



观察到收敛速度随着epsilon值的增大而增快，而收敛到的值随着epsilon的增大而减小。

这大概是因为epsilon值增大时更容易找到最优策略，但是下一步使用最优策略的概率为 $1-\epsilon$ ，这个概率会减小，因此实际上的收敛值也会降低。