Methods for value-based RL (2/2) exercise

Methods for value-based RL (2/2) exercise

Exercise

- Implement SARSA and/or Q-Learning and/or MC evaluation
- Plot the Q-values after training with the provided function
- Try out the cliff environment. Can you verify the different behaviour?
- Create your own environment and play around ☺

Note :

- Use the `step` and `reset` functions of the GridWorld environment (and not the `step_dp` function)