same cell as the ball and the target cell is empty, then the robot and ball will move into the target cell. If the ball is not in the same cell as the robot or the target cell is not empty, then the command will fail.

Shoot commands

If the robot and the ball are in the same cell, then the robot can attempt to shoot at the basket. If the Euclidean distance between the robot and the basket is less than 1 cell, then the shot will succeed with 90% probability. In this case, the robot will receive a reward of +2. If the distance from the robot to the basket is between 1 and 3, then the probability of success is 66%. In this case, the robot will receive a reward +10. If the distance between the robot and the basket is between 3 and 4, then the probability of success is 10%. In this case, the robot will receive a reward of +30.

If the shot is unsuccessful, then the ball will be placed at location (0.8 * WIDTH, HEIGHT//2).

If the robot leaves the playing field, it will receive a penalty of -100.

The episode will end if the robot scores a point or if the robot leaves the playing field.

Evaluate the performance of the Monte-Carlo REINFORCE algorithm for this domain.

Opponents

A number of opponent players occupy squares in the playing field. The opponent players are uniformly distributed across the playing field.

If the robot executes a movement command or dribble command that would result in the robot moving into a cell with an opponent, then the robot will receive a penalty of 5 points and the ball will be moved back to the start position (0,HEIGHT/2).

Evaluation

Evaluate the performance of policy gradient algorithm using the following parameters

	WIDTH * HEIGHT		
	9 * 6 5 Opponents	12 * 10 10 Opponents	15 * 12 50 Opponents
Basketball-v2	REINFORCE	REINFORCE	REINFORCE

Submission

The submission consists of three parts: honesty, code, and report.

Create a directory with your id in the name for the code of your project <student-id>_rl_a2.

Add a README.txt file into the directory, which shows: (a) how to compile and run your program, and (b) any interesting features and extensions of your assignment.

Add the source code for your assignment into this directory.

Add a file <student-id>_rl_a2_report.pdf, which shows the results of the requested evaluations. Discuss the results. Were the results expected or did they surprise you? Did the results highlight shortcomings in the system? Do you have any ideas for how to fix those?

Compress the assignment and send it to the discord id **jacky.baltes#5889** via direct message. Do not post the assignment in the #simultaneous-localization-and-mapping channel.

If your assignment is too large (> 10 MB) to be sent via a direct message, then upload your assignment to Google drive or similar online storage and send the link in the direct message.