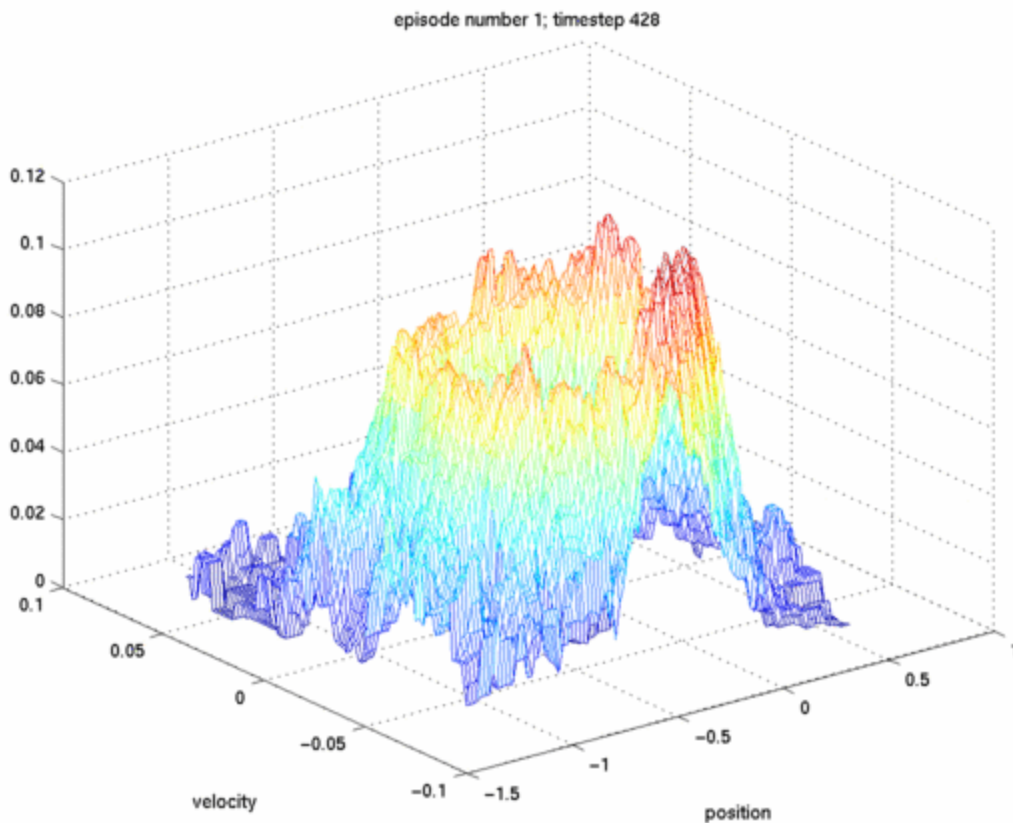
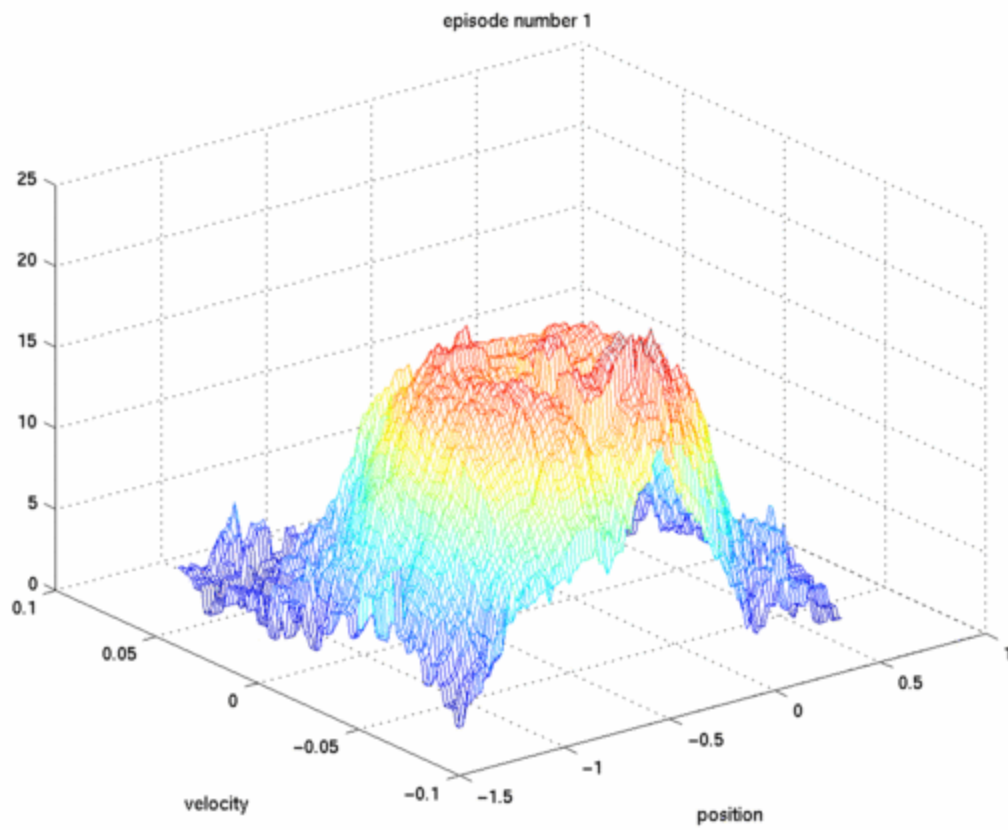


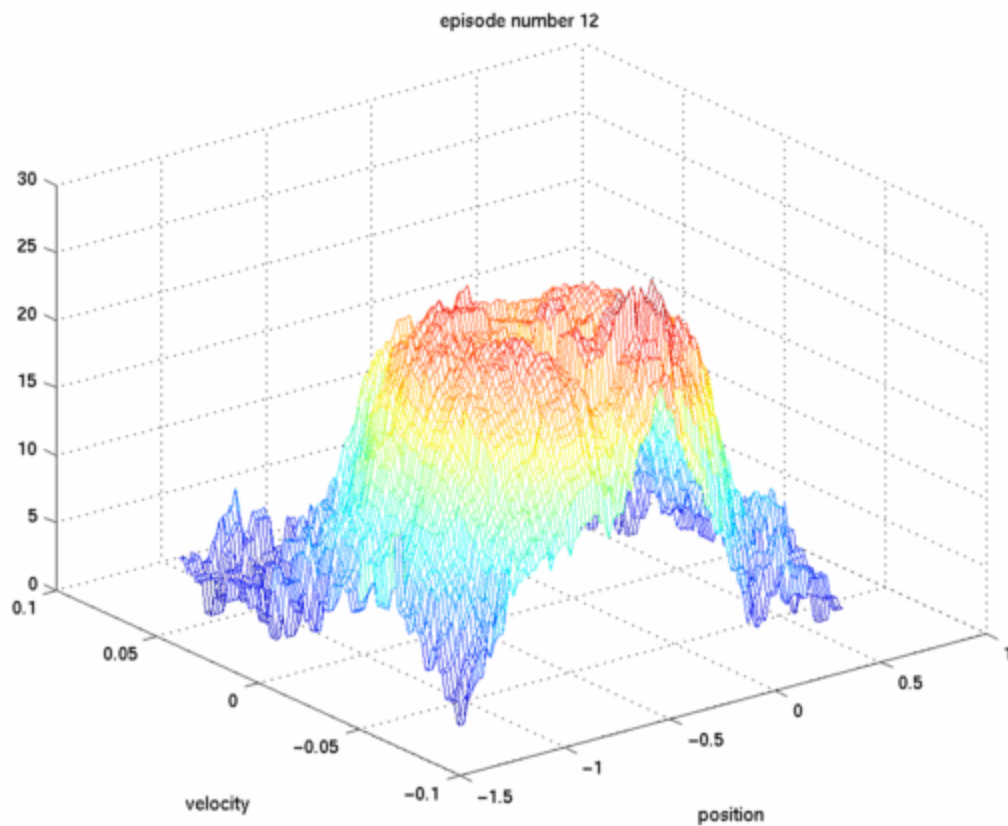
Here you will find replicated experiments from the mountain car example from this chapter in the book. To obtain plots at episode 9000 I choose to run the code `mnt_car_learn.m` with a rather small learning rate $\alpha=0.005$. This will cause learning to proceed very slowly but will also make the convergence to the cost to go surface more stable. I first present mesh plots like those shown in the book. For the first episode during the 328th timestep.



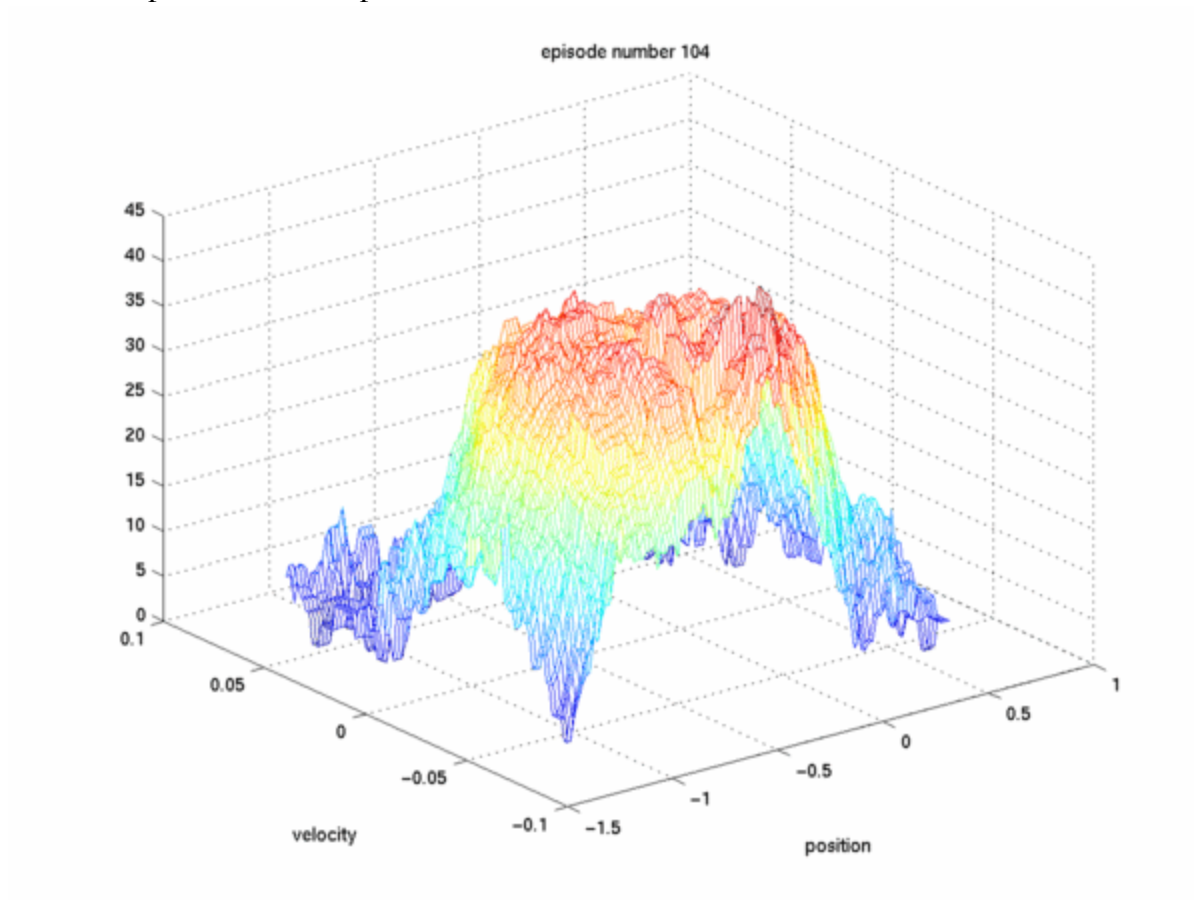
For the last timestep in the first episode we obtain



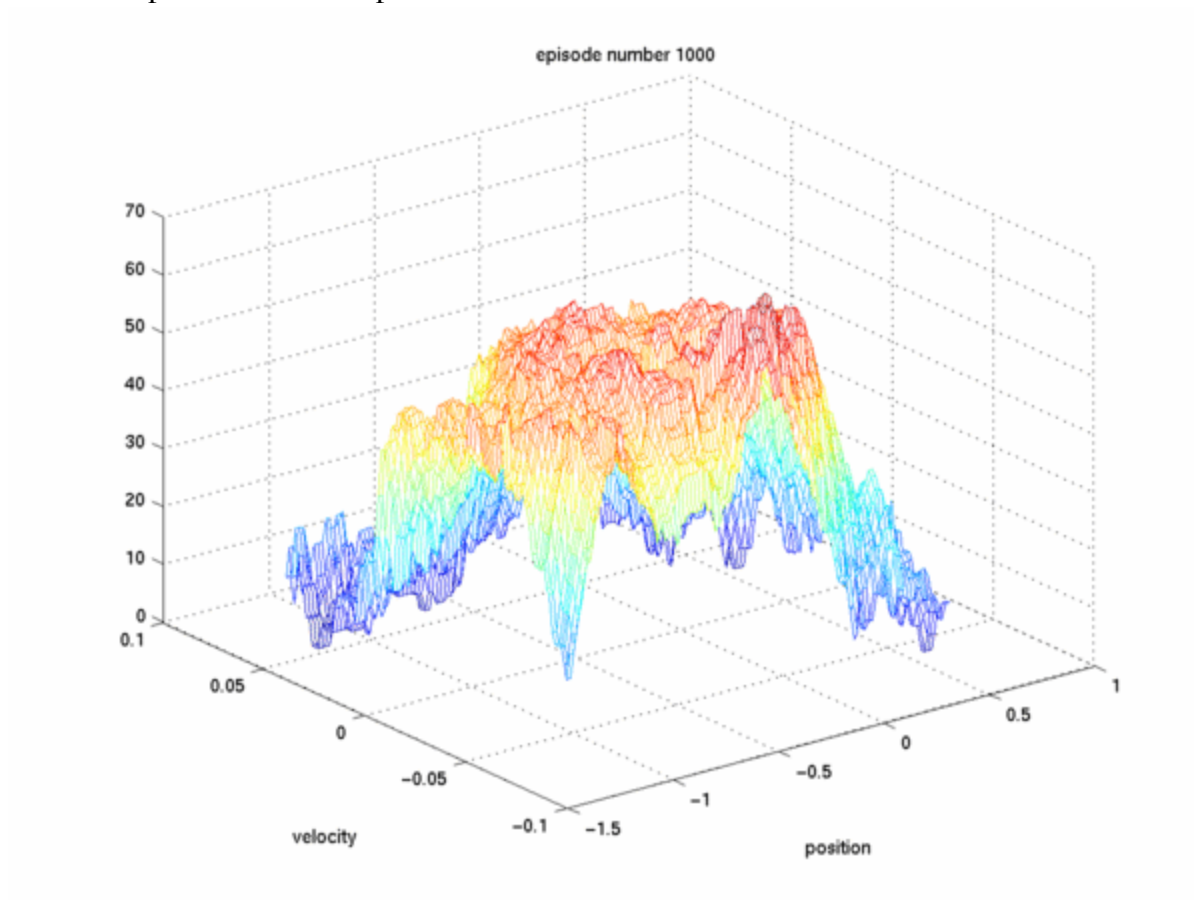
For the last timestep in the 12ths episode we obtain



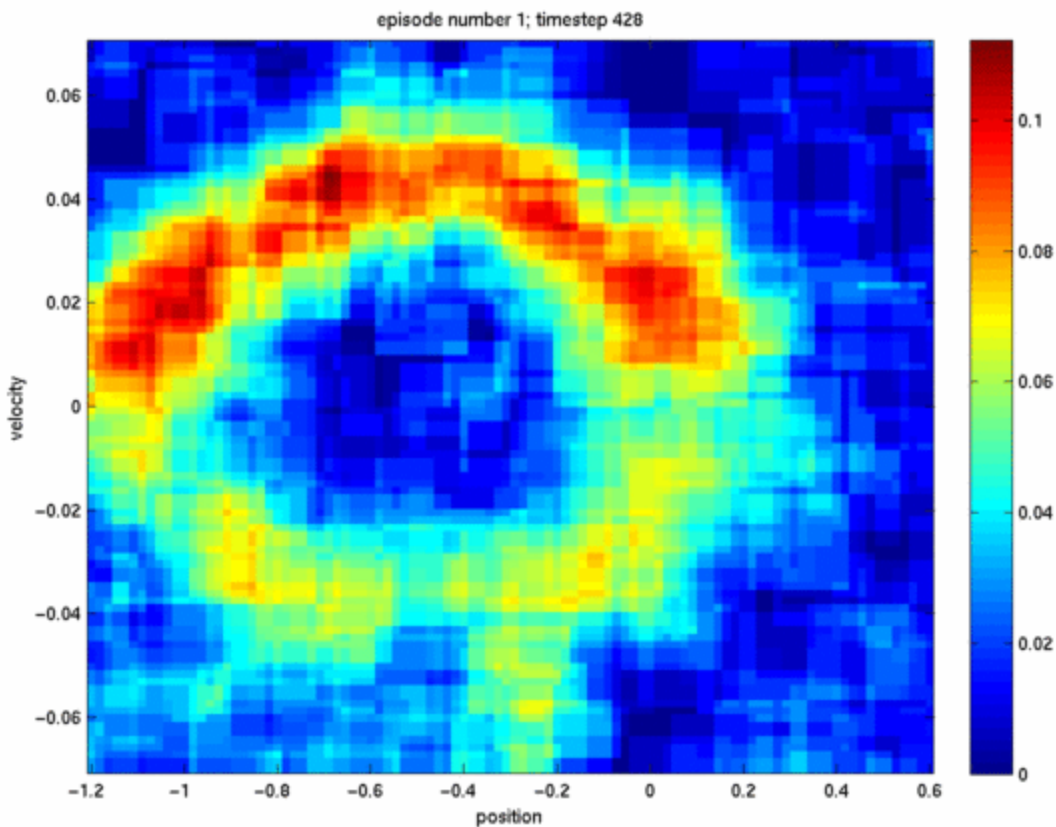
For the last timestep in the 104th episode we obtain



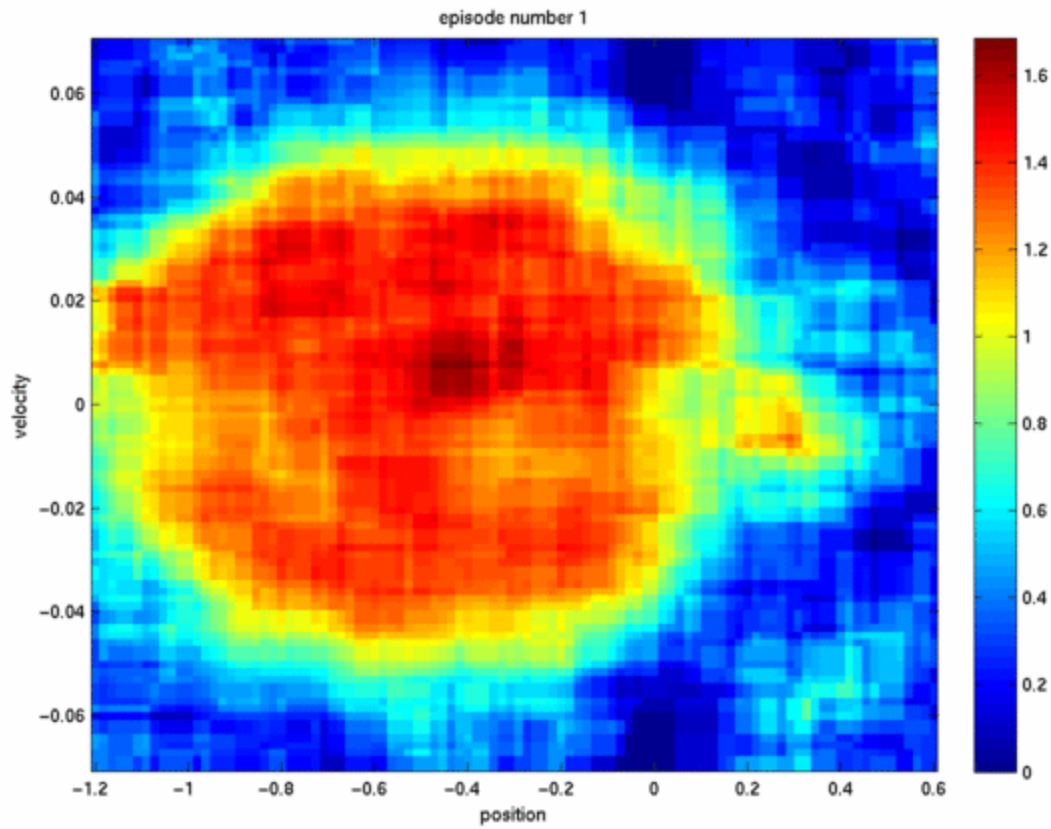
For the last timestep in the 1000th episode we obtain



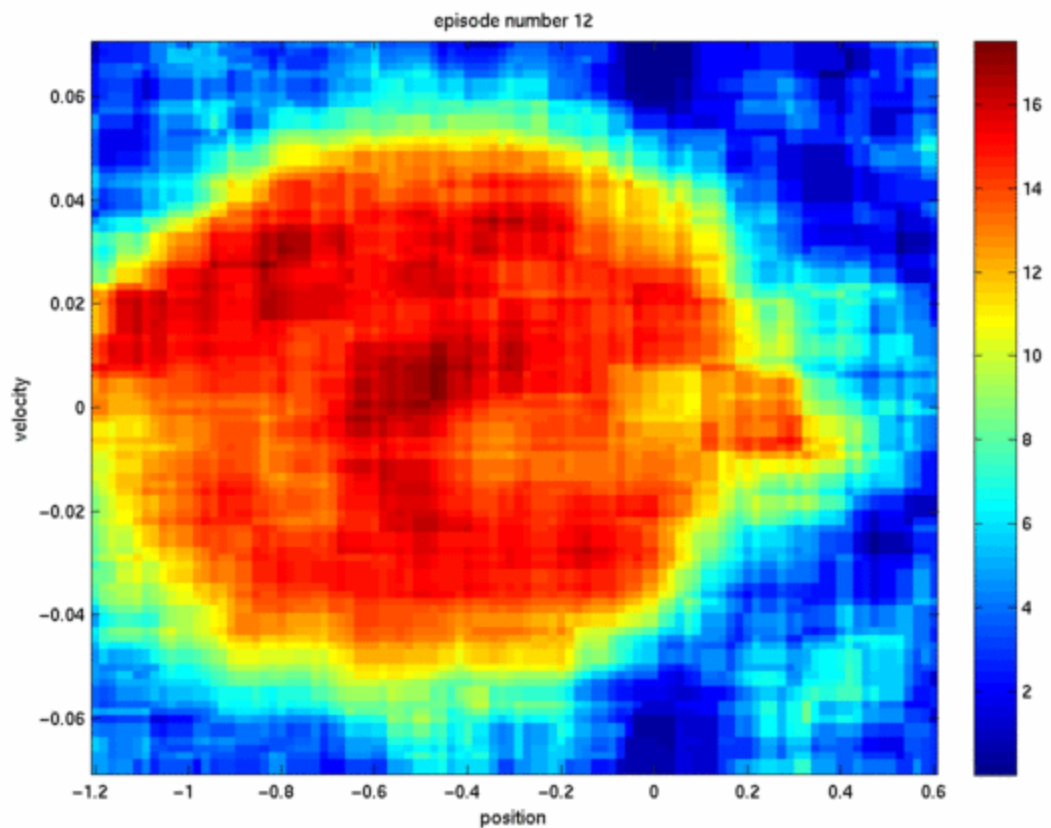
From these plots we see that very quickly the policy obtained from only a few episodes is quite similar in appearance to that obtained in the limiting case (say an infinite number of episodes). These results look quite similar to those presented in the book. To view these plots differently we can present them using Matlabs "imagesc" command. Doing this we obtain, for the first episode during the 328th timestep.



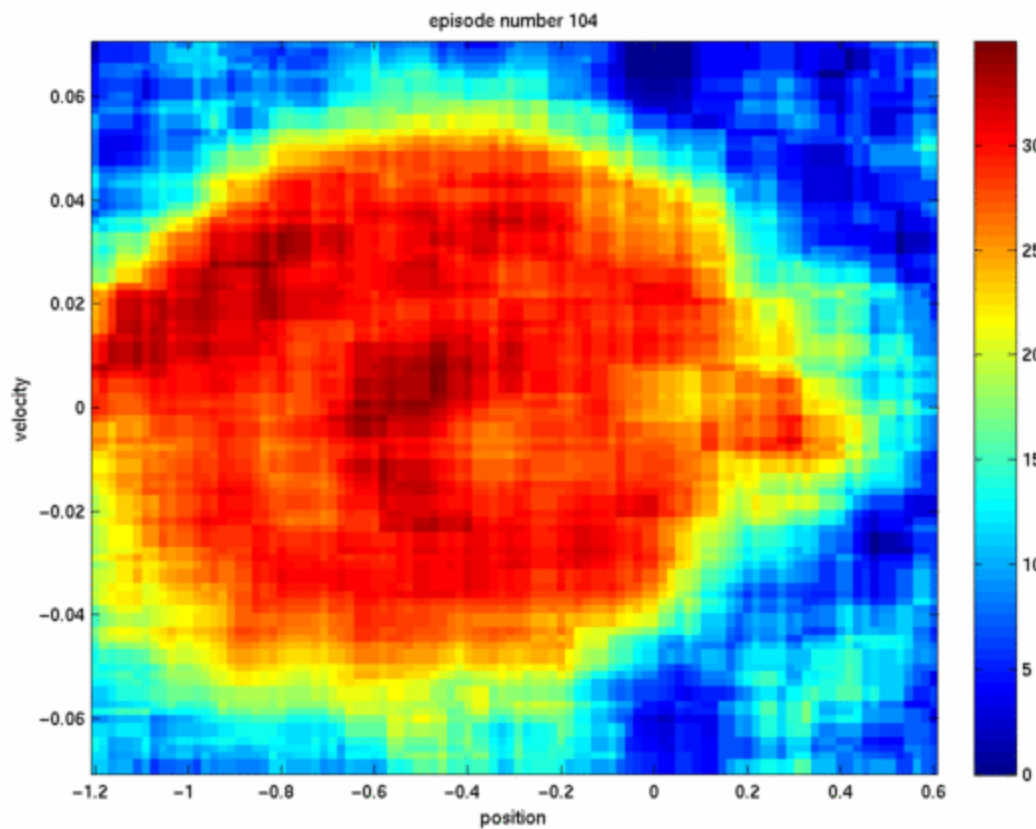
For the last timestep in the first episode we obtain



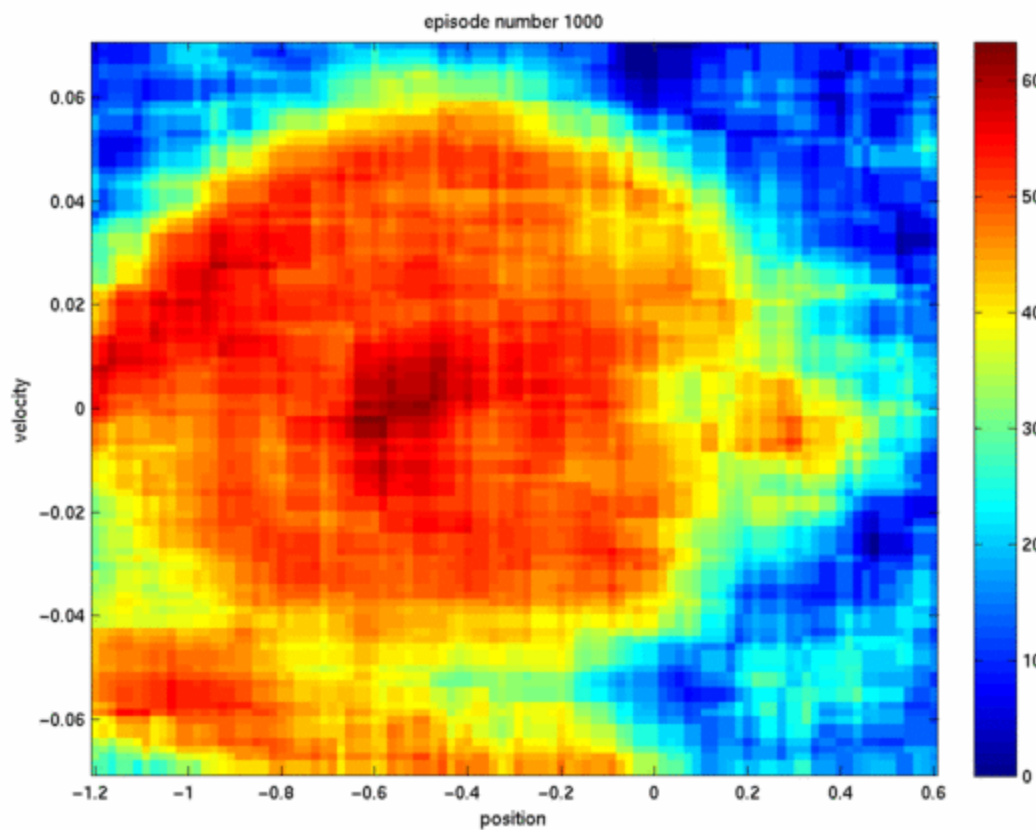
For the last timestep in the 12ths episode we obtain



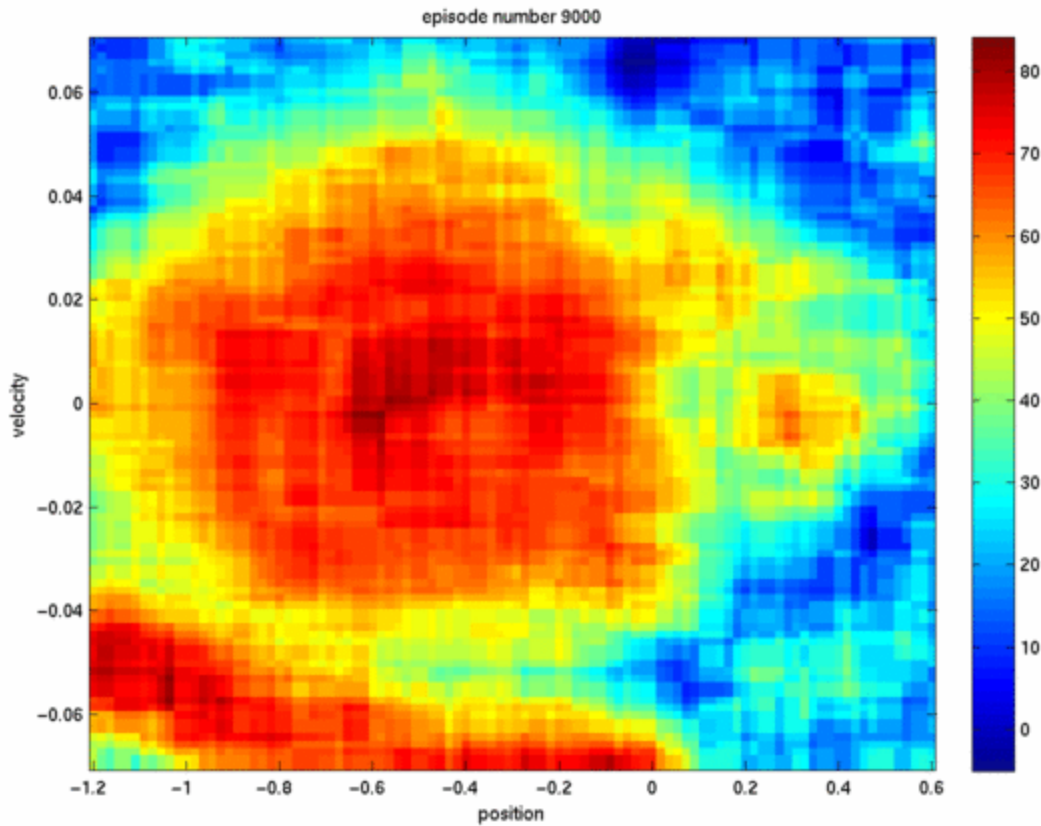
For the last timestep in the 104ths episode we obtain



For the last timestep in the 1000ths episode we obtain



For the last timestep in the 9000ths episode we obtain



These results look very similar to those presented in the book.

[John Weatherwax](#)

Last modified: Sun May 15 08:46:34 EDT 2005