* 1. Title: Using single Q funtion to against multiple reward function in deceptive reinforcement learning
  2. Abstract:

In this paper, we develop a method to deal with multiple reward function by a single Q function. Hiding reward function is one of key methods to slove privacy-preserving problem in deceptive reinforcement learning. Our method hides the detail of multiple reward functions to protect information of reward function better and reach better deception by intergrating global information.

* 1. Introduction:
     1. Background of the study

Reinforcement learning is an area of machine learning. By Maximize the notion of reward, it returns an intelligence result. One of effect methods is using reward function to compute Q-Value for each step and generates a Q-Table for the entire issue. As reinforcement learning has been widely used, the protection of privacy becomes a key problem. To solve privacy problem, deception reinforcement learning is used to generate a fake result to cheat the observer. The performance of deceptive reinforcement learning is measured by it's magnitude, density, and extent.

* 1. Problem Statement

One of the keys to deception reinforcement learning is to protect privacy, especially the relevant information of reward function is not leaked. However, in the deceptive reinforcement learning problem with multiple reward functions, there are usually multiple Q functions that related to reward function used to generates Q-Table, which means the computation of reward function are exposed to agents who are responsible for calculating Q-Table. Therefore, there is a great risk of leaking information about reward function.

* 1. Research Objective

To hide the information of reward function, we design single Q-function which integrated all reward functions to compute Q-Value. In this way, only one agent who is dealing with this single Q-function knows the truth. Therefore, the privacy of the whole system can be protected in a safer way.

* 1. Significance of the study

Both deception reinforcement learning and multiple rewards integration are popular topics in machine learning area. However, there still is not an comprehensive work. Therefore, our innovative research work on this topic will make up for the vacancy in deception reinforcement learning.

* 1. Literature Review

Deception is a traditional topic and it's general theory and knowledge structure is integrated (Barton Whalty, 1982). The deception is fell into two categories: dissimulation and simulation, and some methods to deceit are raised. Currently, the application on reinforcement learning also acquires remarkable achievement (Yue Yang, Zhengshang Liu, Peta Master and Tim Miller, ). The ambiguity model and irrational model are designed based on the Q-value that acquired from reinforcement learning. More specific application of deception reinforcement learning on path-planning gives the process of designing a deceptive path (Peta Masters and Sebastian Sardina). One of key method in deceptive path-planning is maximising the last deceptive point. Also, the reinforcement learning with multiple shared rewards is a hot topic. Douglas promotes an integrated interaction model which deals with multiple reward function and goals properly. (Gouglas M. Guisi, 2016)

* 1. Research Methodology
     1. Research Methodology

There are two important research methodologies in our project. One of the methods is Mathematical Proof which is used to prove the rationality of the algorithm. Another one is Experiments by coding and testing to verify the correction of our ideas in reality.

* 1. Justification

Reinforcement learning is based on the result of a series calculations. To acquire a correct and rational Q-Value after covering multiple reward functions in a single Q function, we must use mathematical methods to get a good result at first. Currently, some works provide the ways to handle multiple reward functions. Our aiming is based on these works to find an effective method that performs on deception reinforcement learning well. On the other hand, we plan to conduct our experiment on P4-simulator. Currently we have reproduced two classical deception reinforcement learning mode: ambiguity model and irrational model. These models provide deceptive path planning according to Q table. Therefore, verifying the performance of the integrated Q function on these models and adjusting based on these to make the model show better deception are reasonable.

* 1. Plan

From 3.2-3.9, we research the basic knowledge of deception reinforcement learning

From 3.10-3.24, we reproduce the models that show in popular paper and decided the specific research topic.

From 3.25-4.8, we start to implement that using a single updating function against multiple reward functions.

From 4.9-4.23, we optimize the performance on various situations and test the result.

From 4.23, we organize the results and write report.

* 1. Expected Outcomes

Our research will make up the gap between reinforcement theory and it's application on deception model. After this method is implemented, the bottom layer of the database does not need to provide complete layout information to the superior, but provides part of the processed information as required. Because our proposed method hides more information from superior agents, this is a more secure and confidential reinforcement learning method.

* 1. Conclusion

We promote a method against multiple reward functions by a single Q function in deception reinforcement learning. This method hides more information to observer to protect privacy in a safer way.