1. The before value of the 4th point is non-zero because it is VERY similar to the 2nd point. So similar, in fact, that it shares 4/8 tiles with the 2nd point. Because of this, half of the 8 theta values corresponding with the tiles have non-zero values from learning the 2nd point, and this will produce a non-zero estimate.

2. The MSE comes down smoothly from 0.25 to 0.01, and then stays there. It does not further decrease because the state spaces are being generalized. This means that the machine can't learn the best thing to do in EVERY state, but only the best thing to do in most states. Because of this, it will always be wrong some of the time.