CS 238 Final Project Update

Matt Foutter Michael Liu Zach Rotzal

The Problem:

The goal of this project is to develop an autonomous self-driving system for a school bus for the Stanford campus. The bus will be able to transport students and staff between different buildings and locations on campus, providing a safe and efficient mode of transportation. This project will focus on the sequential decision-making capabilities of the self-driving bus, as well as addressing potential sources of uncertainty.

Progress:

We have scoped the problem so that the bus drives on a one lane road and is the only vehicle on the road. The goal of the bus is to reach the end of the road as quickly as possible without hitting any pedestrians.

So far, we've developed a value iteration solution in a 1D grid world with deterministic pedestrian actions. We have defined the environment, agents, state space, actions, and reward. Our algorithm currently models the position of the pedestrian with the reward function, which is a feature we'd like to remove after speaking with Marc during OH. Currently, if the pedestrian will be on the road in the next time step, which we know with 100% certainty, then the reward at the preceding road state and forward action is -100. The bus then uses value iteration to solve for the utility in each state and picks the highest utility action from the current state. We've attached plots of the reward function for forward action when the pedestrian is about to cross, and a trajectory plot of the bus state along the road with simulation iteration. Here, the bus yields to the pedestrian and then continues forward (see appendix).

After speaking with Marc, we'd like to generalize our problem to an exploration strategy. Likely, we'll use epsilon-greedy exploration. We'd like to demonstrate that before our exploration phase ends, we are able to draw enough samples of the environment to accurately capture the value of state action pairs in a tabular Q function which we can exploit thereafter to navigate the environment.

Soon we will have our Q-learning with an epsilon greedy approach working. After this initial algorithm is implemented, possible extensions of our project include incorporating stochastic processes, moving into a 2-D space, and incorporating a two lane road with other cars. We have explored the CARLO simulation for extending our state space to a 2-D world and action space.

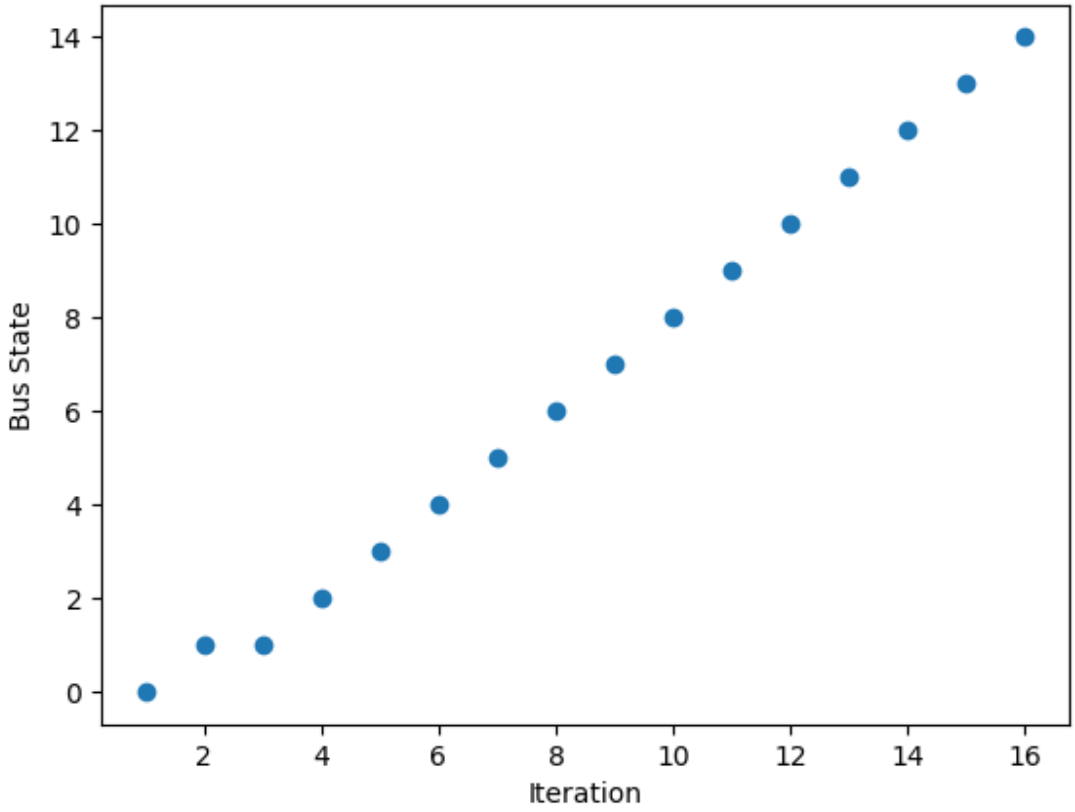
Timeline:

Our goals for future work is as follows:

- 3/6 - define exploration logic for world 1D
- 3/8 - introduce random pedestrian action
- 3/10 - scale project to grid world 2D
- 3/17 - video/paper

Appendix

Bus Trajectory



Reward Function at Pedestrian Crossing