Mean-payoff Games

Barbara Jobstmann
CNRS/Verimag (Grenoble, France)

Bucharest, May 2010

Definition

- Game graph for 2 players (S, S_0, E)
- Reward/weight function $r : E \to [-W, \ldots, 0, \ldots, W]$
- Player-0 value of a play $\rho = s_0 s_1 \ldots$ starting in state s_0 is

$$\mathrm{MP}_0(\rho) = \liminf_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} r(s_{i-1}, s_i)$$

- Player-1 value of a play $\rho = s_0 s_1 \ldots$ starting in state s_0 is

$$\mathrm{MP}_1(\rho) = \limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} r(s_{i-1}, s_i)$$

- Aim of Player 0: maximize $\mathrm{MP}_0(\rho)$
- Aim of Player 1: minimize $\mathrm{MP}_1(\rho)$
- Introduced by Ehrenfeucht and Mycielski

Value of a state and optimal strategies

- Given a game G, we denote by Π_i the set of all possible strategies of Player i.
- The Player-0 value of a state s under strategies $\pi_0 \in \Pi_0$ and $\pi_1 \in \Pi_1$, denoted by $\mathcal{V}_0(s, \pi_0, \pi_1)$, is the mean-payoff value of the play starting in s that is compatible with π_0 and π_1.

$$\mathcal{V}_0(s, \pi_0, \pi_1) := \mathrm{MP}_0(G_{s,\pi_0,\pi_1})$$

- The Player-0 value of a state s under the strategy $\pi_0 \in \Pi_0$, denoted by $\mathcal{V}_0(s, \pi_0)$, is $\mathcal{V}_0(s, \pi_0) := \inf_{\pi_1 \in \Pi_1} \mathcal{V}_0(s, \pi_0, \pi_1)$. (Player 0 want so ensure the value independent of Player 1.)
- A strategy $\pi_0 \in \Pi_0$ is optimal for Player 0 in a state s if $\mathcal{V}_0(s, \pi_0)$ is maximal, i.e., $\forall \pi_0' \in \Pi_0 : \mathcal{V}_0(s, \pi_0') \leq \mathcal{V}_0(s, \pi_0)$.
- Player-1 value and optimal strategies are defined analogously.

Determinacy and positional optimal strategies

For all state s, there exists a value v_s such that there exists a positional Player-0 and a positional Player-1 strategy π_0 and π_1 that ensure

$$v_s \leq \mathcal{V}_0(s, \pi_0) \qquad \mathcal{V}_1(s, \pi_1) \leq v_s$$

v_s is called the value of state s.

Note that π_0 and π_1 are optimal strategies.

Determinacy and positional optimal strategies

We use a result from Gimbert and Zielonka, "Games where you can play optimally without any memory" [CONCUR 2005]

Theorem

Suppose that a value function $\mathcal{V}$ *is such that for each finite game graph* $G = (S, S_0, E)$ *controlled by one player, i.e. such that either* $S_0 = \emptyset$ *or* $S_1 = \emptyset$*, the player controlling all states of* G *has an optimal (uniform) positional strategy in the game. Then for all finite two-player game graph* G *both players have optimal positional strategies in the game* G

One-player games have positional strategies

Consider a game graph G with $S = S_0$ and an arbitrary state $s \in S$ Claim: the best Player 0 can do is go from s to a simple cycle C with maximal average reward r_{max} and stay in the cycle. The payoff Player-0 gets with this strategy it the average reward of the cycle C.

To proof: for all plays $\rho = s_0 s_1 \dots$ starting in s, $\mathcal{V}(\rho) \leq v_{max}$. Consider an arbitrary play $\rho = s_0 s_1 \dots$, first we decompose the play into its cycles as follows: we put the state on a stack and as soon as we revisit a state that is already on the stack, we have found a cycle and we remove the states from the stack. Let $C_0, C_1, \dots$ be the sequence of simple cycle generated like this and let $v_0 v_1 \dots$ be the average reward we obtain in these cycle.

One-player games have positional strategies

Then, we know that $\forall i \geq 0 : v_i \leq v_{max}$, since v_{max} is the maximal average reward we can obtain with a simple cycle.

This prove that $\mathcal{V}(\rho) \leq v_{max}$ for any arbitrary play ρ.

The proof for $S = S_1$ is similar but know we search for the simple cycle with minimal average.

(Note that these cycle can be find with Karp's shortest path algorithm in polynomial time.)

k-Step Game [Zwick and Paterson]

The two players play the game for exactly k steps constructing a path of length k, and the weight of this path is the outcome of the game.

The length of the game is known is advance to both players.

Let $v_k(s)$ be the value of this game started at a $s \in S$.

Theorem

The values $v_k(s)$ for all $s \in S$ can be computed in $O(k \cdot |E|)$ time.

Proof.

The results follows easily from the following recursive relation

$$v_k(s) = \begin{cases} \max_{(s,s')\in E}\{r(s,s') + v_{k-1}(s')\} & \text{if } s \in S_0 \\ \min_{(s,s')\in E}\{r(s,s') + v_{k-1}(s')\} & \text{if } s \in S_1 \end{cases}$$

along with the initial condition $v_0(s) = 0$ for all $s \in S$. □

Convergence of k-step game

Intuitively, $\lim_{k\to\infty} v_k(s)/k = v(s)$, where $v(s)$ is the value of the infinite game that starts at s.

Theorem

For every $s \in S$, we have

$$k \cdot v(s) - 2nW \leq v_k(s) \leq k \cdot v(s) + 2nW$$

Proof.

We use the fact that both players have positional optimal strategies. Let π_0 be a positional optimal strategy for player 0 start at s. We show that if Player 0 plays according to π_0 then the output of the k-step game is at least $(k-n) \cdot v(s) - nW$. $\square$

Consider play compatible with π_0. Push edges played by the players onto a stack. Whenever a cycle C is formed, the mean weight of C is at least $v(s)$ (since π_0 is an optimal strategy for player 0). The edges part of C lie at the top of the stack. They are removed and the process continues. At each stage the stack contains at most n edges with weight at least $-W$. Player 0 can therefore ensure that the total weight of the edges encountered in a k-step game starting from s is at least $(k-n) \cdot v(s) - nW$. This is at least $k \cdot v(s) - 2nW$ as $v(s) < W$. Similarly, if player 1 plays according to a positional optimal strategy π_1, she can make sure that the mean weight of each cycle closed is at most $v(s)$. At most n edges are left on the stack and the weight of each of them is at most W. She can therefore ensure that the total weight of the edges encountered in a k-step game starting at s is at most $(k-n) \cdot v(s) + nW \leq k \cdot v(s) + 2nW$.

Algorithm

Theorem

Let $G = (S, S_0, E)$ be a game graph with a reward function $w : E \to \{-W, \ldots, 0, \ldots, W\}$. The value $v(s)$ for every state $s \in S$ can be computed in $O(|S|^3 \cdot |E| \cdot W)$ time.

Proof.

Compute the values $v_k(s)$, for every $s \in S$, for $k = 2n^3W$. This can be done, according to the previous theorem, in $O(|S|^3 \cdot |E| \cdot W)$ time. For each state $s \in S$, compute the estimate $v'(s) = v_k(s)/k$:

$$\begin{aligned} v_k(s) - 2nW &\le k \cdot v(s) \le v_k(s) + 2nW \\ v'(s) - \frac{2nW}{k} &\le v(s) \le v'(s) + \frac{2nW}{k} \\ v'(s) - \frac{1}{n(n-2)} \le v'(s) - \frac{2nW}{k} &\le v(s) \le v'(s) + \frac{2nW}{k} \le v'(s) + \frac{1}{n(n-2)} \end{aligned}$$

Proof (cont.)

The value $v(s)$ is a rational number, with a denominator whose size is at most n. The minimum distance between two possible values of $v(s)$ is at least $\frac{2}{n(n-2)}$. The exact value of $v(s)$ is therefore the unique rational number with a denominator of size at most n that lies in the interval $[v'(s) - \frac{1}{n(n-2)}, v'(s) + \frac{1}{n(n-2)}]$.