CS3630 Project 2 Report (Fall 2022)

Name: Moungsung Im
GT Email: mim30@gatech.edu
GT Username: mim30

1.1 State Abstractions [2 pts each]

Identify whether the following situations follow a Bayesian or Frequentist approach. Give a short explanation why.

a) Find whether a coin is biased or not

It follows a Frequentist approach. We can flip a coin many times and see whether the ratio of head and tail is 5:5. If so, the coin is fair, and if not, the coin is biased.

b) Find the probability of getting an A in CS 3630

It follows a Bayesian approach. If we know the prior data of the ratio of A in the course has been X for certain number of semesters, we can assume the probability of getting an A grade in this semester would be X.

c) Find how long it will take for RoomBuzz to run out of charge

It follows a Frequentist approach. If X minutes takes for RoomBuzz to run out of charge, we can assume that it will run out of charge in X minutes for the next execution.

1.2 State Abstractions [3 pts]

Suppose RoomBuzz needs to charge every night, but the charger is located in the Office. RoomBuzz's robust software will consistently route it to the charger at the end of the day. What is the prior probability distribution of the robot's state every morning?

0/0/1/0/0
in order of Living Room, Kitchen, Office, Hallway, Dining Room

2.1 Actions over time [4 pts]

It's the beginning of the day, and RoomBuzz undocks from its charger in the Office. It chooses action "R" in order to clean the hallway. What is the probability that RoomBuzz does not end up in the hallway?

0.2

2.2 Actions over time [4 pts]

It's dinnertime and RoomBuzz has an 80/20 chance of being in the Dining Room or Kitchen, respectively. It takes action "L". What is the PMF over RoomBuzz's resulting belief state?

0.16/0.04/0/0.64/0.16
in order of Living Room, Kitchen, Office, Hallway, Dining Room

3.1 Dynamic Bayes Nets [4 pts]

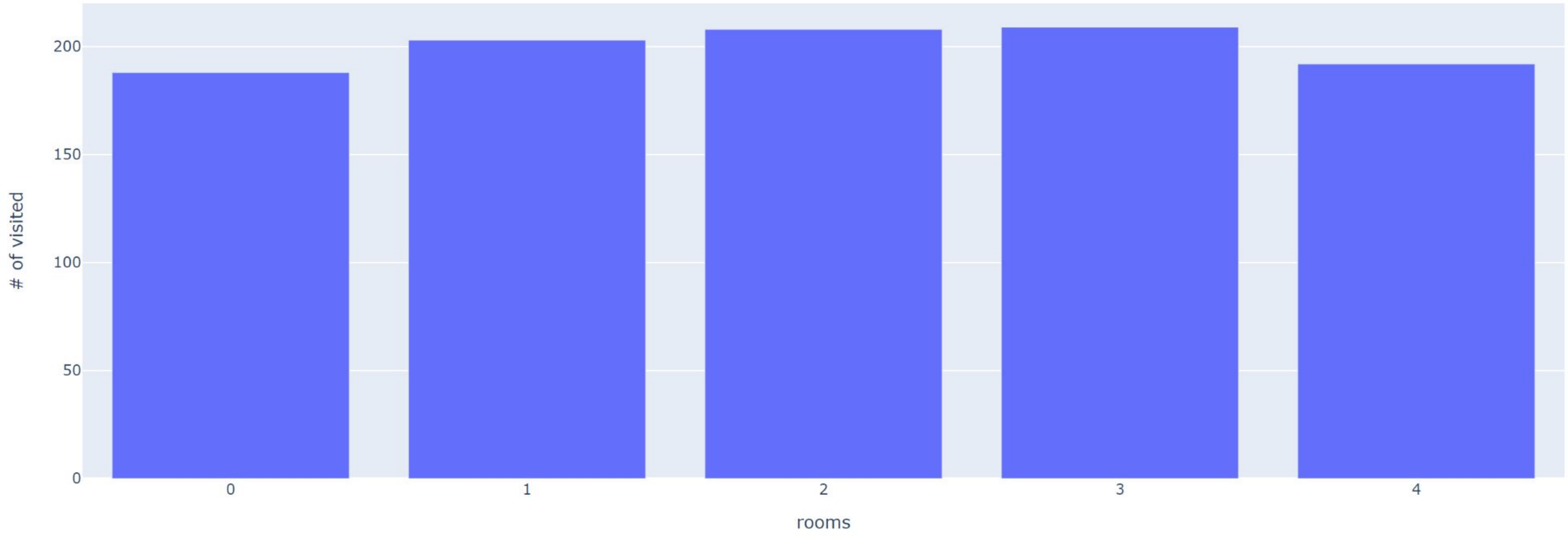
Notice that sometimes, multiple different light observations are made in the same state. Explain why. Can we use these light readings to infer what state we are in?

It is sometimes ambiguous that the state of robot although we have light measurement value. We cannot only use light measurement to infer what state we are in because for example, if measurement of being light is 0.8, the state can be Living Room or Kitchen. We need sensor and action statement to infer where we are.

3.2 Dynamic Bayes Nets [4 pts]

Run ancestral sampling 1000 times to find the **initial** state using the prior from get_prior() and the action sequence from create_all_left_action_sequence(). Plot a histogram with the results. Why do we see the resulting distribution? Paste your histogram in your report.

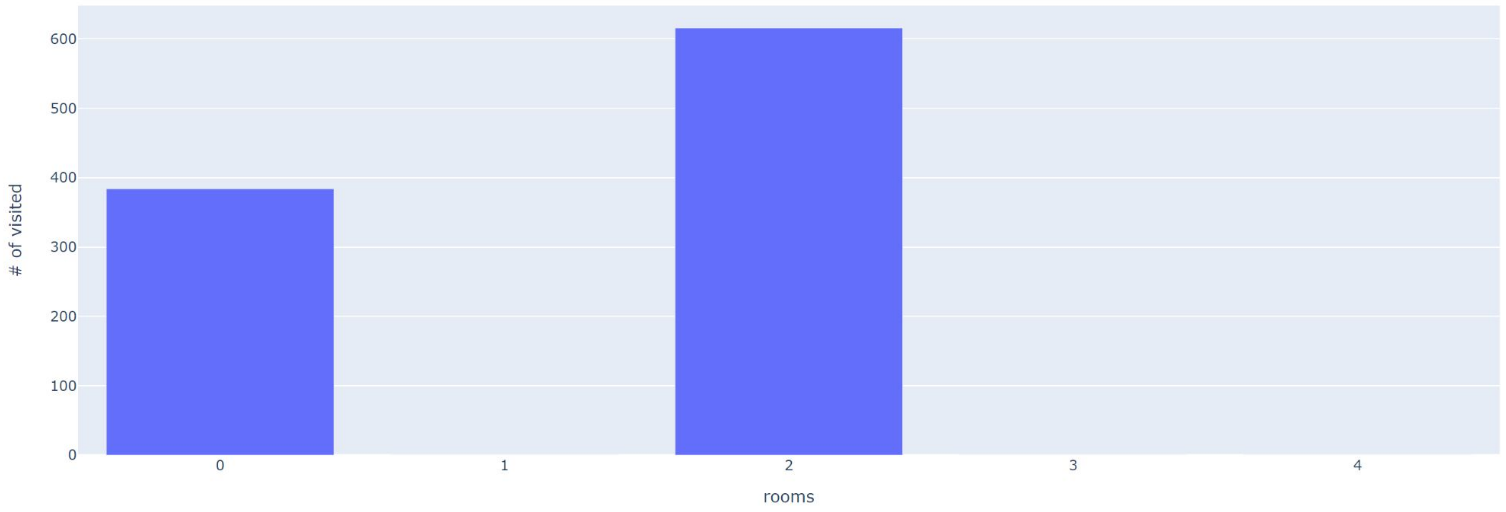
Because get_prior returns the priors to have uniform distribution over all the rooms, the initial state will be equally distributed. It means the robot can start from one of every location equally likely. The graph would look slightly different from each sampling, but the point is every state has equal probability of being the initial state.



3.3 Dynamic Bayes Nets [4 pts]

Run ancestral sampling 1000 times to find the **final** state using the prior from get_prior() and the action sequence from create_custom_action_sequence(). Plot a histogram with the results. What do you notice? Explain why we see the resulting distribution. Paste your histogram in your report.

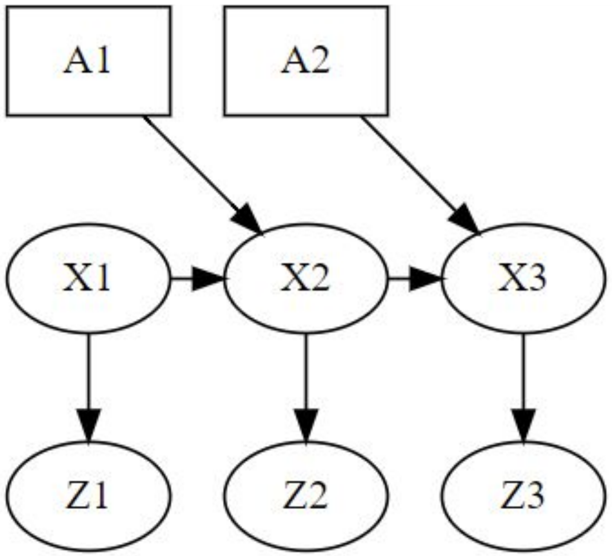
The final state would be Living Room or Office. Since the custom_action_sequence makes the robot goes left multiple times and then goes up, the robot would reach to Living Room or Office. Although it moves up from there, the location of the robot will remain the same. The ratio of Living Room and Office would be 2:3 because get_prior returns uniformly distributed probability of initial state, therefore P(Living Room) + P(Kitchen) = ⅖ and P(Office) + P(Hallway) + P(Dining Room) = ⅗.

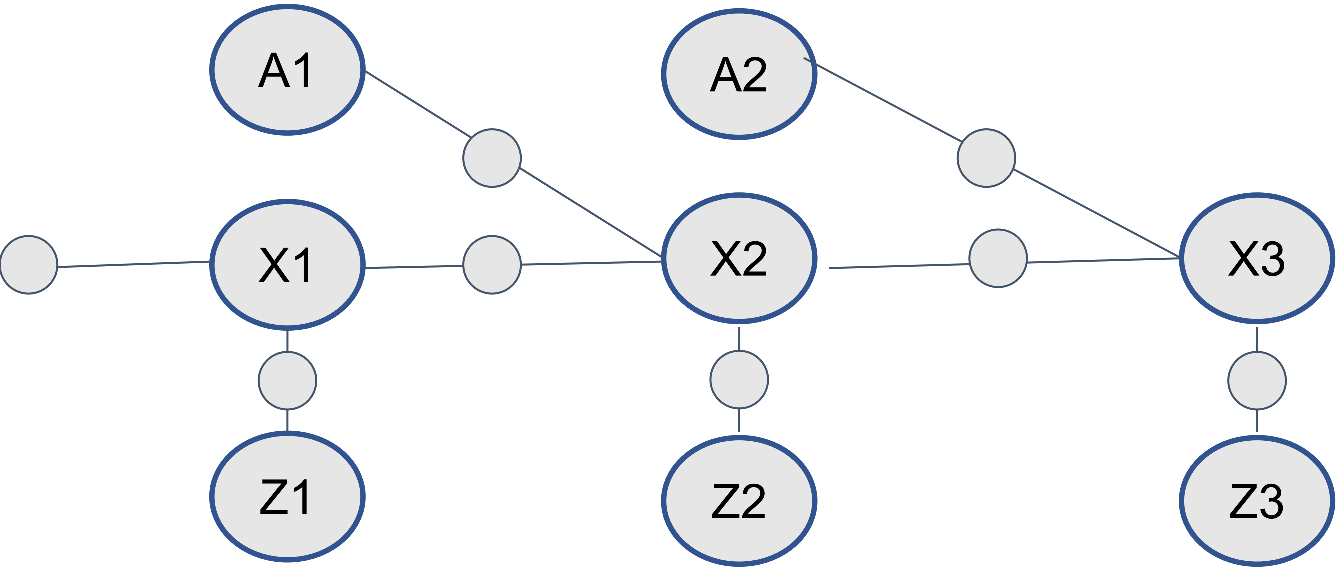


4.1.A Perception w/ Graphical Models [2 pts]

Given the DBN from the textbook,

- If all actions and measurements are unknown, what does the factor graph look like?

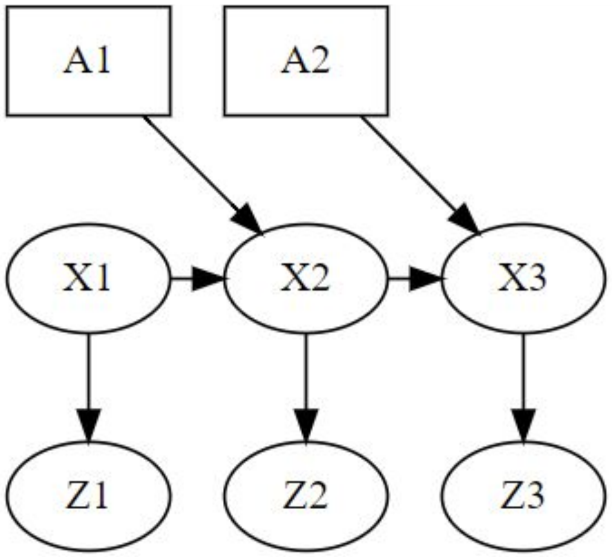


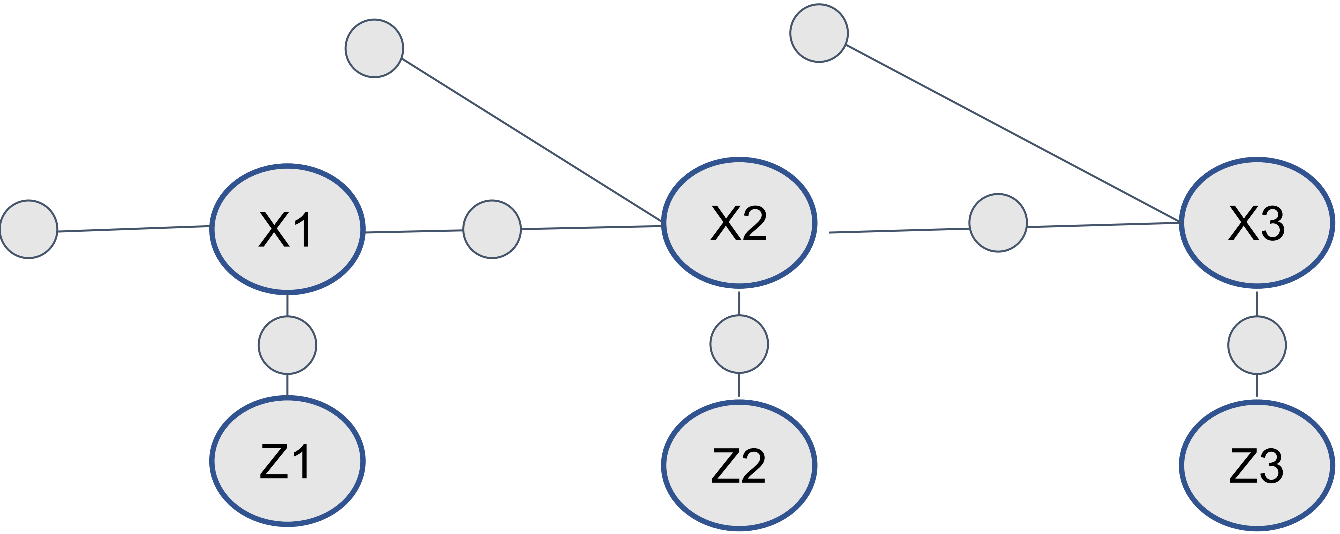


4.1.B Perception w/ Graphical Models [2 pts]

Given the DBN from the textbook,

- If the actions are known but measurements are unknown, what does the factor graph look like?

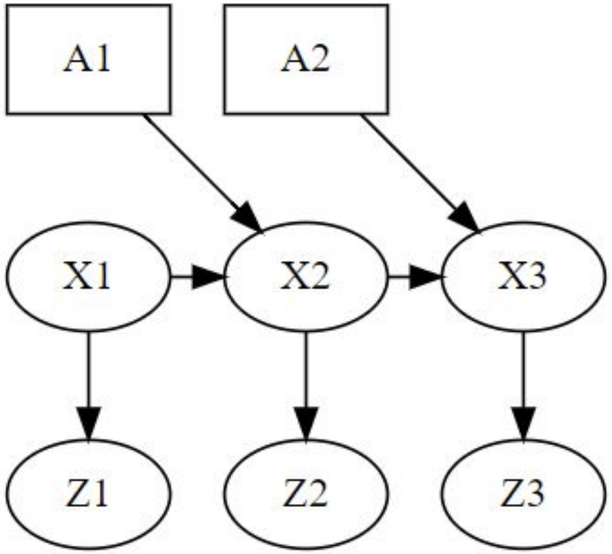


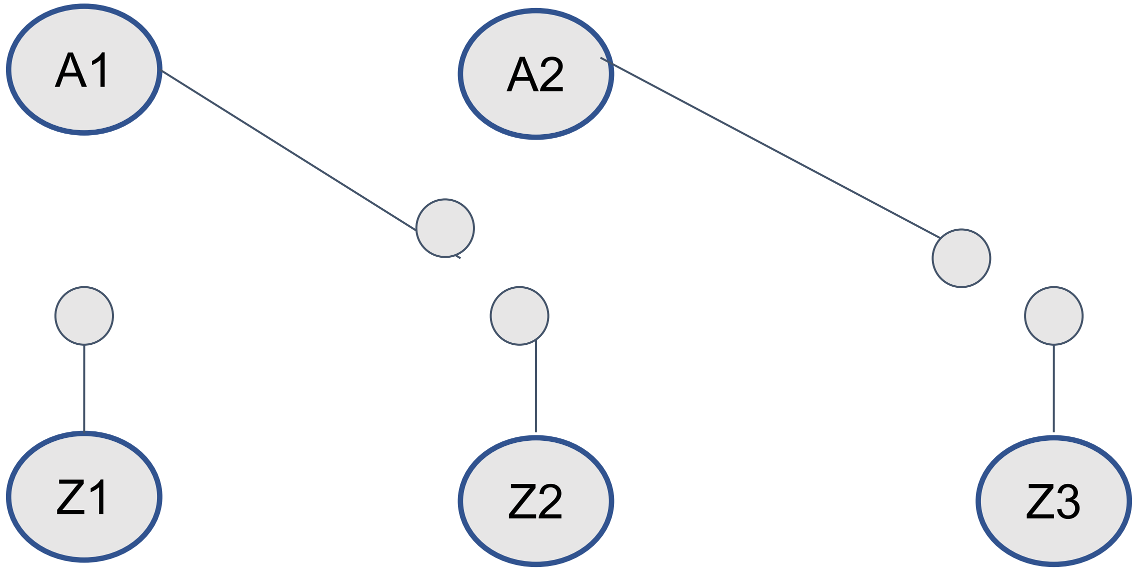


4.1.C Perception w/ Graphical Models [2 pts]

Given the DBN from the textbook,

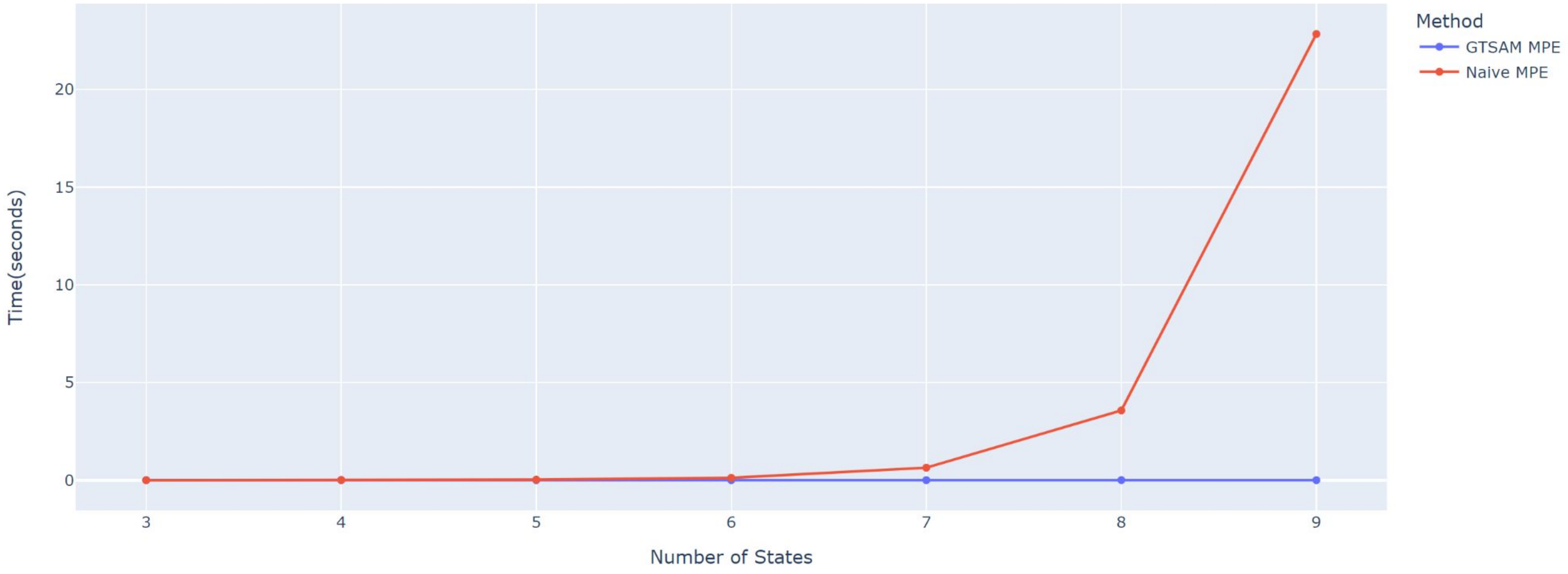
- If the actions and the measurements are unknown, but the trajectory of the states is known, what does the factor graph look like? Do you observe any direct findings from the factor graph?





4.2 Perception w/ Graphical Models [2 pts]

Please plot the graph from the coding section which compares the time complexity of Naïve MPE implementations vs GTSAM implementation



4.3 Perception w/ Graphical Models [2 pts]

What is the time complexity of MPE when enumerating over an N different number of states?

- Exponential(e^x+c)

5.1 Markov Decision Process [2 pts]

What is a rollout? What is rollout reward?

Rollout is the process of evaluating the discounted reward for such a sequence. Each rollout produces one sample trajectory, and one corresponding discounted reward.
Rollout reward is a reward that a robot can eventually obtain from the expected trajectory of a rollout following given actions. In our example, if the final state of a trajectory from a rollout is Living Room, it will have rollout reward which is 10. If not, it will have zero reward.

5.2 Markov Decision Process [3 pts]

Compare the two control tapes provided after TODO 18. What is the optimal control tape? Why?

The second control tape is the optimal control tape because it has higher average reward. It is obvious that starting from office, if we go Right, and then try 3 UPs, unless the first right or all of three up fails after Right, we will reach to Living Room and get the reward. The first control tape, if all of the given actions succeed, the robot eventually reach to Kitchen. In order to obtain the reward, the first two actions must succeed and the last two must fail. The probability of fail is 0.2, therefore it is harder to fail than to succeed. In conclusion, the second control tape is optimal to get reward because it is easier to succeed all actions and end up in Living Room rather than succeeding half and failing half.

5.3 Markov Decision Process [3 pts]

What is the optimal policy? Give a short explanation on what this policy tells us.

The optimal policy is that a policy of the optimal actions at any moment in time depends on the state in which the action is executed.
It tells us which action is better to be executed at each state to gain a reward.
The optimal policy for our example is L L R U U.
In Living room, the robot is better stay there.
In Kitchen, the robot is better to go left.
In Hallway, the robot is better to go up.
In Office, the robot can go to Living room through Hallway, so it has to go right to go to Office.
In Dining Room, it can go to Living room through Kitchen, so it has to go up to go to Kitchen.

5.4 Markov Decision Process [3 pts]

What's the objective of using policy iteration? What's the objective of using value iteration?

The objective of using policy iteration is improve a policy. Starting with a guess, the policy iteration improve our guess until until no further improvements are possible.
The objective of using value iteration is to approximate optimal value. Value iteration operates by iteratively using our current best guess of the optimal value along with the known expected reward to update the approximation.

6.1 *EXTRA CREDIT* RL [2 pts]

In the given equation, please define the variables x, a, x', γ, V*, P, and $\bar{R}$.

$$Q^*(x,a) \doteq \bar{R}(x,a) + \gamma \sum_{x'} P(x'|x,a) V^*(x')$$

x = state
a = action performed
x' = future state after doing action a at state x
γ = discount factor
V* = optimal value function
P = probability of x' given that x, a
$\bar{R}$ = reward function

7. Feedback

Please provide feedback on the coding portion of the project. How did it help your understanding of the material? Is there anything that you think could have been made more clear?

I wish we can test our codes in colab, not gradescope. Gradescope is too slow to check my code.
I want to say the textbook and lectures are very good and appreciate for your hard work.
The material is hard, but I enjoy this class.