CS 5100 – PS7

Yijing Xiao

**Problem: Q-Learning**

a) **Give an equation for $V^\pi$ in terms of $Q^\pi$ and $\pi$ (a | s).**

$$V^\pi(s) = \sum_{a \in A} \pi(a|s) \; Q^\pi(s,a)$$

b) **Give an equation for $Q^\pi$ in terms of $V^\pi$, T (s, a, s'), and R (s, a, s').**

$$Q^\pi(s,a) = R(s,a,s') + \gamma \sum_{s' \in S} T(s,a,s') \, V^\pi(s')$$

c) **Derive the general Bellman equation for $Q^\pi$, in terms of $\pi$(a' | s'), T (s, a, s'), R (s, a, s'), and $Q^\pi$ (s', a').**

According to equations we have on question a and b, we could derive that:

$$Q^\pi(s,a) = R(s,a,s') + \gamma \sum_{s' \in S} T(s,a,s') \sum_{a' \in A} \pi(a' | s') \, Q^\pi(s',a')$$

d) **Give an equation for $V^*$ in terms of $Q^*$.**

$$V^*(s) = \max_a Q^*(s,a)$$

e) **Give an equation for $Q^*$ in terms of $V^*$, T (s, a, s'), and R (s, a, s').**

$$Q^*(s,a) = R(s,a,s') + \gamma \sum_{s' \in S} T(s,a,s') \, V^*(s')$$

f) **Derive the general Bellman optimality equation for $Q^*$, in terms of T (s, a, s'), R (s, a, s'), and $Q^*$ (s', a').**

According to equations we have on question d and e, we could derive that:

$$Q^*(s,a) = R(s,a,s') + \gamma \sum_{s' \in S} T(s,a,s') \max_{a'} Q^*(s',a')$$

g) **Give an equation for $\pi^*$ in terms of $Q^*$.**

$$\pi^*(s) = \operatorname*{argmax}_a Q^*(s,a)$$

h) **Give an equation for $\pi^*$ in terms of $V^*$, T (s, a, s'), and R (s, a, s').**

According to equations we have on question e and g, we could derive that:

$$\pi^*(s) = \operatorname*{argmax}_a (R(s,a,s') + \gamma \sum_{s' \in S} T(s,a,s') \, V^*(s'))$$