

Homework

Task Frozen Lake MDP

Now you will implement value iteration and policy iteration for the Frozen Lake environment from `OpenAI Gym`. We have provided custom versions of this environment.

1. **(coding)** Read through `vi_and_pi.py` and implement `policy_evaluation`, `policy_improvement` and `policy_iteration`. The stopping tolerance (defined as $\max_s |V_{old}(s) - V_{new}(s)|$) is $\text{tol} = 10^{-3}$. Use $\gamma = 0.9$. Return the optimal value function and the optimal policy.
2. **(coding)** Implement `value_iteration` in `vi_and_pi.py`. The stopping tolerance is $\text{tol} = 10^{-3}$. Use $\gamma = 0.9$. Return the optimal value function and the optimal policy.

Submission

Submit a zip file that includes your source code and PDF of report to the email(zhangyc8@mail2.sysu.edu.cn)

Naming format of zip file: RL_Student ID_Name_Assignment ID, e.g.: 20220315_张三_实验9.zip

The environment code can be available at

https://github.com/ZYC9894/SYSU_RL2022 or <https://github.com/joenghl/2022AI/tree/main/Homework/Experiment>