Task 1

I would like to note, the assignment says to treat uninitialized Q values as 0 but this gives wrong results for utilities. The utility function is ultimately a max function, and a max function between ntr=-0.04 and 0 will return 0 when ntr would be more accurate. I made this change in the code and got much more accurate results. I set the default Q value for utility only to -1 because it doesn’t know if it is able to reach a terminal state. Any significantly negative value works the exact same. Q values everywhere else have a default value of 0 as defined in the assignment description.

**AgentModel\_Q\_Learning("environment2.txt", -0.01, 0.99, 10000, 1000)**

utilities:

0.923 0.955 0.977 1.000

0.901 0.000 0.881 -1.000

0.870 0.847 0.858 0.830

policy:

> > > o

^ x < o

^ < ^ <

Active Q-Learning took 0.25 seconds to run.

**AgentModel\_Q\_Learning("environment2.txt", -0.04, 0.9, 3000000, 50000)**

utilities:

0.510 0.649 0.796 1.000

0.399 0.000 0.489 -1.000

0.297 0.253 0.344 0.129

policy:

> > > o

^ x ^ o

^ > ^ <

Active Q-Learning took 77.83 seconds to run.