

The final learned function after 10000 iterations has 3 main peaks, and 3 main troughs. Our learned function after only 20 iterations, is mostly flat. This is because most of the state spaces have not been visited at all yet. The few bumps and dips represent states that have been visited within the 20 iterations. On the left we see a tall peak. This peak isn't as tall as it will become, since the x,y corresponding to it's highest point has not yet been visited. This is the same for all peaks and troughs in the 20 iteration sample. Most aren't as tall or as deep as they should be, but they're all in roughly the right place. The estimate of a particular state space depends quite heavily on the estimates of its immediate neighbours. As seen in the graph, (x,y) pairs who's immediate neighbours have also been visited (are non-zero) correspond very closely to the same value on the 10,000 iteration graph.

If the tiling was input space was changed to 11×21 , with the x dimension being divided twice as finely as the y dimension, the width of the peaks and troughs in the y dimension would be learned much faster. This occurs because there are so many more state spaces in the x dimension to visit, so an even higher percentage of them will be unvisited after 20 iterations. There is an upside to this alternate tiling however. Though the x 's will be less accurate in the 20 iteration graph, they will have twice the granularity on the 10,000 iteration graph.