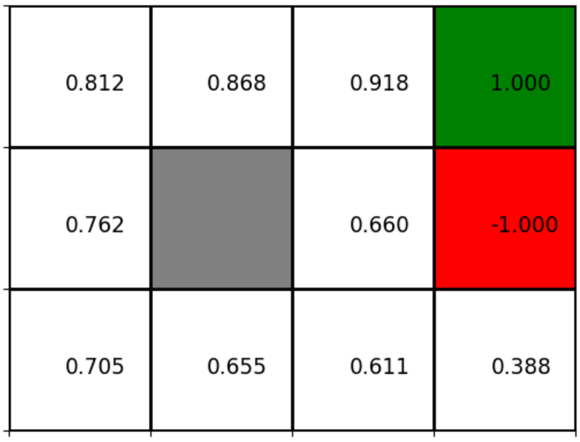
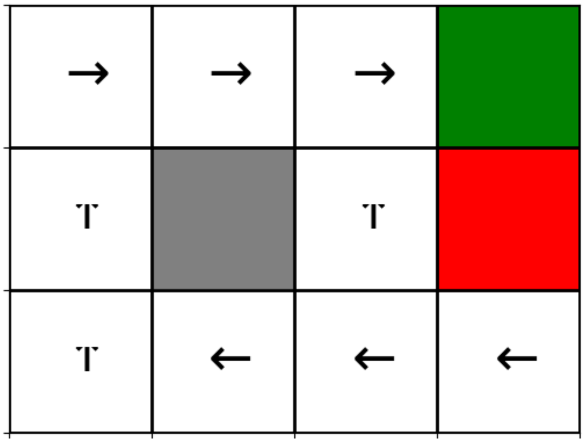
TTK23 - Assignment 1 – Simen Keiland Fondevik

Problem 1: Value Iteration

a), b)

Gamma = 1.0 gives the following value function and policy.

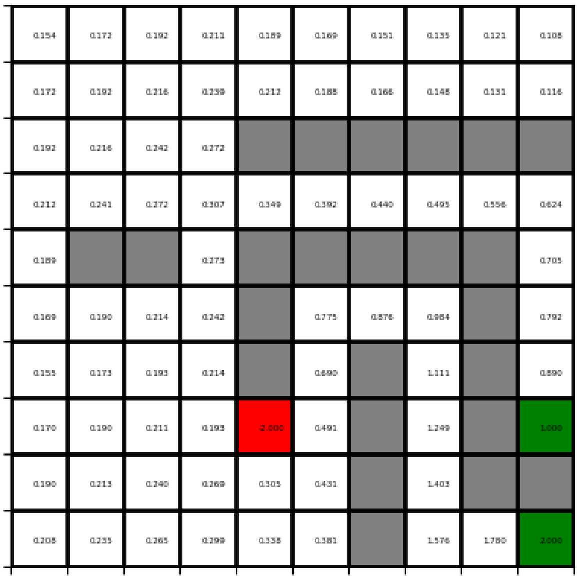


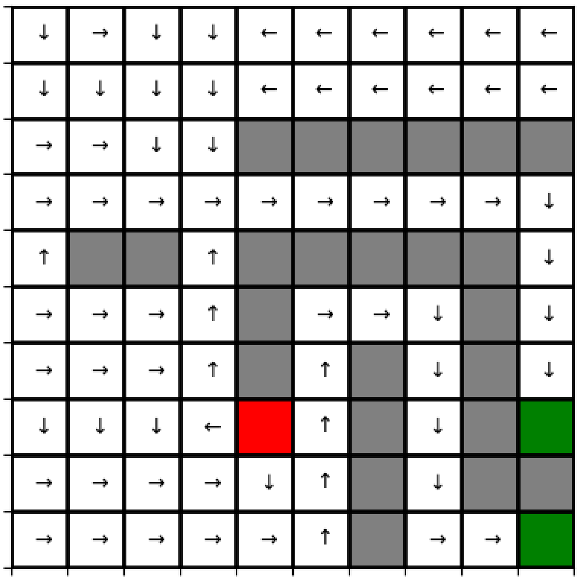


c)

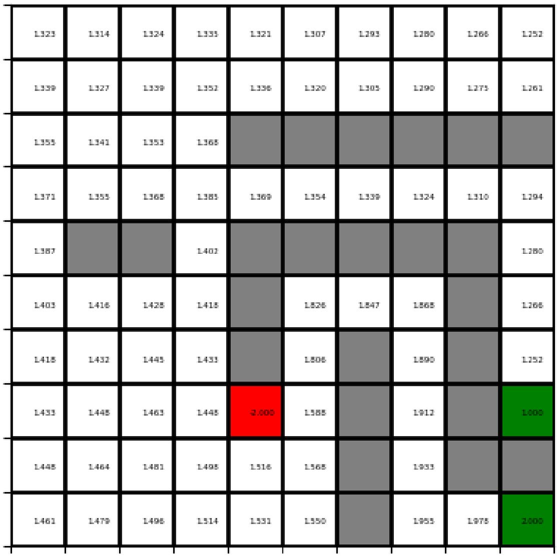
See figures. With greater discount factor (bigger gamma) it is easier to see the +2 terminal state even when close to the +1 state. Thus, gamma = 0.99 finds a better policy.

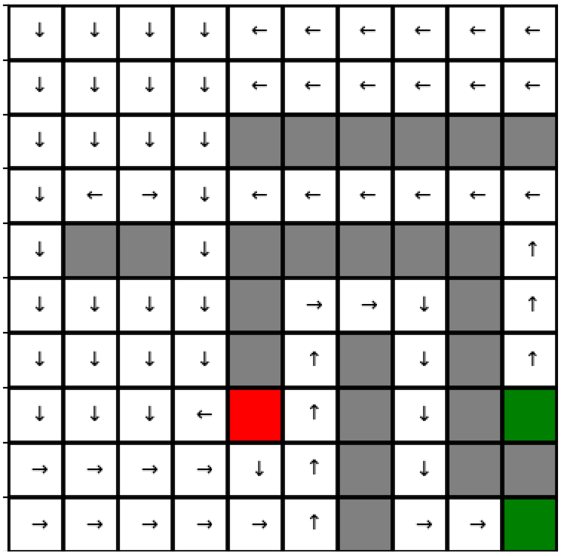
Gamma = 0.9





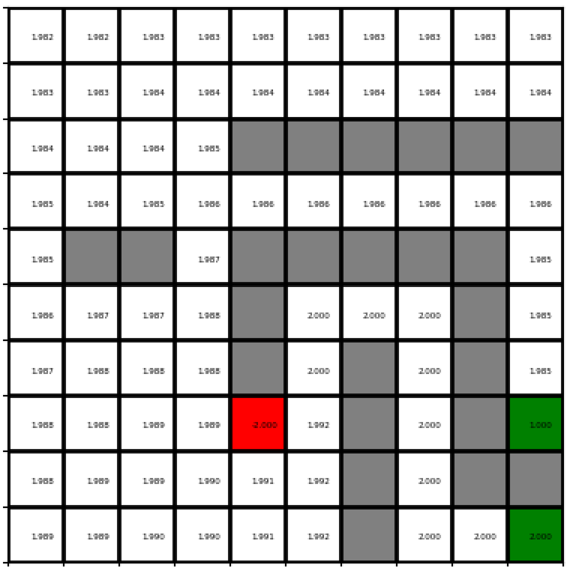
Gamma = 0.99

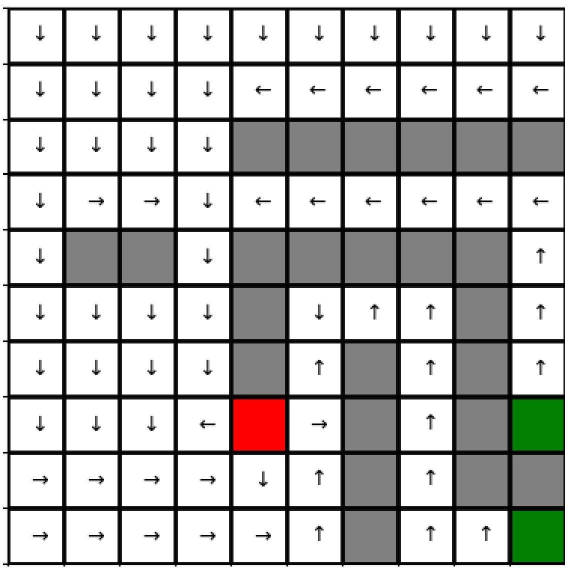




d)

With gamma = 1.0 we should from any state converge towards the +2 reward. Thus all states should have a reward close to 2 and the policy becomes kind of random in the cases where neighboring states are equally good. This is confirmed by the following results. As shown, the policy is unable to reach any terminal state.





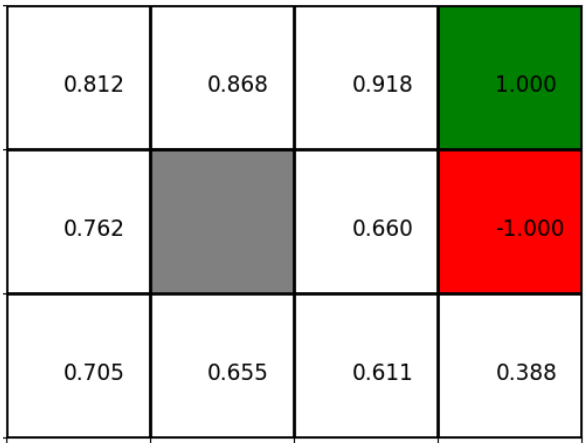
Problem 2: Policy Iteration

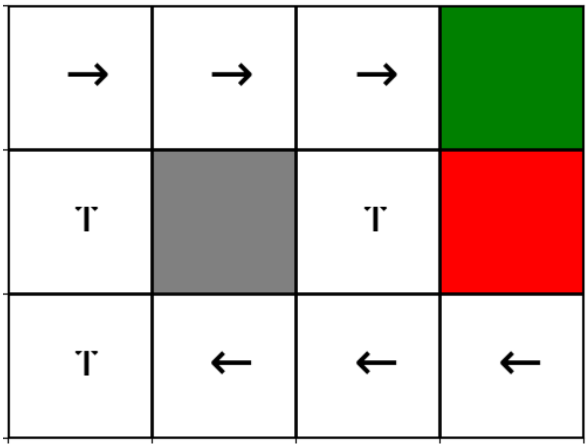
a)

I choose Iterative Policy Evaluation because it was the first algorithm in the assignment. Well, it also looks easier :P

b)

gamma = 1.0. Identical with 1a).





c)

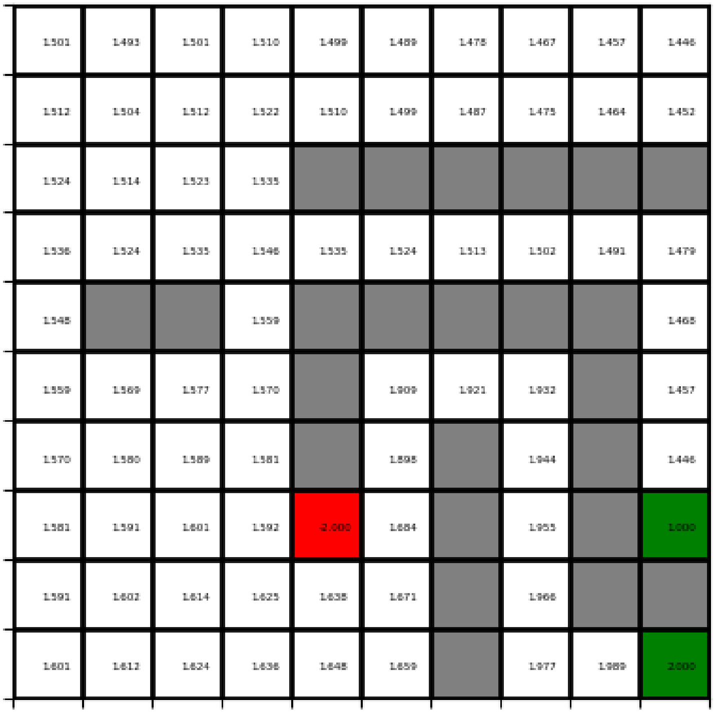
Policy iteration can change policy to an equally good one for each iteration, thus never satisfy the termination requirement of two consecutive identical policies.

d)

Value iteration uses 13 iterations and its biggest error (except from changing terminal states from 0 to +1/-1) was 0.792 on the tiny grid. For policy iteration the corresponding numbers are 7 iterations and 8.156 error.

Problem 3

Adding a negative reward of -0.01 for every non-terminal state finds a useful and optimal policy. As can be seen below, we are guaranteed to end in the +2 reward even when next to the +1 reward.


1.501 1.493 1.501 1.510 1.499 1.489 1.478 1.467 1.457 1.446
1.512 1.504 1.512 1.522 1.510 1.499 1.487 1.475 1.464 1.452
1.524 1.514 1.523 1.535
1.536 1.524 1.535 1.546 1.535 1.524 1.513 1.502 1.491 1.479
1.548 1.559 1.468
1.559 1.569 1.577 1.570 1.909 1.921 1.932 1.457
1.570 1.580 1.589 1.581 1.898 1.944 1.446
1.581 1.591 1.601 1.592 -2.000 1.684 1.955 1.000
1.591 1.602 1.614 1.625 1.638 1.671 1.966
1.601 1.612 1.624 1.636 1.648 1.659 1.977 1.989 2.000