

# CS4013/5013 Assignment 6

Fall 2024

Due Date: Nov 24, 2024.

## Task 1. Use Monte Carlo to Estimate Utility

Implement first-visit Monte Carlo and use it to estimate the utility of each non-terminal state based on the policy in Figure ?? (left) and stochastic move in Figure ?? (right). Set discount rate  $\gamma = 0.8$  and the reward at each non-terminal state  $r = -0.04$ . To estimate the utility of each state, run at least 10 experiments. Report your estimated utilities in the following table.

State	(1,1)	(1,2)	(1,3)	(2,1)	(2,3)	(3,1)	(3,2)	(3,3)	(4,1)
Utility	-0.1996	-0.1992	-0.1992	-0.1995	-0.1944	-0.1973	-0.1898	-0.1928	-0.1845

## Task 2. Optimize policy based on your estimates.

Based on your estimated utilities, optimize your policy at each state using Bellman's equation. (See detailed method in page 31 or 35 of '1104\_mod7a.pdf'.) You can write a program to automatically estimate it or manually estimate it. In either case, report your updated policy in the following table. You may use R for right, L for left, U for up and D for down.

State	(1,1)	(1,2)	(1,3)	(2,1)	(2,3)	(3,1)	(3,2)	(3,3)	(4,1)
Old Policy	R	U	L	U	D	L	L	R	L
New Policy	R	R	D	D	D	D	U	L	D