CSCI 4511/6511 - Exam Prep 7

Instructions:
This is ungraded exam prep.

1 Markov Processes

The Ignatius Coffee Company in Cleric, AZ, has two espresso machines and two baristas. It takes about five minutes for one barista to brew one cup of coffee. There is room for four people total to wait for their coffee.

- For this problem, model each time step as five minutes in length.
- If a customer enters during a time step *and* an espresso machine is free:
  - The customer's coffee is brewed during that time step
- Otherwise, if a customer enters during at time step and no machine is free:
  - The customer waits
- In any time step, there is an equal probability that 0, 1, 2, or 3 customers will arrive.

Ignatius himself has decreed that, after 5 PM, the shop will stay open until no more customers are waiting, after which the shop will close (and no more customers will be accepted to queue for coffee).

Model the number of customers waiting for coffee as a Markov chain.

2 Stationary Markov Processes

During the day (from 8 AM until 5 PM), there is an equal probability that 0, 1, 2, 3, 4, or 5 customers will arive in any five-minute step. After 5 PM, the probabilities for customer arrival are the same as the previous problem.

What is the probability distribution for the number of customers waiting at the shop at 5 PM? Model the problem as a Markov chain and explain how you would solve it– you do **not** need to compute a solution, but your solution must explain every step of the computation.

3 Markov Reward Processes

The coffee shop nets $1 per cup of coffee served (after expenses). Assuming no other changes in the business, how much do profits rise if a fifth waiting spot is added?

Model the problem as a Markov reward process and explain how you would solve it– you do **not** need to compute a solution, but your solution must explain every step of the computation.

4 Markov Decision Processes

The Ignatius Coffee Company in Cleric, AZ roasts coffee for Carl's Diner in South Cleric. Every week, Ignatius has to order coffee beans from his supplier in Honduras. The beans arrive the following week.

Ignatius has noticed a pattern in Carl's ordering:

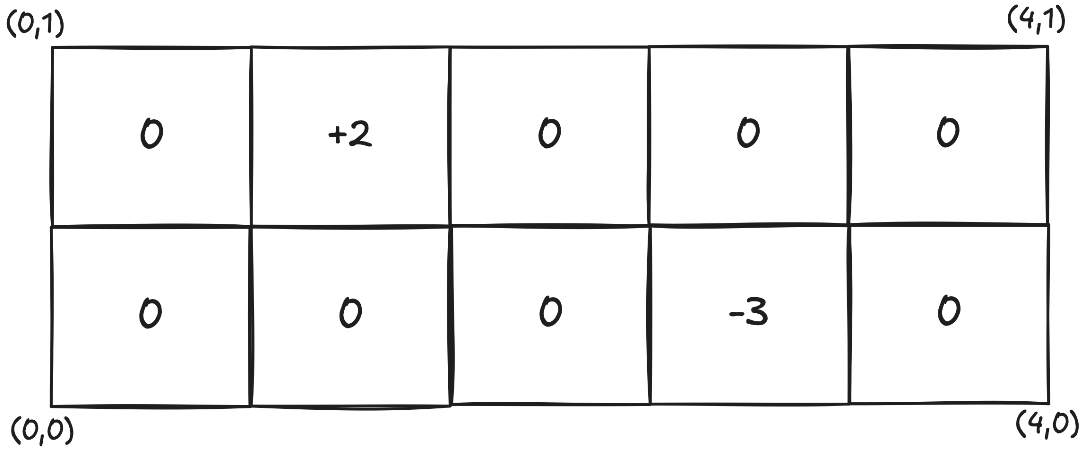
- Carl orders at least 40, and not more than 50 bags of coffee each week.
- Each week's order is within two bags' of the previous week's order, with equal probability within these values:
  - If Carl ordered 45 bags last week, the next week is equally probable to be 43, 44, 45, 46, or 47 bags.
  - If Carl ordered 40 bags last week, the next week is equally probable to be 40, 41, or 42 bags.
- Ignatius can order any number of bags of coffee from his supplier:
  - Bags of unroasted coffee cost Ignatius $2
  - Ignatius sells roasted coffee bags to carl for $10 each
  - Ignatius values every missed order[1] at negative $3.

Model Ignatius's decision for coffee ordering as a Markov Decision Process. You only need to set up the problem: show all states, state transitions, and rewards, and describe mathematically an algorithm to solve the problem. You do not need to provide a solution.

[1]A missed order is a bag of coffee that Carl would like to buy, but can't, because Ignatius is out of stock

5 Reinforcement Learning

For a problem with states arranged on the following grid, with rewards for entering each grid square shown:



- Consider an action space where an agent can deterministically move to any adjacent (shares an edge) state.
- Use a learning rate α of 0.5
- The problem will terminate when reward exceeds 4 or falls below -4; do not discount.
- Start with $Q(s, a) = 0 \quad \forall s, a$

5.1 Sarsa

Show how the Sarsa algorithm updates $Q(s, a)$ for an agent starting at point $(0, 0)$. moving one state to the right, and then moving one state up.

5.2 Sarsa-λ

Show how the Sarsa-λ algorithm updates $Q(s, a)$ for an agent starting at point $(0, 0)$ and moving right three times. Use a λ value of 0.5.

Intentionally blank.

This is the back of the exam prep.