- 1. Every tile has a specific tile number, and the second tiling creates 121 new tiles (since we have 11x11 grid). Since the first tiling fills tile indices 0-120, then the second tiling ranges from 121-241.
- 2. Each tiling adds a new set of 121 tiles, starting at the next index that the previous tiling ends at. Since the first tiling ends at 120, the next first tile starts at 121. If you number the 8 tilings from 0-7 and multiply 121 to their index, it will give the first tile for the first 7 tilings. This is essentially saying that the point is found in the first tile for each of the first 7 tiling schemes.

- 3. For the first seven tilings, the top right corner of the first tile is higher than 0.1,0,1 which means it contains the point. For the 8th tiling, the top right corner of the first tile is (0.6/8) which is 0.075, which is smaller than point 0.1, 0.1 and therefore does not contain the point. The point is then contained in the tile diagonally attached (12 indices higher)
- 4. The 13th tile in the 8th tiling is tile 859 because the point is in the second column of the second row. The second column in the second row is 12 tile indices higher than the first tile (1st column in first row), if we are counting to the right by 1's. Since the first tile index of the 8th tiling is 847 (which is 7*121), then the tile diagonally above that is tile 859.
- 5. The maximum tile index is 967 because there are 121 tiles per tiling, with a total of 8 tilings. One would assume the maximum to be 968 because (121*8) = 968, however we are starting from index 0, so the last index is 967.
- 6. The second and fourth examples have very similar outputs because their inputs are extremely similar. The point in y axis differs by 0.1, which means they are most often found in the same tile for each tiling. However, since there is 0.1 difference in the y axis, this affects which tile the point is found in for the 6th and 7th tiling. For these tilings, the point is found 1 row higher (11 tiles greater).

Explanation of part 2

The plot after running the learn function for 20 episodes shows roughly 6 peaks, and 6 valleys. There are two strong peaks, the tallest having a height of 0.14 and the second tallest having a height of 0.11. Other peaks have similar heights around 0.03. The two tall peaks have somewhat similar heights because the step size is only 0.1/8 and only being ran for 20 episodes. This means that the updates for each episode will only update the height a maximum of ~1% of it's true value. This also explains why the widths of the peaks and valleys are similar. The area on the plot where there are no peaks or values are simply due to the fact that the 20 episodes did not cover any points in those tiles. If the dimensions of the tiles were changed to 11x21, the plot would have more peaks/valleys. This means that examples 2 (in1=4, in2=2, target=-1.0) and 4 (in1=4, in2=2.1, target=-1.0) will be effecting less of the same tiles, since there are more tiles covering the same amount of area.

Answers to questions about part 2

- 1. The before value of the fourth point is non zero because the second example has many similar tile indices, and after learning from the second example, the indices it's covered has already been updated.
- 2. The MSE does not decrease further toward zero because of the randomly generated number with a mean of 0 and standard deviation of 0.1. This is effecting the target in that it does not dip below 0.01 because 0.01 is the MSE and the MSE is the squared standard deviation.

 $MSE = SD^2$

0.01 = 0.1^2

The estimated MSE: 0.250624172553
The estimated MSE: 0.0570193140199
The estimated MSE: 0.02085333843
The estimated MSE: 0.0145199834427
The estimated MSE: 0.0124026289134
The estimated MSE: 0.0117035810798
The estimated MSE: 0.0120987100577
The estimated MSE: 0.0116460483132
The estimated MSE: 0.0113511501288
The estimated MSE: 0.0112610216573
The estimated MSE: 0.0116863283344



