CS3630 Project 2 Report (Fall 2022)

Name: Anthony Wong
GT Email: awong307@gatech.edu
GT Username: awong307

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

Frequentist. Every time you flip a coin, there is an equal chance of heads or tails. It is not affected by previous flips.

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

Bayesian. We can say that the probability of getting an A can depend on how long you spend a week doing work for the class.

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

Bayesian. We can say that the length of time it takes RoomBuzz to run out of charge depends on how long it was charged for previously.

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?

If we define our states as such,

["Living Room", "Kitchen", "Office", "Hallway", "Dining Room"]

our prior probability distribution would look like:

[0, 0, 1, 0, 0]

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?

According to the conditional probability table defined in 3.2.2, there is a 20% chance that our “R” action is unsuccessful and we are unable to reach the hallway.

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?

The pmf for the resulting belief state is :

Hallway - 64%
Dining Room - 16%
Kitchen - 4%
Living Room - 16%

3.1 Dynamic Bayes Nets [4 pts]

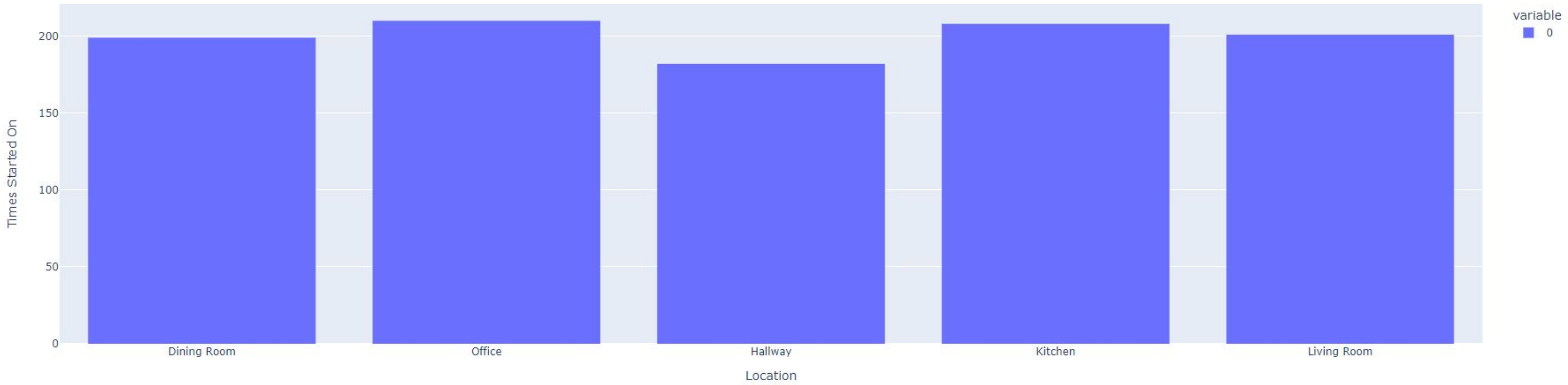
Notice that sometimes, multiple different light observations are made in the same state. Explain why. Can we use these light readings to infer what state we are in?

This is because the can be both error and variation in the light observations depending on factors such as the time of day. We can still use these light readings to help infer what state we are in. However, if we look at Living Room and Kitchen for example, both which have the same probability distributions, we can only use light readings to supplement things such as actions.

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.

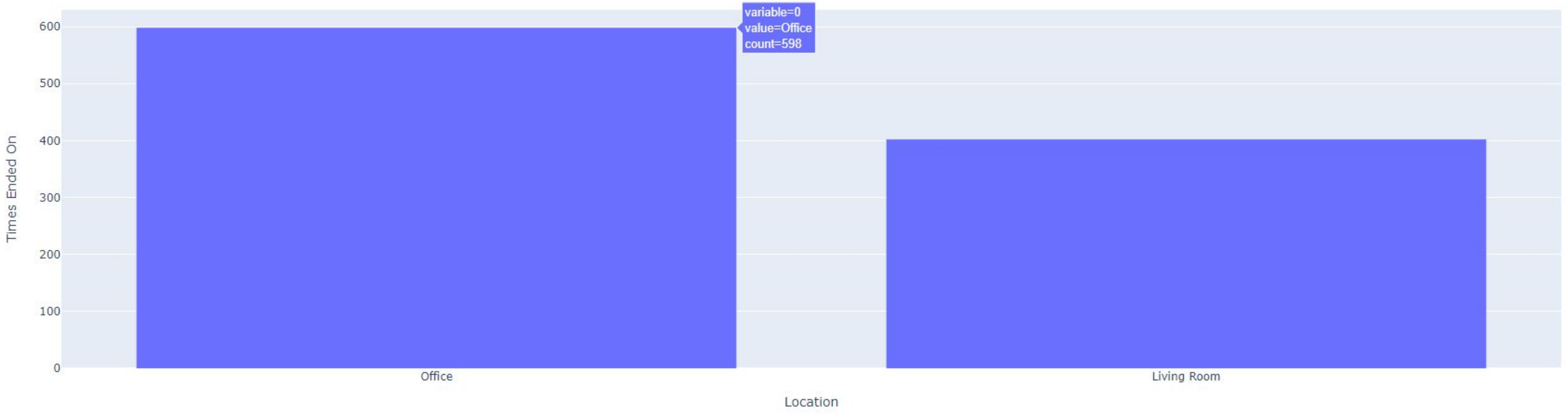
The resulting histogram is the one below. Because we are finding the initial state, the locations we get in the histogram depend solely on our prior probabilities from get_prior(). Our prior probability defines an equal chance of starting in any location. This is why we see a decently even distribution.



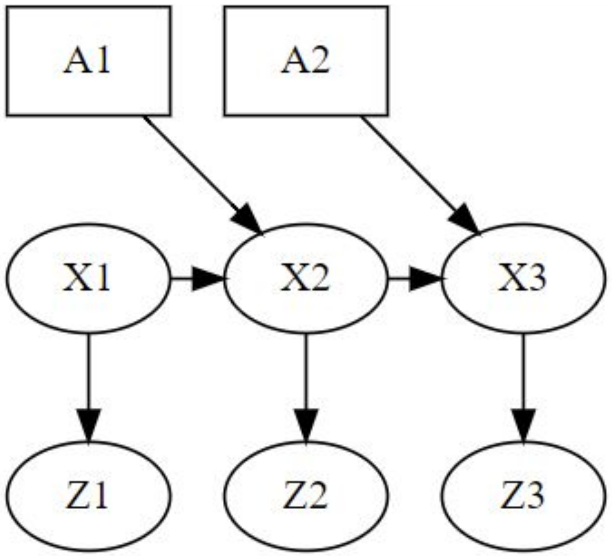
3.3 Dynamic Bayes Nets [4 pts]

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 resulting histogram is the one below. The final state in all 1000 trials was either the Office or the Living Room. This is because our custom action sequence does all LEFT actions up until the last action which is UP. The only two rooms that can be reached using only LEFT actions are the Office or the Living Room. Because you cannot go UP in either of those rooms you will stay in either the Office or the Living Room. This is why we see this distribution.

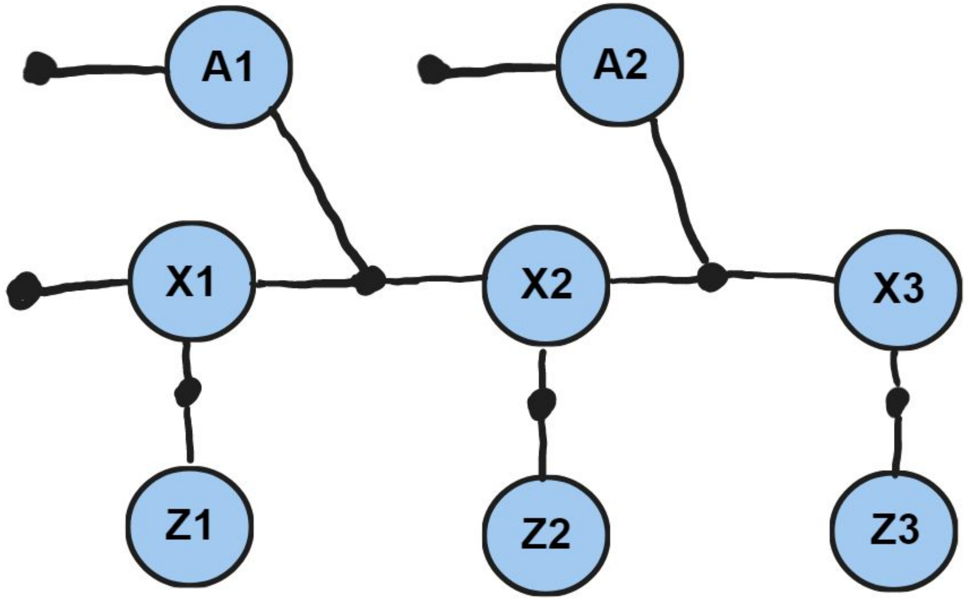


4.1.A Perception w/ Graphical Models [2 pts]



Given the DBN from the textbook,

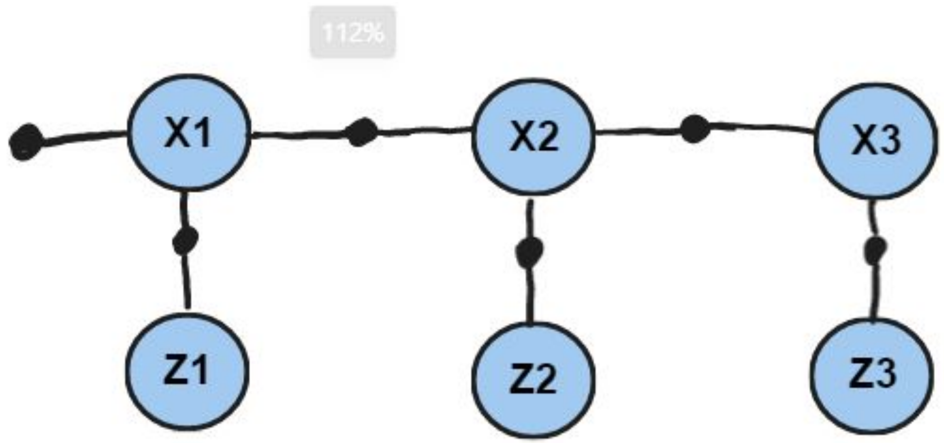
- If all actions and measurements are unknown, what does the factor graph look like?

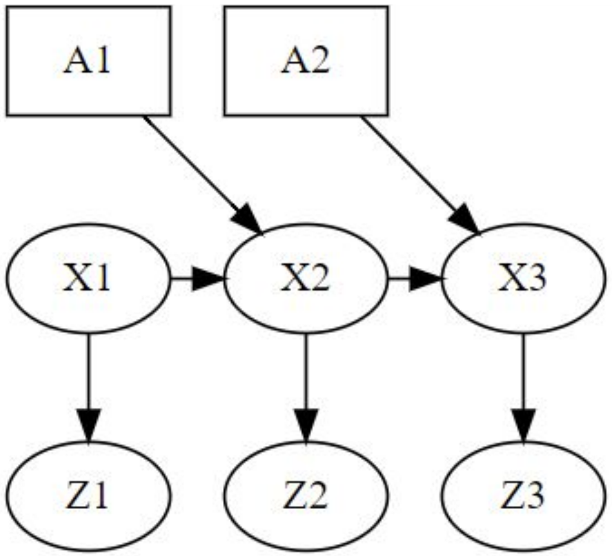


4.1.B Perception w/ Graphical Models [2 pts]

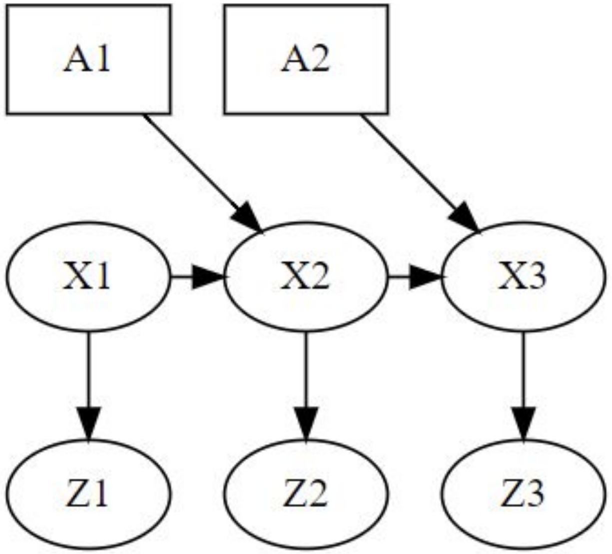
Given the DBN from the textbook,

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





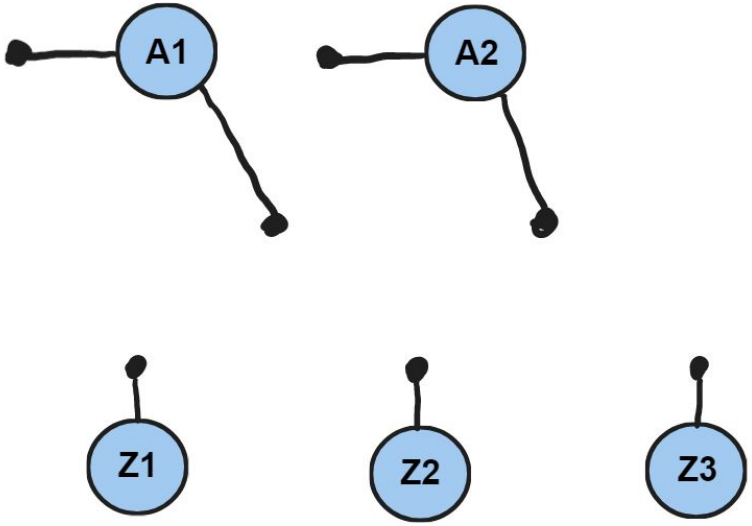
4.1.C Perception w/ Graphical Models [2 pts]



Given the DBN from the textbook,

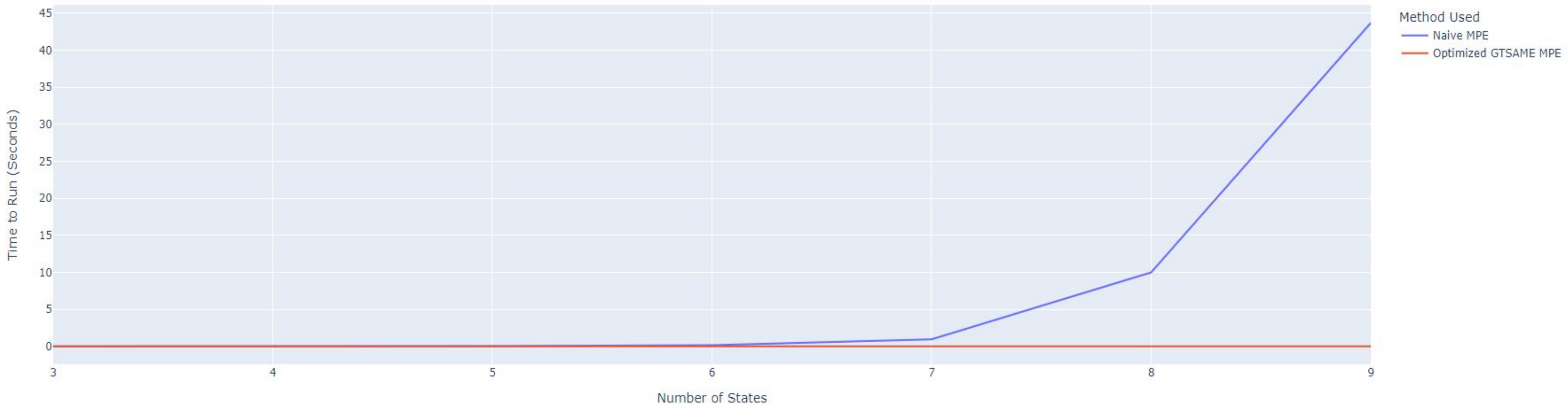
- If the actions and the measurements are unknown, but the trajectory of the states is known, what does the factor graph look like? Do you observe any direct findings from the factor graph?

From the factor graph, we can say there is no relationship between actions and measurements (they are independent).



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)

- Linear(ax+c)
- Cubic(ax^3+c)
- Quintic(ax^5+c)
- Exponential(e^x+c)

5.1 Markov Decision Process [2 pts]

What is a rollout? What is rollout reward?

A rollout is a discounted reward for a given sequence of actions. It produces a sample trajectory and a corresponding discounder reward. Given some reward function, a rollout reward is the reward we expect to receive from a rollout.

5.2 Markov Decision Process [3 pts]

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

The two control tapes only differ in the 3rd action where the first is RIGHT and the second is UP. The optimal control tape was the second control tape. It is optimal because it produced an average reward of 13.0 over 10 trials compared to control tape one which averages 4.0 over 10 trials.

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: LEFT LEFT RIGHT UP UP. This policy gives us the best chance to end up in the Living Room no matter what room we start in (Living Room is our only chance for reward).

5.4 Markov Decision Process [3 pts]

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

In policy iteration, we start with a random policy, find a value function for that policy, and find a new improved policy based upon the value function. We go until we converge on an optimal policy, or an optimal mapping of states to actions. In value iteration we start with some value function and continually improve it based upon the known rewards in the environment until the changes we make in our value function become negligible, thus becoming optimal.

6.1 *EXTRA CREDIT* RL [2 pts]

In the given equation, please define the variables x, a, x', γ, V*, P, and R̄.

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

x: current state
a: action taken
x’: next state
gamma: discount factor to prevent infinite time horizons
V*: optimal value function
P: probability - in this case, conditional probability of next state given current state and an action
R: Reward function returning the reward for a given state and action

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've take CS3600 Intro to AI previously and that class was extremely theory-based and had few in-depth examples of MDPs, Q-learning, etc. I think the structure this class takes makes understanding the difficult material somewhat easier, since I usually learn from working out examples and coding instead of looking at formulas. What I think would make the project more clear is more worked out examples in the form of diagrams. Like for Reinforcement Learning, rollouts, etc., it would be nice to have a written out example to show what is happening. Additionally, the gtsam 101 sections are helpful but I would prefer even more resources on it because it can be slightly confusing.