Reinforcement learning course
Home assignment #3: report

Mohammad Ali Nazari and Tomasz Kosiński

Task #2:

After running the Q-learning agent on a medium grid with 2000 training games and unchanged parameters (epsilon, gamma and alpha), it was impossible to reach the 95-100 % win level. The reason for it is that the exact Q-learning approach requires building a Q-function for the entire state space, which tends to impossible with the raising problem complexity.

In an attempt to fix it, we have conducted a small-scale empirical experiment, considering a very basic approach, namely increasing number of training games.

# of training games	Win rate [%]	Avg. score of last 100 games
2 000	6.5	-459.59
11 000	58	86.92
12 000	69	202.97
13 000	78	297.98
14 000	78	299.85
17 500	85	369.37
20 000	89	414.52
25 000	98	506

Another approach we could take is to tweak the Q-learning agent parameters (i.e. alpha, gamma and epsilon). Increasing, for instance, epsilon, the probability of taking random actions increase but with the expense of losing the game in many episodes (e.g. when Pacman moves around in the immediate vicinity of the Ghost).

Task #3:

The methods “getQValue” and “update” have been implemented. The program works for the smallGrid and Identity Feature Extractor. However, if fails when we use the Simple Extractor. Despite our serious effort to find the mistake in our implementation/reasoning, we submit this initial version of our work, and we will continue with the correction phase.

Questions:

- In our Pacman game, two approaches to features are proposed. The first one is called Identity extractor, and the feature function matches the state-action pairs. The second one offers the state, x, y, and action features/attributes. It seems that the features do not have meaningful interpretation (to our current understanding).
- The update equation 2, takes into account the maximum value we can have in the next state (delta), and updates the weights w_i so that the feature function with more contribution in the value of the next state, gets more weight.