Abstract

Introduction to the Model

Stochastic Games

Model

Computation and Linear Program

Design Problem and Heuristics

Experiments

Q-Learning Competition

Experiments

Partial Observability

Experiments

**Partially Observable Introduction**

**Partially Observable**

Next, we will introduce a partially observable variant of the GTGR scenario. In partially observable scenarios, the rules of the game remain unchanged, except for addition of “shadow nodes,” special nodes which hide the position of the adversary, until the adversary enters an observable portion of the graph.

In Fig. X, the agent starts the game on node S. The 4 nodes in grey are shadow nodes. When the adversary moves to node 4 ,5, 6 or 7, the observer is unable to determine their position until the adversary re-enters a visible portion of the graph.

**Solutions to The Partially Observable Model**

We will examine two solutions to the partially observable model, both of which involving linear programming.

**Fat Solution**

The first solution, referred to as the “fat” solution, is the easiest to conceptualize. The fat solution groups connected shadow nodes into “fat” nodes, so that the observer can treat non-observable sections of the graph as single states.

Fig. X2

In the example illustrated in Fig. X2, when the adversary moves to nodes 4, 5, or 7 (the three entrance points to the hidden portion of the graph) the observer will only know that the adversary has entered the state *Z1*. From the observer’s point of view, the adversary will remain in *Z1* until the adversary moves to nodes, 3, 8, or S, (the three exit points from the hidden portion of the graph). When the adversary enters *Z1,* the observer knows what states the adversary could possibly occupy without directly observing them. With the fat solution, the observer takes the same action for every turn the adversary spends in a hidden portion of the graph. We can easily modify the linear program to accommodate the fat solution.

//linear program and follow up

**Graph Transformation**

By making a number of assumptions about the adversary’s strategy, can generate a fully observable graph to simulate a graph that is partially observable.

\!\!\!\!\max\_{V, \{f\_i(s)\}\_{i,s}} \sum\_{\theta}&P(\theta)V(\theta, s\_{o}) \**label{eq:game\_obj}**\\

V(\theta, s) &\leq \sum\_{i \in B} r(s, i, j, \theta)f\_{i}(s) + V(\theta, j) &\forall\theta\in B,\forall s \mid s\neq \theta, \forall j\in\nu(s)\**label{eq:game\_incentive}**\\

V(\theta, s) &= 0 \quad &\mbox{when} \,\,\, s=\theta\**label{eq:game\_end}**\\

\sum\_{i} f\_{i}(s) &= 1\quad &\forall s \**label{eq:game\_sum}**\\

f\_{i}(s) &\geq 0\quad &\forall s,i\**label{eq:game\_nonneg}**