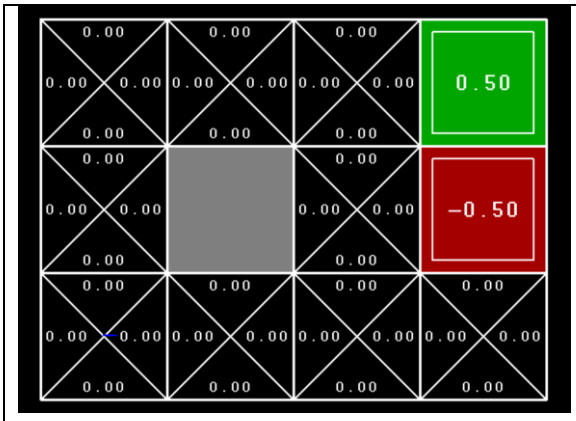
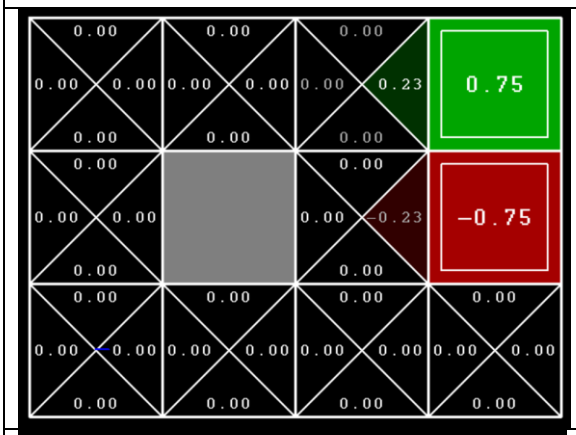
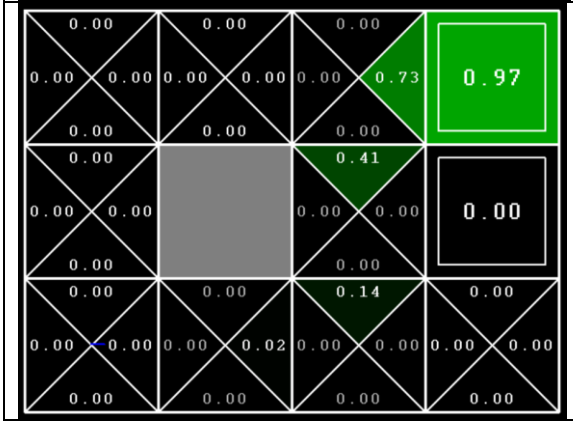
Your Name	
Your UnetID	

Active Reinforcement Learning

- $Q_{k+1}(s,a) \leftarrow (1-\alpha)Q_k(s,a) + \alpha \cdot [R(s,a,s') + \gamma \max_{a'} Q_k(s,a)\,]$

	After two episodes, one exiting from the negative exit, and the other exiting from the positive, we have these Q-values. What Q-values do we have if we exit from the (1,3) square through the positive exit? Hint: two values change. α=0.5, γ=0.9, $R(s,a,s') = 0$
	After two more episode2, one exiting from (2,3) through the negative exit, and one exiting from (1,3) through the positive exit, what are our Q values? What values change if we exit one more time through (1,3) through the positive exit? What are the new Q values?
	These are the values after 5 episode moving long the bottom edge and then up column 3 to exit through the positive exit. Why are there no negative Q values?

Exploration:

What does $f(u,n) = u + \frac{c}{n}$ with constant c do as n->infinity?	

Feature Driven Models:

What does $Q(s,a) = 4.0 \cdot f_{\mathrm{DOT}}(s,a) - 1.0 \cdot f_{\mathrm{GST}}(s,a)$ Become after the update? *F_{dot}(s,a) = distance to closest big dot with action a.* *F_{gst}(s,a) = distance to closest ghost with action a*	