CS 4100
artificial intelligence

Summer 2 2024, Khoury College, Northeastern University

Problem Set 2

Instructions

- All submissions must be typed. No exceptions are made to this rule.
- Hand-drawn figures are acceptable only where specified in the question, and provided all labels are clear and legible. If we can't read your text, we can't assign points. Please scan and insert any such figures in the final PDF document.
- You may not, as a general rule, use generative AI for problem sets. Any use of GenAI that is **solely** intended to improve writing clarity or sentence structure is acceptable, but must be accompanied by an appendix containing your original, unedited answers. If you use generative AI in this manner, start a new page titled 'Appendix' at the end of the written submission, and paste your original answers here with the corresponding question numbers clearly indicated.
- If you discuss this problem set with one or more classmates, please ensure that all parties must declare collaborators on their individual submissions. Such discussions must be kept at a conceptual level, and no sharing of written answers is permitted.

Deadlines

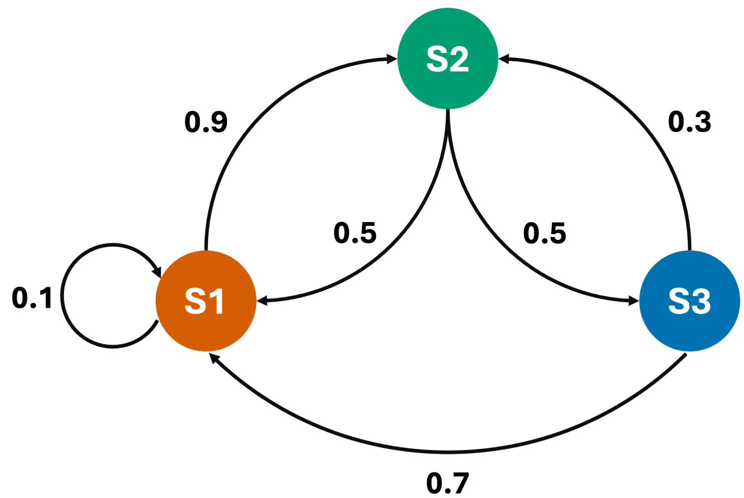
- Written submissions should be uploaded to Gradescope by 6:00 PM on 7/28/2024.
- Gradescope will show a 'late' deadline of 7/31. This is intended solely for any students who may wish to invoke the freebie, outlined in the course policies.
- Any submissions received after 6:00 PM on 7/28 will be considered late, and will automatically invoke the use of your freebie.
- Regrade requests must be submitted on Gradescope within 1 week of receiving your grade, after which no further requests will be entertained.

Reach Out!

If at any point you feel stuck with the assignment, please reach out to the TAs or the instructor, and do so early on! This lets us guide you in the right direction in a timely fashion and will help you make the most of your assignment.

Markov Models

Consider the following Markov model with 3 states, **S1**, **S2** and **S3**. The transition probabilities of the model are represented along the edges (note that each edge is directional).



Q1. Given the above Markov model, explain whether the stationary distribution depends on the start state (without actually computing the stationary distribution). (2)

Q2. Given the initial probability distribution $p_0 = [0.1, 0.6, 0.3]$ for states **S1**, **S2**, and **S3** respectively, compute the stationary distribution for the given Markov model. You may use a program to do the computations. Report only the final stationary distribution π. (2)

Q3. What is the probability of the following transition sequence: **S1** $\rightarrow$ **S1** $\rightarrow$ **S2** $\rightarrow$ **S3**? (Use the stationary distribution computed above, and show your calculations.) (2)

Q4. Given that we start in the state **S2**, what is the probability of returning to **S2** after two transitions? (Show your calculations.) (2)

Q5. Given the starting probability distribution $p_0 = [0.1, 0.6, 0.3]$, what is the probability of ending up in **S2** after exactly two transitions? (Show your calculations.) (2)

Hidden Markov Models

Recall Hidden Markov Models (HMM) from class, where we applied this approach to sequence labeling tasks such as parts-of-speech tagging or named entity recognition. Here, your task is to construct and use an HMM model to make inferences about a coin-flipping game with the following rules.

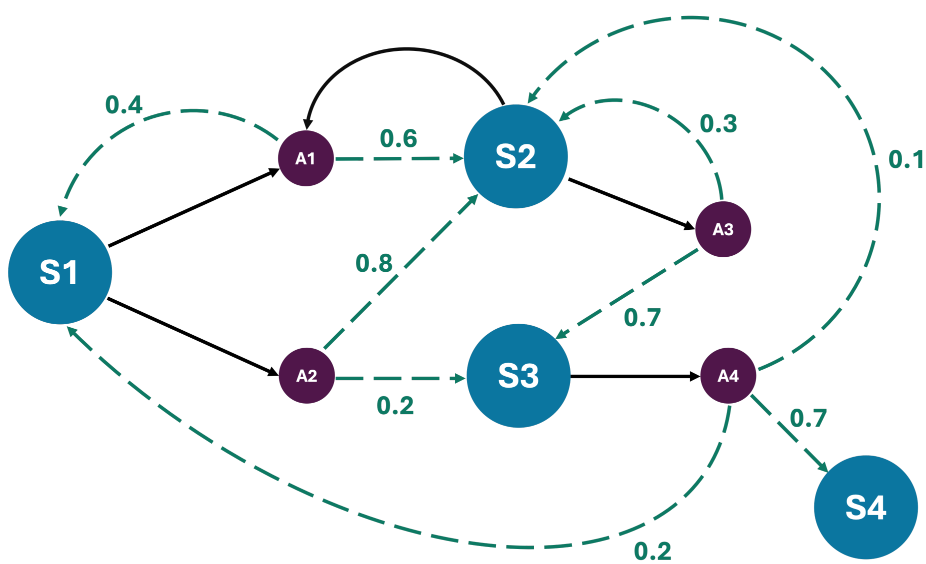
Your professor produces two identical looking coins. However, only one of the coins is a fair coin, and the other is a biased coin that produces an outcome of **Heads** 70% of the time. The professor always knows which coin is the fair one, and will perform three coin flips in total. Between each flip, the professor may swap the coin, following the rule that if a fair coin is flipped in one round, then in the next round, the professor chooses a coin completely at random. If, however, the biased coin is flipped in any round, then the professor is twice as likely to choose the biased coin again in the next round, as compared to the fair coin. As the three flips are performed, you observe the outcomes **Heads**, **Heads** and **Tails** respectively.

Q6. Draw the HMM diagram for this game showing transitions between hidden states, and emissions to outcomes with the corresponding probabilities labeled along the edges. Hand-drawn figures accepted for this question, provided the grader can read everything clearly. (For reference, see the second diagram in our HMM notes, showing the *Very Late*, *Late* and *On Time* hidden states, and the *Happy* and *Sad* outcomes.) (5)

Q7. Given the observed outcome sequence, predict which coin was most likely flipped in each of the three turns (i.e., compute the most likely hidden sequence). Show all your intermediate calculations and use the stationary distribution to reason about which coin was flipped in the very first round. (10)

Markov Decision Processes and Reinforcement Learning

Consider the following MDP with 4 states, and 4 actions. A dashed line represents a transition from a chosen action to some next state. Transition probabilities are specified along each dashed edge.

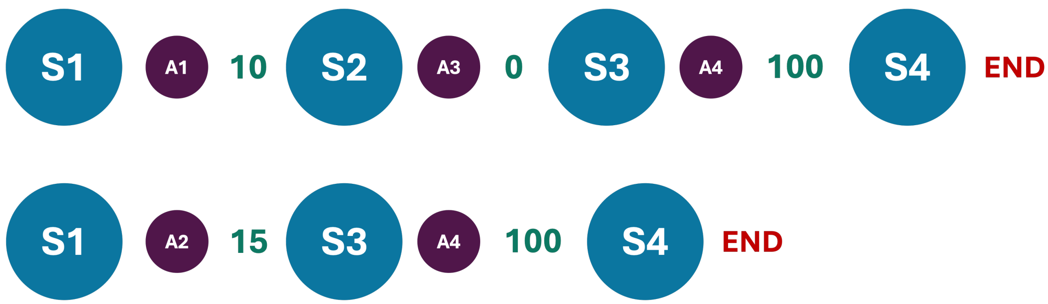


Q8. How many unique policies does this MDP have? Explain your reasoning and list all policies. Use **X** to indicate no possible action from a state. (2)

Q9. If from any state, all valid actions are equally likely, then what is the total probability of reaching S4 from S1 using paths of at most length 3? List all such paths and compute the total probability. Show your calculations. (An action followed by a transition into a next state counts as a total of one move.) (4)

Q10. Given that $R(S1, A1, S2) = 10$, $R(S1, A2, S2) = 10$, $R(S1, A2, S3) = 15$, $R(S3, A4, S4) = 100$, and that rewards for all other transitions are 0, write and expand the optimal value function equation for $V_{opt}(S1)$. Assume that the discount factor is γ, and leave your final answer in terms of $V_{opt}(S2)$ and $V_{opt}(S3)$. (4)

Q11. Assume that by simulating this MDP using some exploration policy π, we obtain the two following episodes:



Use Q-learning updates to calculate the agent's final optimal policy given this data stream, and show all intermediate steps. Assume $\gamma = 1$. For your reference, the Q-learning update equation is given by: (10)

$$\eta = \frac{1}{1 + \text{number of updates to } \hat{Q}_{opt}(s, a)}$$

For each observed (s, a, r, s'):

$$\text{Estimate, } \hat{Q}^{(t)}_{opt}(s, a) = (1 - \eta)\hat{Q}^{(t-1)}_{opt}(s, a) + \eta[R(s, a, s') + \gamma \hat{V}^{(t-1)}_{opt}(s')]$$
$$\text{where } \hat{V}_{opt}(s') = \max_{a'} \hat{Q}_{opt}(s', a')$$

Academic Integrity

Q12 Review, and copy/paste the following academic integrity acknowledgement in your final submission as the answer to Q12.

I have read and understood the academic integrity policy as outlined in the course syllabus for CS4100. By pasting this acknowledgement in my submission, I declare that all work presented here is my own, and any conceptual discussions I may have had with classmates have been fully disclosed. I declare that generative AI was not used to answer any questions in this assignment. Any use of generative AI to improve writing clarity alone is accompanied by an appendix with my original, unedited answers.