Treasure Keeper: Cooperation and Competition in a Multi-Agent Game

Gonçalo Lopes
goncalo.chincho.lopes@ist.utl.pt

Maxwell Junior
maxwell.junior@tecnico.ulisboa.pt

1. PROBLEM DEFINITION

Games enable the simulation of behaviors in hypothetical situations that can't be set in the real world, with which humans are capable of interacting with.

We propose a game called "Treasure Keeper", where a group of cooperative agents with the same properties, the "Treasure Hunters", compete against a single agent with different properties and goals, the "Treasure Keeper". Both teams are placed on a square grid world environment with treasure chests and jail cells, where the hunters must collect as much treasure as possible before the keeper locks every hunter in the jail cells.

The rules are the following:

- The keeper knows the location of every chest.
- The keeper can grab a hunter if they are in adjacent cells.
- The keeper can lock a hunter in a jail cell if it is carrying a hunter and standing next to a jail cell.
- A hunter dies the second time they get locked.
- A free hunter can free locked hunters, but not dead ones.
- The amount of treasure retrieved from a chest at each tick by a hunter is inversely proportional to the amount of hunters retrieving gold from that same chest.
- Treasure chests award a significant amount of gold when fully searched.

2. MULTI-AGENT SYSTEM

We propose a solution based in a multi-agent system, where the treasure hunter agents share information among themselves, as cooperation may lead to strategies that increase the chances of success in achieving the common goal they share, and on the other hand, the treasure keeper must devise strategies to keep the hunters away from the treasures.

Given the stochastic aspect of each run at the beginning of the game during the search for treasure by the hunters and search for hunters by the keeper, the agents must devise a strategy for each different instance of the game, making it possible to study how these adapt to different states, and which strategies they employ to maximize the following utility function u for a run r of the hunters:

$$u(r) = \frac{obtained\ treasure}{total\ treasure} \tag{1}$$

Or minimize (1) in the case of the keeper.

The agents deliberative capabilities will be trained through reinforcement learning, by playing the keeper against the hunters, iteratively optimizing the behavior of both types of agents. This learned behavior can then be studied by changing different parameters of the environment and the agents, such as the number of agents, gold collection speed or the keeper's walking speed for example.

3. PROPERTIES

The agents' properties are the following:

- Adaptive - the agents learn how to interact with the environment and with other agents.
- Rational - the agents must, in the case of the hunters, maximize the amount of treasure collected, and in the case of the keeper, to minimize this amount.
- Proactive - agents have initiative to fulfill their goals.
- Cooperative - among hunters, which have common goals.
- Competitive - between the group of hunters and the keeper.
- Mobile - both hunters and the keeper move around the grid.

The environment's properties are the following:

- Partially Accessible - only some aspects are available to the agents, such as the treasure chests' locations.
- Non-Deterministic - the environment is asynchronous, thus leading to unpredictability.
- Dynamic - while one agent is deliberating, another one can act.
- Discrete - finite number of percepts and actions.
- Non-Episodic - previous actions influence future ones.