**Explanation**: The f20 plot is a very rough estimate of the f10000 plot because after only 20 examples, there only exists many extreme values where as the f10000 plot shows a much more fluid picture of the algorithm due to generalization of many values. There are many peaks and valleys in both the f20 and f10000 plots when appear to range from just under 1.0 and just above -1.0 in height. There appears to be 3 peaks and 3 valleys corresponding to dark red to dark blue on the plots respectfully. This is due to the target sin(x-3)cos(y) function we are trying to approximate. Confirmed by plotting on Wolfram Alpha. The width of them is related to the number of examples we run as the more examples we have, the more prominent the peak or valley will become due to generalization. F20 has very little generalization so the width of the peaks is very small. This is due to the limited values that the function approximator has.

If we were to use a 11x21 grid instead of a 11x11grid as in this experiment, the results would appear much different such that there would be more rounded and smooth peaks and valleys. This is because the tiling has a finer granulate and its able to generalize must smoother than with larger tiles. Learning after 20 examples would change to even less error as we are dealing with a more fine range of values and the slope would be less steep compared to the 11x11 grid. Overall an alternative tiling with more tiles will provide a smoother generalization.