Exercise 14.1 (a) Uncontrolled Episodes
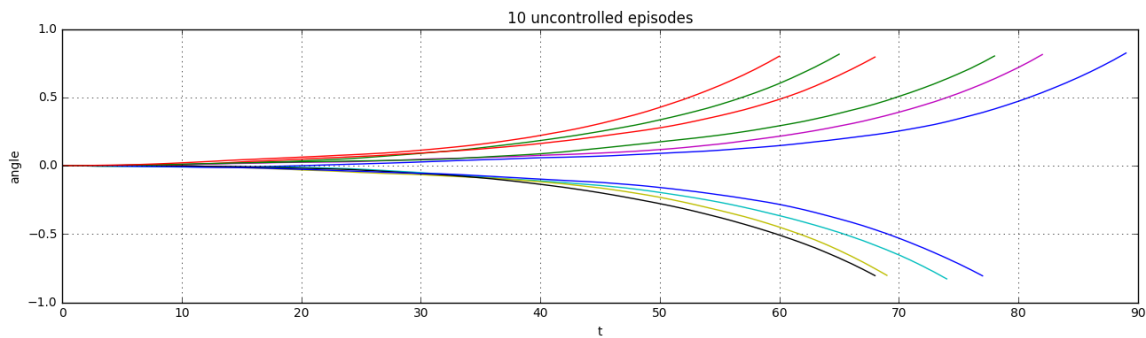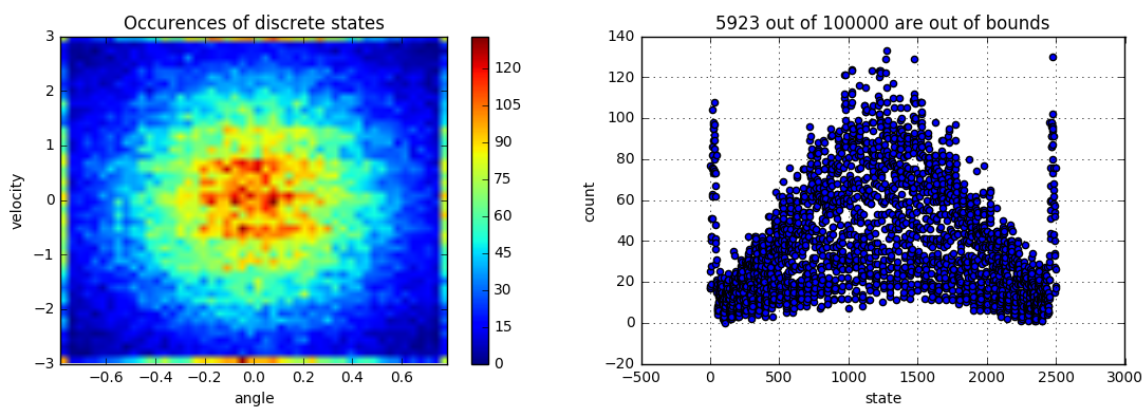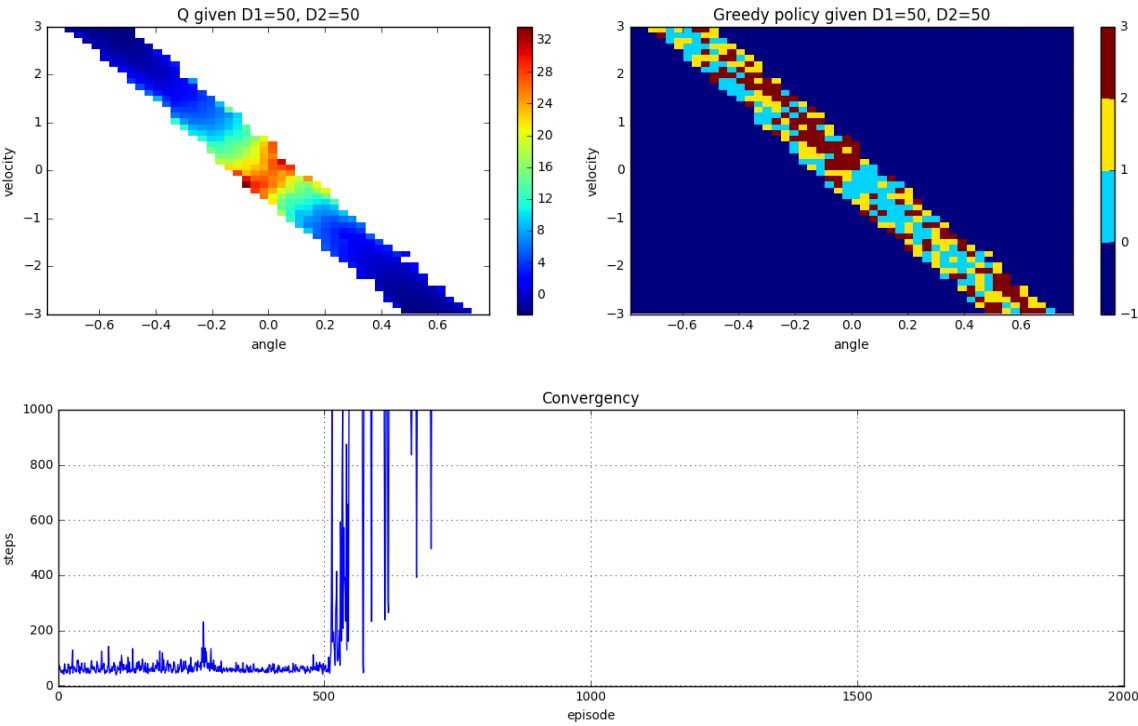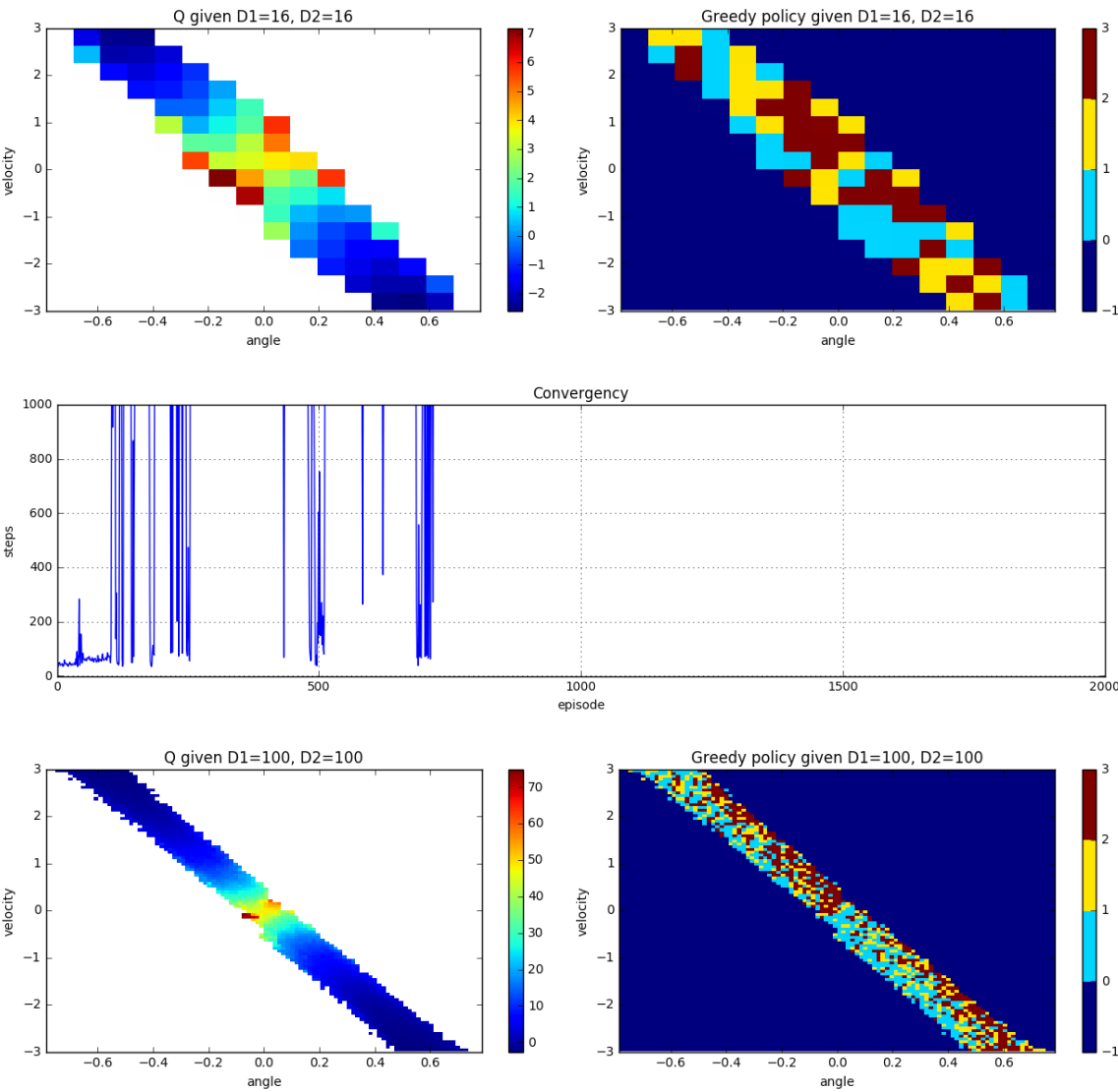

Exercise 14.1 (b) Discrete Spaces
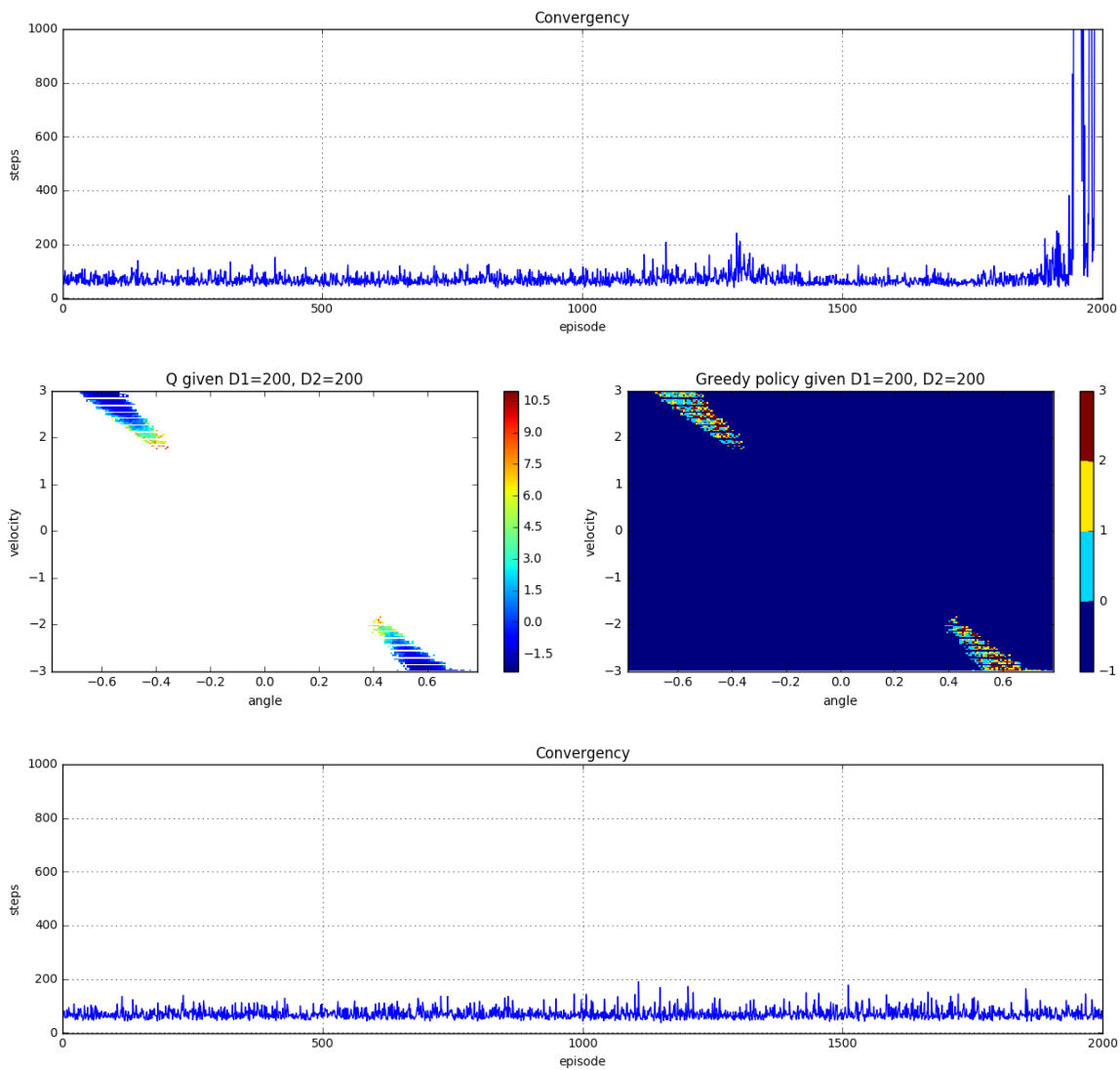

Exercise 14.2 (a) Q-Learning with e-greedy policy

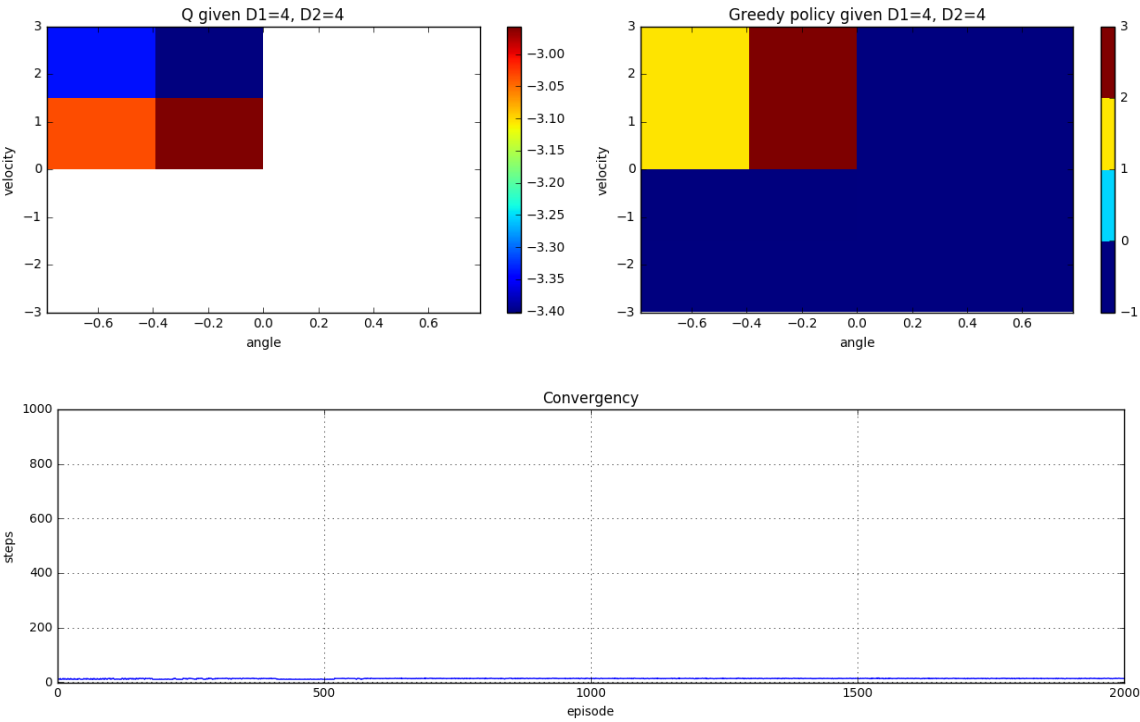## Exercise 14.2 (b) Changing D1 and D2

Q given D1=16, D2=16

Greedy policy given D1=16, D2=16

Convergency

Q given D1=100, D2=100

Greedy policy given D1=100, D2=100

Convergency

Q given D1=200, D2=200
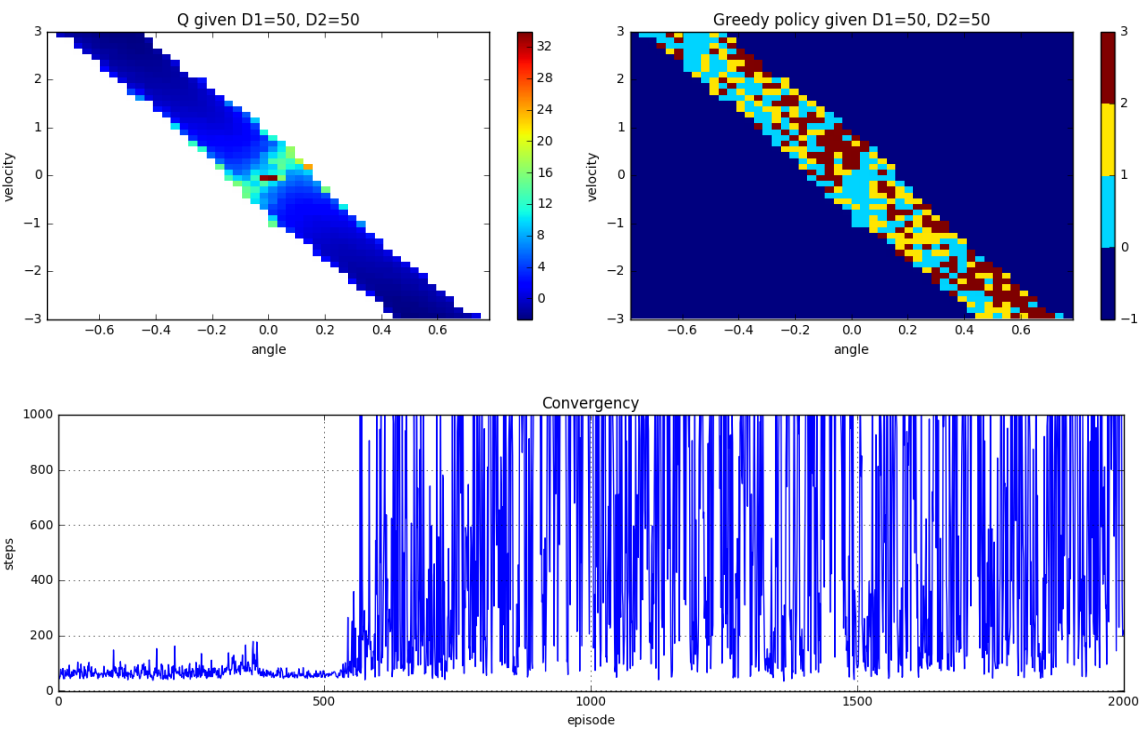

Greedy policy given D1=200, D2=200

Convergency

14.2 (c)

Smallest number that earns a good policy: Should be D1=D2=2, for me it was around 16

14.2 (d) e-greedy policy (e = 0.1)

14.2 (e) fast learning rate (eta = 1)