ARC notes: Minimax Visibility **Zhongshun Zhong**
Date: October 31, 2018

Problem Statement

Given a known environment(graph, grid-based, etc), the initial position of the agent and the guard. For a given time length 0 to T, design for a path for the agent to maximize the total visibility region and minimize the detectability caused by the guard.

Problem Assumption

We study a path planning algorithm for the agent, it takes place in a 2-D graph or a visibility polygon based environment. We make the following assumptions:

- The time T is discretized, the agent and the guard move in turn of a unit time step(such as 1s).
- The agent knows the initial distribution probability of the guard.
- The agent and the guard have known to sense ranges. we may assume that both sensing ranges are unlimited for ease of illustration,
- The guard will move randomly before detecting the agent. When the agent was detected, the guard will start to catch the agent.
- The guard can move faster than the guard. (Optional)
- The game ends when the agent was caught by the guard.

Problem Formulation

The agent's objective can be written as:

$$\max_{\pi_a(t)} \min_{\pi_g(t)} E\left[R(\pi_a(t)) - \eta_d(\pi_a(t), \pi_g(t))P_d - \eta_c(\pi_a(t), \pi_g(t))P_c\right]. \tag{1}$$

where,

- $\pi_a(t)$ denotes an agent's path.
- $\pi_g(t)$ denotes a guard's path.
- $R(\pi_a(t))$ denotes the positive reward collected by the agent.
- P_d is the penalty reward whenever it is detected by the guard.
- P_c is the penalty reward whenever it is caught by the guard.
- $\eta_d(\pi_a(t), \pi_g(t), \eta_c(\pi_a(t), \pi_g(t))P_c$ are either 0 or 1 to indicate detect and catch of the agent.

Since the agent don't know the exact position of the guard initially, we convert the problems into two sub-problems: 1) Path planning with estimation the position of the guard (before detected), 2) Path planning with an adversarial guard(after detected).

0.1 Stage 1: Before detected

At stage one, the agent does not know the position of the guard since the guard is invisible. The objective function we consider in stage 1 is,

$$\max_{\pi_a^1(t)} E\left[R(\pi_a^1(t)) - \eta_d(\pi_a^1(t), \pi_g(t))P_d\right]. \tag{2}$$

With known initial position of the guard, the probability of its moving direction, the goal of stage 1 is to design an path $\pi_a(t)$ off-line.

The agent will execute $\pi_a^1(t)$ at this stage until the agent was detected. Then the problem will move to stage 2.

0.2 Stage 2: After detected

When the agent was being detected by the guard, then problem becomes a standard pursuit-evasion game. The guard will try to catch the agent, the agent needs to avoid being caught and also explores more undetected area. The objective function in stage 2 is,

$$\max_{\pi_a^2(t)} \min_{\pi_g(t)} \left\{R(\pi_a^2(t)) - \eta(\pi_a^2(t), \pi_g(t))P_c\right\}. \tag{3}$$