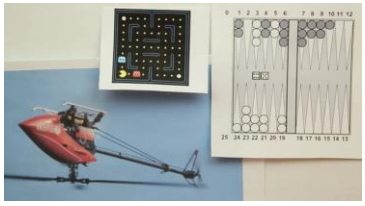
Computational Intelligence
Unit # 15

Reinforcement Learning



Acknowledgement

- Several examples of this lecture have been taken from Stanford AI class and Stanford Machine Learning class.

Q-learning

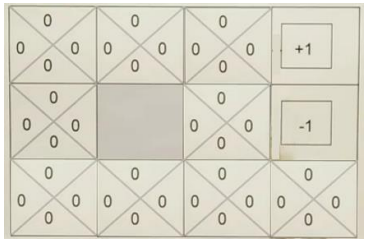
- Q-learning is a reinforcement learning algorithm that does not need a model of its environment and can be used online. Q-learning algorithms work by estimating the values of state-action pairs.
- The value Q(s, a) is defined to be the expected discounted sum of future payoffs obtained by taking action a from state s and following the current optimal policy thereafter. Once these values have been learned, the optimal action from any state is the one with the highest Q-value.
- The agent can perform adaptively in a world without understanding it. All it tries to do is sort out good actions to perform from bad ones.

QLearning

Initialize $Q(s, a)$ arbitrarily
Repeat (for each episode):
  Initialize s
  Repeat (for each step of episode):
    Choose a from s using policy derived from Q (e.g., ε-greedy)
    Take action a, observe r, s'
    $Q(s, a) \leftarrow Q(s, a) + \alpha\left[r + \gamma \max_{a'} Q(s', a') - Q(s, a)\right]$
    $s \leftarrow s'$;
  until s is terminal

Figure 6.12: Q-learning: An off-policy TD control algorithm.

Q-learning



Instead of having 'value' for each state. We now have a value for each state, action (s,a) pair.

QLearning

- Q-learning is *model-free*(calls this *primitive learning*). We ignore the question of what *x*,*a* actually lead to, and the problem of storage of that large and complex mapping, and just concentrate on scoring how *good x*,*a* is.
- The agent can perform adaptively in a world without understanding it. All it tries to do is sort out good actions to perform from bad ones.

Some Videos

- http://www.youtube.com/watch?v=tovrpoUkzYU
- http://www.youtube.com/watch?v=Xf_IhCbTQGY&feature=BFa&list=FLTnVTxTVQSBDkVSWPhBpRTQ

How to deal with continuous domains?

- Discretization
- Function Approximation

References

- R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.
- http://reinforcementlearning.ai-depot.com/Main.html
- http://en.wikipedia.org/wiki/Markov_decision_process