15-883 Computational Models of Neural Systems

Homework 7

This assignment gives you some practice with the Temporal Difference (TD) learning rule. Review the equations in Lecture 5.1.

Assume we have a single stimulus x_1 that comes on at the start of a trial (time t=0) for one time unit.

Assume a reward of 100 is supplied at time t=3.

Recall that V(t) is the total discounted future reward:

$$V(t) = \sum_{\tau=0}^{\infty} \gamma^{\tau} r(t+\tau+1) = r(t+1)+\gamma V(t+1)$$

1. Assume a discount factor of 0.9. Assume that r(t) and V(t) are zero if t<0 or t>4. Fill in the values for V(t) in the table below:

t	**x_1(t)**	**r(t)**	**V(t)**	**δ(t)**
0	**1**	0		
1	0	0		
2	0	0		
3	0	**100**		
4	0	0		

2. Recall that the reward prediction error δ(t) is the difference between the predicted future reward at time t and the actual reward at time t+1 plus the discounted expected future reward γV(t+1):

$$\delta(t) = r(t+1)+\gamma V(t+1)-V(t)$$

Use this formula to fill in δ(t) in the table above.

[Continued on next page.]

3. Let's use a single linear neuron to calculate V*(t), our estimate of V(t). Assume a complete serial compound representation of x_1, so we have weights $w_{1,0}$ through $w_{1,4}$ and V*(t) is given by:

$$V^*(t) = \sum_{i=0}^{4} w_{1,i} \cdot x_1(i)$$

Assume that all the weights start out at 0. Assume a learning rate of η=0.1. The TD learning rule is:

$$\Delta w_{1,i}(t) = \eta \cdot \delta(t) \cdot x_1(i)$$

A trial takes 5 time steps. Using the learning rule above, you can calculate how the weights will look at the end of each trial. Fill in the tables below. Use the current row of the *w* table to calculate the next row of the V* table. Then use the latest row of the V* table to calculate the next row of the δ table. Finally, use the latest row of the δ table to calculate the next row of the *w* table.

After *n* Trials	**$w_{1,0}$**	**$w_{1,1}$**	**$w_{1,2}$**	**$w_{1,3}$**	**$w_{1,4}$**
0	0	0	0	0	0
1					
2					
3					

Trial	**V*(0)**	**V*(1)**	**V*(2)**	**V*(3)**	**V*(4)**
1					
2					
3					

Trial	**δ(0)**	**δ(1)**	**δ(2)**	**δ(3)**	**δ(4)**
1					
2					
3					