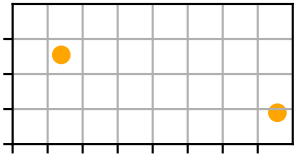
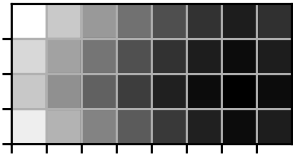
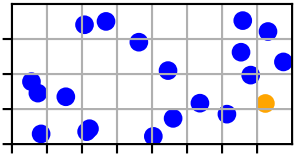
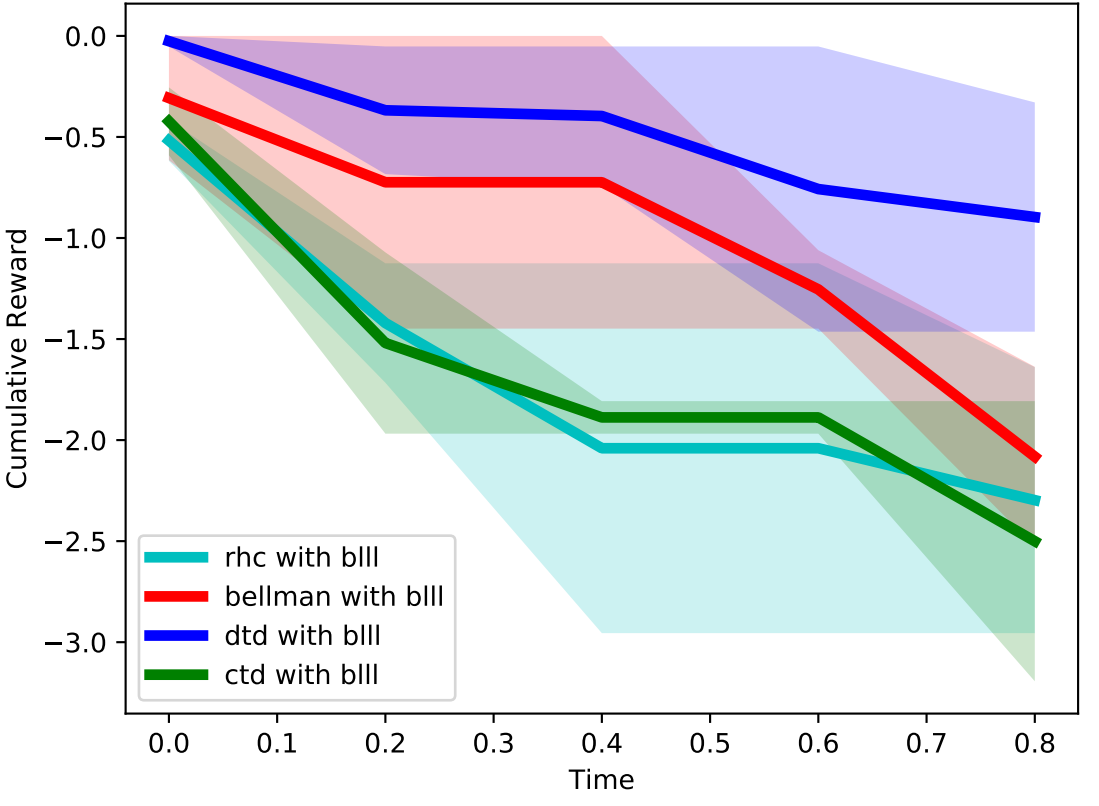
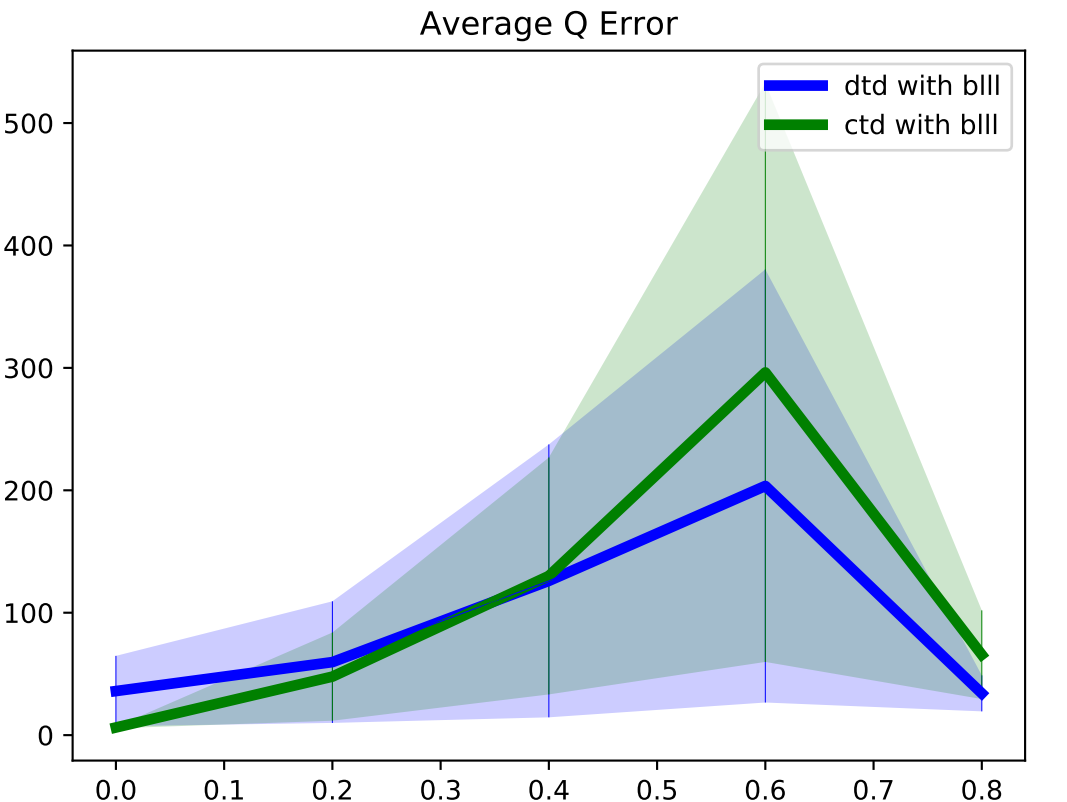
rhc with blll at t/T = 0.0/1.0








0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
Cumulative Reward
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Time
rhc with blll
bellman with blll
dtd with blll
ctd with blll


Average Q Error
dtd with blll
ctd with blll
500
400
300
200
100
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8