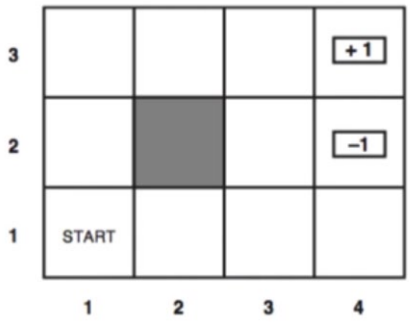
Individual assignment – solve the gridworld problem using learning the Q-function by a neural network
Due on 10/25 midnight

Solve the 3 by 4 gridworld problem using learning the Q-function by an NN.
What to report (either a 2-page IEEE format paper or a 5-page ppt):

- **Summarize your workflow to include important steps, such as data generation and minibatch in training, neural network design (initialization, how many layers, how many neurons/layer, what thresholding function used..), target network, training algorithm implementation, stopping criteria, evaluation results.**
- **Provide learning curves of training & evaluation.**
- **Provide your hyperparameters, and how you determined your hyperparameters. Illustrate the effect of different hyperparameters on your learning curve.**
- **Include a link to your code so results can be verified.**
- **In case you have multiple files, combine them into 1 pdf and upload to canvas.**

The gridworld:



- State transition: a desired direction occurs 80% of the time, 10% of the time to the left and 10% of the time to the right.
- Collision with walls results in staying at the same spot
- Two terminal states have reward +1 and -1, respectively.
- Each move at all other states, except the two terminal states, has a reward of -0.04
- The controls/actions at each state can be a move in one of the four directions N, S, W, E
- For example, in position (3,2), if the desired direction is N, then 80% of the time you will move in that direction, but 10% of the time you end up moving into the wall, while another 10% of the time you end up moving into the trap with a big penalty of -1.