ORIE 6750 Final Project

Yucheng Chen

**Project Problem (comes from Project idea 1):**

A collection of individuals and their social connections are represented as a graph, where each node represents an individual and the edges in the graph represent the social connections among them.

Each node is associated with a binary random variable , representing if the individual is infected at time . An individual with known prior probability in the graph is infected at time 0. An infected individual infects his or her neighbors with probability and each infected induvial has a probability of reporting to the public health department. Also, the public health department has test kit at each time step and can determine if individuals are positive. All infected people (self-reported and tested) are removed from the graph. The objective is to minimize the infected people.

**Formulation as a POMDP:**

Suppose there are nodes in the graph, then

***Time horizon***:

***Belief state***:

where is probability that the individual is infected.

***Physical state***: is a binary matrix where  indicates there is a social connection between individual at time step and indicates otherwise.

***Reward***:

where

The reward is zero in intermediate steps since the objective is to minimize the infected people at the end. Thus, in the final step the reward is the number of un-infected people. Some of them may connect to infectious individuals but consider this is the find step (maybe because we run out of the test kits), we can remove the connections between the infectious and un-infected people at this final step.

***Observation***:

test kits are applied to the individuals and find out whether they are positive. Denote the set of all individuals as and as the set of individuals that are tested at time step . Note that the items in should be distinct. Then the observation vector , is a **binary** vector of size and represents the test results (positive or negative) for the tested individuals.

***Transition Dynamics***:

-Update belief state according to the graph connection:

where .

The meaning of the dynamics is as follows:

The probability of the individual being infected at comprised of two terms: 1). The probability of being infected at ; 2). The probability of not being infected at but infected by his or her neighbors at . The term represents the probability of being infected by ’s neighbors, where each individual is ’s neighbor in the graph.

-Update the belief state according to the observation :

-Update Physical state :

For ,

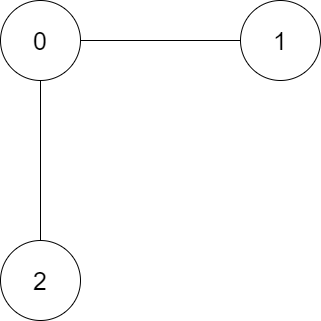
**Bellman Recursion:**

where

This Bellman recursion is a little bit different than the classical POMDP Bellman recursion since the transition dynamics of belief state in this problem is stochastic. Therefore, the Bellman recursion of this problem is a double integral: 1) the inner integral over the distribution of the belief state , which doesn’t appear in the classical POMDP Bellman recursion; 2) the out integral over the distribution of observation .

**Optimal policy:**

***Case study 1:***



The first case study is for the graph shown as above: there are three people in the group and person #0 has connections with both person #1 and person #2. Set the number of test kits for each step as and let the initial belief state as .

Then plot of the value of this initial belief state versus the horizon remaining with different set of probability of infecting neighbors and the probability of self-reporting is as follows:

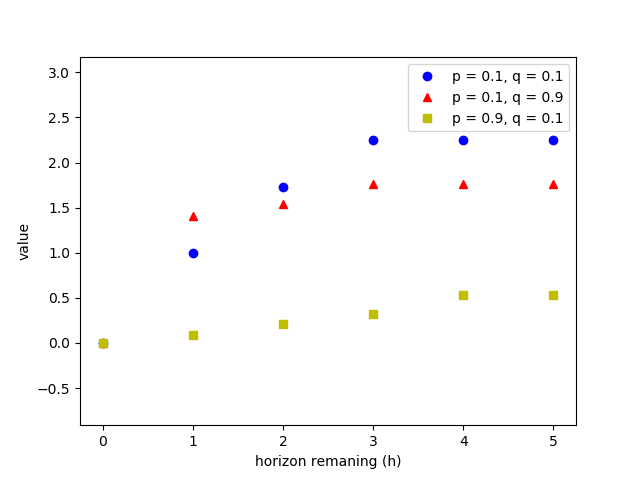
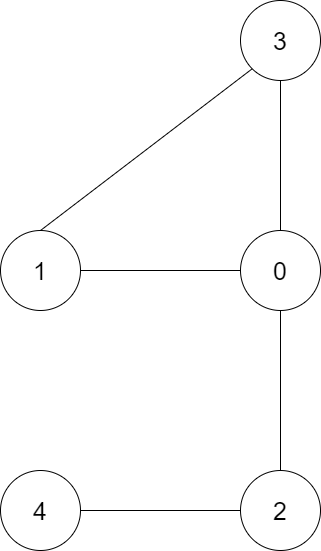


Figure 1. Value of versus horizon remaining

As shown in Figure 1, if no planning horizon remains, the value is zero as no person is determined as non-infected based on . As the planning horizon increases, the value increases accordingly. We can observe that if both and are small, the value is large because the probability of infecting neighbors if low and people are reluctant to self-report of being infected. However, if the probability of self-reporting increases, the value can reduce as the planning horizon increases because the people with low probability of infecting of virus has high probability of self-reporting. As the probability of infecting neighbors increases, the value decrease because people are more likely of being infected. The running time for case 1 is ~10 seconds

***Case study 2:***



The second case study is for the graph shown as above: there are five people in the group. Set the number of test kits for each step as and let the initial belief state as .

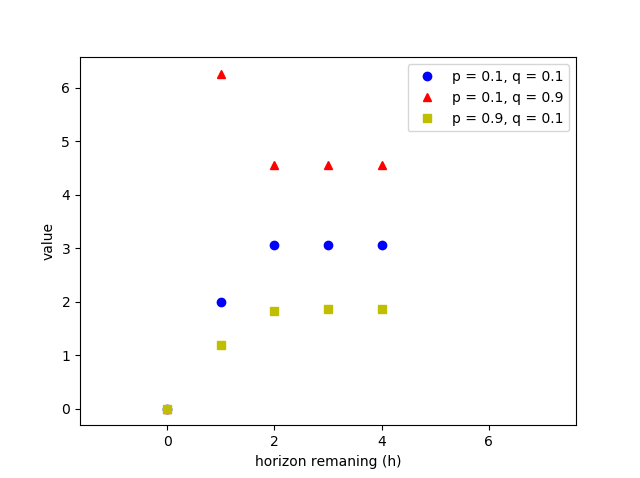


Figure 2. Value of versus horizon remaining

As shown in Figure 2, if no planning horizon remains, the value is zero as no person is determined as non-infected based on . As the planning horizon increases, the value increases accordingly except for , where value of is greater than the value of any . The value of small and large has the smallest value because the people are less likely to be infected and has lower probability of self-reporting. The running time for case 2 is ~2447.39 seconds.

**Heuristic policy:**

As we can observe from the previous section that the optimal policy suffers from the curse of dimensionality. Thus, two heuristic policies are developed for the problem:

Given test kits for each time step,

1. Greedy policy:

Test the first individuals whose posteriors of being infected are highest.

1. One step lookahead policy

Evolve the belief state for one step ahead (assuming no self-reporting) and test the first individuals whose posteriors of being infected are highest.

The graph in case 2 is used for the comparison between the heuristic policies and the optimal policy. The parameters used are: . As shown in Figure 3, the optimal policy is better than both heuristic policies and heuristic policy 2 is better than heuristic policy 1.

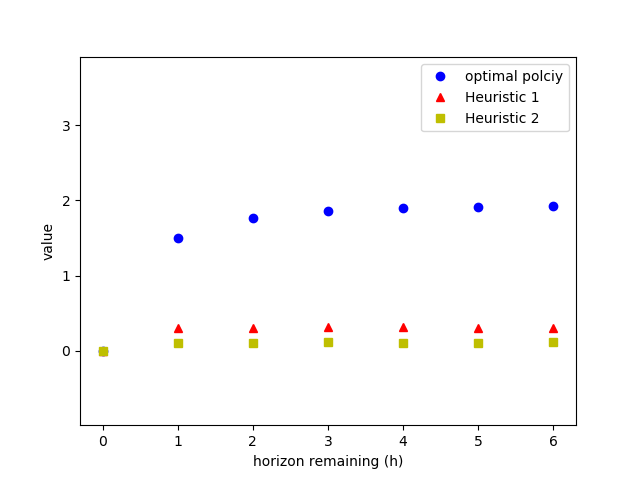


Figure 3. The comparison between Opt Policy and Heuristic Policy