Project 1: Markov Processes and Dynamic Programming

*Collaboration in the sense of discussion is allowed, however, the assignment is **individual** and the work you turn in should be entirely your own. See the collaboration and academic integrity statement here:* https://natanaso.github.io/ece276b. *Books may be consulted but not copied from. Please acknowledge **in writing** people you discuss the problems with and provide references for any books or papers you used.*

Submission

Please submit the following files on **Gradescope** by the deadline shown at the top right corner.

1. Programming assignment: upload all code you have written for the project and a README file with a clear, concise description of how to run your code.
2. Report: upload your report in pdf format. You are encouraged but not required to use an IEEE conference template[1] for your report.

Problems

In square brackets are the points assigned to each part.

1. This project will focus on autonomous navigation in a **Door & Key environment**. The goal is to get our agent (red triangle) to the treasure location (green square). The environment may contain a door which blocks the agent's motion. The agent needs to pick up a key to unlock the door. See Fig. 1 for an illustration. The agent has three regular actions, *move forward* (MF), *turn left* (TL), and *turn right* (TR), and two special actions, *pick up key* (PK) and *unlock door* (UD). Taking any of these five actions costs energy (positive cost). The objective is to design a Dynamic Programming algorithm that minimizes the cost of obtaining the treasure. Further description of the environment and functionality is provided in the accompanying starter code and README file.

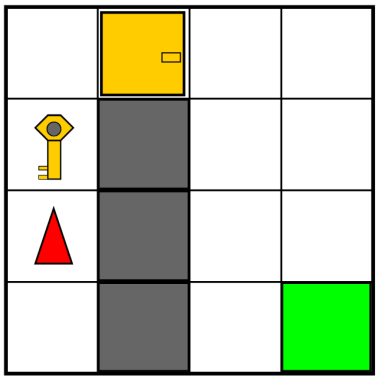

Figure 1: Door & Key Environment: An agent (red triangle) needs to navigate to the treasure (green square). The white squares are traversable, while the gray squares are obstacles that the agent cannot go through. The agent may need to pick up a key and unlock a door along the way to the treasure.

   Write a project report describing your approach to the door and key problem. Include the following sections:

   (a) [5 pts] **Introduction:** provide a short introduction to the problem, its potential application to robotics, and your approach at a high level.

   (b) [30 pts] **Problem Statement:** formulate the problem as a Markov Decision Process. Clearly define the state space $\mathcal{X}$, control space $\mathcal{U}$, motion model f or p_f, the initial state x_0, the planning horizon

[1] https://www.ieee.org/conferences_events/conferences/publishing/templates.html

T, the stage and terminal costs ℓ, $\mathfrak{q}$, and any other elements necessary to make this a well defined problem.

(c) [25 pts] **Technical Approach:** describe your algorithm for obtaining an optimal control policy that minimizes the agent's cost. This should be a clear description of the algorithm that you implemented in python.

(d) [20 pts] **Results:** use your python implementation to determine an optimal navigation policy on the 7 test environments provided in the starter code. You results should include a visualization of the value function, policy function, and the action sequence taken by the agent to obtain the treasure. For example, an action sequence minimizing the cost in Fig. 1 is:

$$PK \rightarrow MF \rightarrow MF \rightarrow TR \rightarrow UD \rightarrow MF \rightarrow MF \rightarrow TR \rightarrow MF \rightarrow MF \rightarrow MF \rightarrow TR \rightarrow MF$$

Your results section should also include a discussion: what worked, what did not, and why.

(e) [20 pts] **Code:** your code will be evaluated based on correctness and efficiency. The results presented in the report should be consistent with the output of your code. The code should have a clear structure and comments. **Copying, rephrasing, or even looking at anyone else's code is strictly forbidden. Writing your own code is the best way to avoid plagiarism discussions or consequences.**