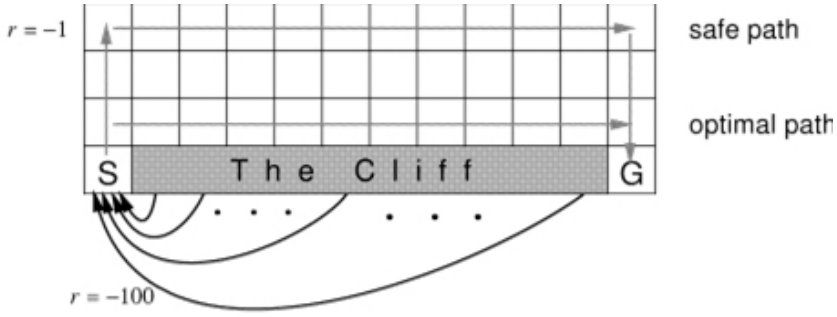
Cliff Walk

The agent lives on a regular grid with a width of 12 fields and a height of 4 fields. It starts in the field $S = (1,1)$. The agent can move one field left, right, up or down per time step. It cannot leave its world. The agent wants to maximize its reward. There is one field $G = (12,1)$ which on arrival resets the agent to the start without any reward/penalty ($r = 0$). Also, there is a cliff region $(2..11,1)$ which instantly resets the agent to the start field S but causes a reward/penalty of $r = -100$. For being on any other field the reward is $r = -1$ per time step.



Given:

- The agent has no prior knowledge.
- 500 episodes
- ϵ-greedy strategy

Tasks:

1. Decision Trees
    a. How deep is the decision tree?
    b. Is the branching factor for every node the same?
    c. Provide an upper bound for the size of the decision tree.
2. Implement an agent using the Q-Learning algorithm (see `qlearning.m`)
3. Implement an agent using the SARSA algorithm (see `sarsalearning.m`)
4. Empirical Analysis (see `compare.m`)
    a. run compare.m (ϵ is set to 0.1 in `sim_const.m`) and compare the results
    b. set ϵ to 0.05 or 0.0 and compare the results again

Attachments:

- MATLAB Scripts (`qlearning.m`, `sarsalearning.m`) which have to be filled at the marked positions
- MATLAB Script `sim_const.m`: contains constants like ϵ etc.
- MATLAB Script `compare.m`: simulates both agents and plots graphs

Solution will be discussed on July, 2