

1 Partially Observable Environments

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

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 observers 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.

$$\max_{V, \{f_i(s)\}_{i,s}} \sum_{\theta} P(\theta) V(\theta, s_o) \quad (2)$$

$$V(\theta, s) \leq \sum_{i \in B} r(s, i, j, \theta) f_i(s) + V(\theta, j) \forall \theta \in B, \forall s \mid s \neq \theta, \forall j \in \nu(s) \quad (3)$$

$$V(\theta, s) = 0 \quad \text{when } s = \theta \quad (4)$$

$$\sum_i f_i(s) = 1 \quad \forall s \quad (3)$$

$$f_i(s) \geq 0 \quad \forall s, i \quad (1)$$