Exercise 6.9: ***Windy Gridworld with King's Moves (programming)*** First, we reproduced the results of the ε-greedy Sarsa from Example 6.5, see Fig. 1. After that, we added the four diagonal steps to the set of actions. As can be expected, the addition of the diagonal steps yields much shorter path. It turns out that we have actually obtained several optimal paths of the length of 7 steps (for different random initializations), see Fig. 2. Notice that the length of the direct path from S to G is 7 steps. Thus, the addition of a ninth action which causes no movement at all except for the passive move caused by the wind cannot possibly lead to any further improvement.

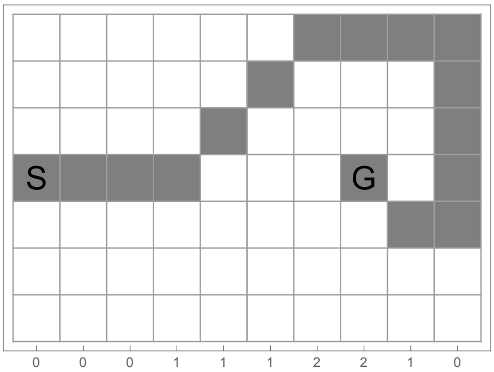

Figure 1: *Example 6.5*: The optimal trajectory (with the length of 15 steps) found by the ε-greedy Sarsa algorithm for the Windy Gridworld, where only four actions are allowed: up, down, left, and right.

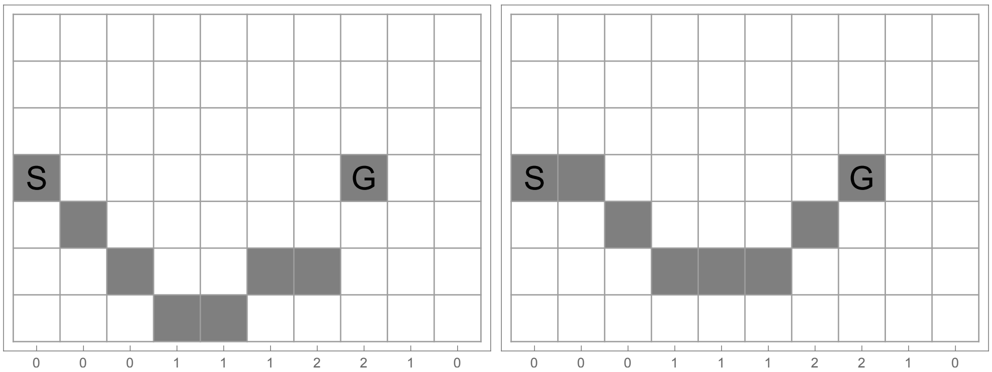

Figure 2: *Exercise 6.9*: Examples of some of the optimal trajectories found by the ε-greedy Sarsa algorithm for the Windy Gridworld with King's Moves. The length of the optimal trajectories is 7 steps.

Exercise 6.10: ***Stochastic Wind (programming)*** We have trained the ε-greedy Sarsa on the Gridworld task with King's moves in the case of the stochastic wind. See Fig. 3 for the example trajectories for two cases: with the stochastic wind turned off and with the stochastic wind turned on (however, both are trained for the stochastic case).

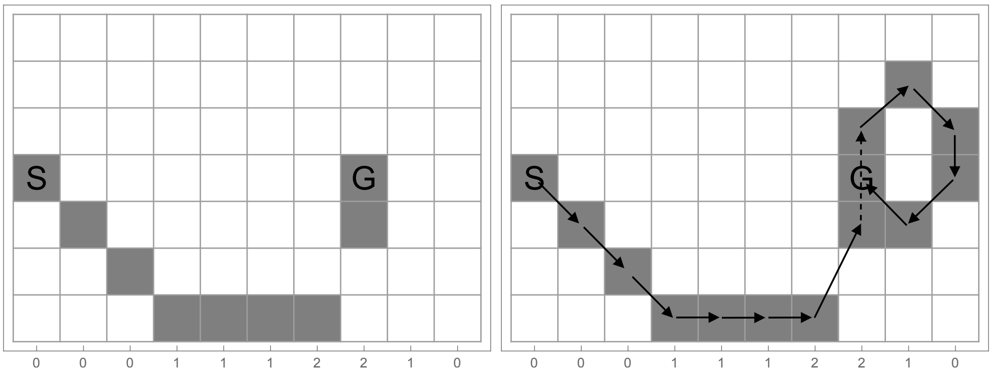

Figure 3: *Exercise 6.10*: Stochastic Wind. Example trajectories for two cases: with the stochastic wind turned off (on the left) and with the stochastic wind turned on (on the right). In both cases, training was performed with the stochastic wind.

Example 6.6.: ***Cliff Walking*** See the comparison between the solution found by Sarsa vs. Q-learning in Fig. 4.

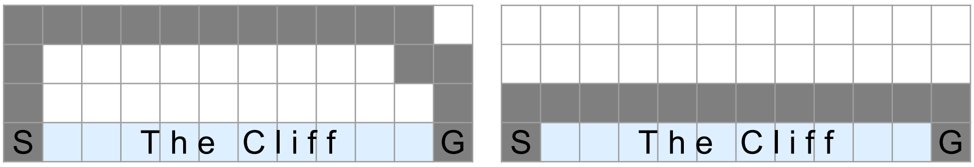

Figure 4: *Example 6.6*: Cliff Walking. The solution found by Sarsa (left) is much "safer" than the solution found by Q learning (right); however, Q-learning yields the optimal path.