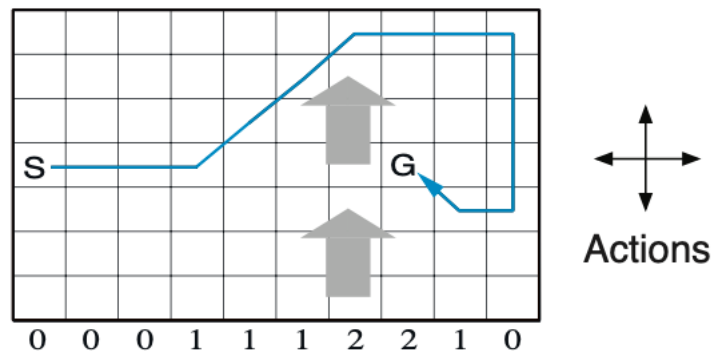


Assignment 1  
Reinforcement Learning  
CISC 453/474, Fall, 2019

Due Friday, November 22, 2019 before Midnight

**Windy Grid World (Sutton & Barto, pg. 130,131)**

Consider the game depicted in the following diagram:



You are to implement several algorithms to solve this problem:

1. Sarsa
2. Q-learning

Compare all solutions in terms of the optimal policies and episodes necessary for convergence. Select the best values for  $\epsilon$  and  $\alpha$  for each case. If they are different, discuss why.

Re-solve the windy gridworld task with King's moves, assuming that the effect of the wind, if there is any, is stochastic, sometimes varying by 1 from the mean values given for each column. That is, a third of the time you move exactly according to these values, as in the previous exercise, but also a third of the time you move one cell above that, and another third of the time you move one cell below that. For example, if you are one cell to the right of the goal and you move left, then one-third of the time you move one cell above the goal, one-third of the time you move two cells above the goal, and one-third of the time you move to the goal.

**Bonus marks:** Implement solutions using the following algorithms:

1. Sarsa( $\lambda$ )
2. Q( $\lambda$ )

## **Grading**

The assignment is worth 100 points divided as follows:

1. Implementation of all algorithms for the first case is worth 30 points. This includes the quality of your code and its correctness.
2. Comparisons of solutions for the first case is worth 20 points.
3. Implementation of all algorithms for the second case (stochastic wind) is worth 30 points. This includes the quality of your code and its correctness.
4. Comparisons of solutions for the stochastic case is worth 20 points. Your report should have a maximum of 5 pages.
5. The bonus marks are worth 10 points (5 for each correct implementation).