Winning a Party Game: Optimal Policy in White Elephant Gift Exchange

Junye Li
Mathematics and Computer Science
Mentor: Matt Kretchmar

Introduction

White Elephant Gift Exchange (WEGE) is a party game that people will play during Christmas Holiday. The goal of our project is to find the best strategy for each player and use the strategy to determine the average gift value that each player can get.

Rules of WEGE

- White Elephant Gift Exchange widely plays in the West by different rules. Normally, there are n players playing this game and each player will bring a wrapped gift. Each player can recognize their own gift.

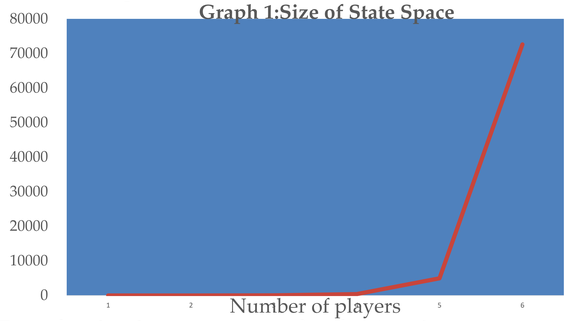
1) There are n people playing game and everyone will bring a wrapped gift.
2) There are n rounds. In each round, one specified player will go first. The player generally will have two choice: open a wrapped gift or steal from other player.
3) In order to avoid infinite cycle, each gift can only be stolen once in each round.

Basic Definition

1.State Space

- Definition: The state represents the status of each player in the game. Namely, the state gives the status of which player has which gift in the game.
- Size of State Space

Graph 1:Size of State Space



- From the plot above, we can see that the size of state space increase almost exponentially.

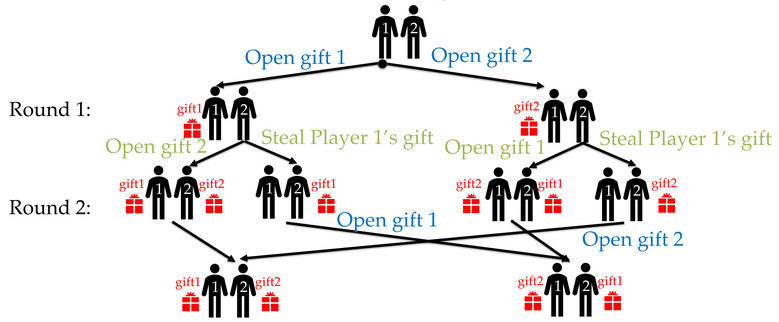
2. Action Space

- Definition: Action is the transition that happens between two states. In general, each player will face two kinds of action. The player can either open a new gift or steal another player's gift.

3.Trajectory

- Definition: sequence of states and actions that moves the game from the start state to end state.

Tree Diagram



- The plot above shows all possible situations will happen if two players are playing.
- At the beginning of the game, each player has no gift.
- In the Round 1, Player 1 open a random gift, which means he/she has 50 percent chance to get gift 1 and 50 percent chance to get gift 2. Then, Round 1 ends.
- In the Round 2, Player 2 needs to take an action. He/she can steal Player 1's gift or open a random gift. If Player 2 steals gift from Player 1, then Player 1 has no gift. Due to the rule that any gifts cannot be stolen back in the same round, Player 1 needs to open a random gift.
- From the plot, we can see that there are seven different states. At the end of the game, it will have two possible outcome. Player 1 gets gift 1 and Player 2 gets gift 2. Or Player 1 gets gift 2 and Player 2 gets gift 1.

Analysis

Dynamic Programming is the main approach that we will use to find the best strategy. It's very clear that which player win how much gift value in the end of the game. Then, Dynamic Programming helps us calculate the value of all states by updating the value backwards. If the player choose to open a gift, then we will calculate the expected value of all next state and use that value to update the value of current state. If the player has two kinds of action, then the value of current state is the maximum value between two next states.

Dynamic Programing can help us get the accurate value of each player. However, it has a very obvious drawback is that it will sweep all the state. From the Graph 1, we can see that the number of states increase very fast. With the number of players increase, it will take longer time for us to calculate the value. Also, if the number of players exceeds 9 players, it will become impossible for us to get the optimal strategy.

Future Work

1. Reinforcement Learning

Since Dynamic Programming cannot give us the accurate value for each player, we need to find an alternative way to get the average gift value and optimal strategy. Reinforcement Learning will simulate n players play the game. Then, it will learn from the experience. It will update the value of each state as it takes action. Then, if next time it goes to the same state, the function will give the action that has maximize value in the future.

Reinforcement Learning is a good way for us to get the value of states, but the value now we get still not as accurate as the value Dynamic Programming does. Therefore, we needs to use function approximation to get more precise value in the future.

2.Pattern of value

We try to find the pattern of the value for the beginning state. It is obvious that the last person always get the highest value. We think there may exist a pattern that can help us now the accurate value for each player, when Dynamic Programming does not work. However, we have not found now.

Acknowledgements

Who helped you in important ways? This section should be short and should include…

- Source of your summer funding (who paid the bills?)
- Those who contributed important data, feedback, taught you a skill that you needed, etc. Refer to people by their full name.

Works Cited

Please follow your discipline's rules for citing the literature. Give credit to others for their work in order to avoid plagiarism & maintain academic integrity. This section could also be labeled Literature Cited or Sources Cited.