A polynomial algorithm for the parity game

Bruno Scherrer*

April 6, 2022

Abstract

...

Parity and mean-payoff games For any integers $i \leq j$, write $[i, j]$ for the set of integers $\{i, i+1, \ldots, j\}$. A parity game between two players, Player 0 and Player 1, can be described by a tuple $\mathcal{G} = (X = [1, n] = X_0 \sqcup X_1,\ E = [1, m],\ p : X \to [1, d])$ with $(n, m, d) \in \mathbb{N}^3$. (X, E) is a directed graph. X is a set of n nodes and E a set of m directed edges such that each node has at least one successor node. The set of nodes X is partitioned into a set of nodes X_0 belonging to Player 0 and a set of nodes X_1 belonging to Player 1. The function $p : X \to [1, d]$, known as a priority function, assigns an integer label to each node of the graph. A play is an infinitely long trajectory $(x_0, x_1, \ldots)$ generated from some starting node x_0: at any time step t, the player to which the node x_t belongs chooses x_{t+1} among any of the outgoing edges of E starting from x_t. The winner of the game is decided from the infinte sequence of priorities $(p(x_0), p(x_1), \ldots)$ occuring through the play: if the highest priority occurring infinitely often is even (resp. odd), then Player 0 (resp. Player 1) wins.

A *mean-payoff game* between two players, Player 0 and Player 1, can be described by a tuple $\mathcal{G} = (X = [1, n] = X_0 \sqcup X_1,\ E = [1, m],\ g : X \to [-W, W])$ with $(n, m, W) \in \mathbb{N}^3$ similar to that of a parity game ; the only difference is that the priority function is replaced by a cost function $g : X \to [-W, W]$. The dynamics of the game is the same as above. On potential plays $(x_0, x_1, \ldots)$ induced by the players' choices, Player 1 wants to maximize $\liminf_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} g(x_i)$ while Player 0 wants to minimize $\limsup_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} g(x_i)$. Ehrenfeucht and Mycielski [1979] have shown that for each starting node x_0, such a game has a value $\overline{v}(x_0)$, the optimal mean-payoff from x_0, such Player 1 has a strategy to ensure that $\limsup_{t\to\infty} \frac{1}{t}\sum_{i=0}^{t} g(x_i) \geq \overline{v}(x_0)$ and Player 0 has a strategy to ensure that $\limsup_{t\to\infty} \frac{1}{t}\sum_{i=0}^{t} g(x_i) \leq \overline{v}(x_0)$.

For both games, it is known (cf. Zielonka [1998], Ehrenfeucht and Mycielski [1979]) that there exist optimal strategies that are positional (i.e. that are mapping from nodes to outgoing edges). In particular, when both players follow these positional strategies from some node x_0, the play follows a (potentially empty) path followed by an infinitely-repeated cycle $(x_1^*, \ldots, x_{c(x_0)}^*)$ for some $c(x_0) \in [1, n]$.

1 Mean payoff game with costs in $\{0, 1\}$

Before we consider parity games, we shall spend some time on the specific class of mean payoff games where the costs all belong to $\{0, 1\}$. Let $\overline{v} : X \to \mathbb{Q}$ be the value of such a mean-payoff game. For any starting state x, the value $\overline{v}(x)$ from node x necessarily has the form $\frac{a}{b}$ for some $0 \leq a \leq b \leq n$. In this context the aim of this section is to describe a computationnally efficient characterizaton of 1) the set of nodes x for which Player 1 can ensure that $\overline{v}(x) > 0$ and 2) the *set of policies* such that Player 0 can ensure that the mean-payoff is 0. To do this, we shall focus on the closely related total cost of the h-horizon payoff game, played on the same graph with the same cost function, but with a finite duration h. We begin by a bit of terminology for this setting.

*INRIA, Université de Lorraine, bruno.scherrer@inria.fr

Let N and M be the set of (positional) strategies for Player 0 and Player 1:

$$N = \{ \ \mu : X_0 \to X \ ; \ \forall x \in X_1, \ \mu(x) \in \{ \ y; (x,y) \in E\} \ \},$$
$$M = \{ \ \mu : X_1 \to X \ ; \ \forall x \in X_0, \ \mu(x) \in \{ \ y; (x,y) \in E\} \ \}.$$

For any pair of strategies $\mu \in M$ and $\nu \in N$, consider the operator $T_{\mu,\nu}$ that acts on functions $v : X \to \mathbb{N}$:

$$\begin{cases} \forall x \in X_0, \ [T_{\mu,\nu} v](x) = g(x) + v(\nu(x)), \\ \forall x \in X_1, \ [T_{\mu,\nu} v](x) = g(x) + v(\mu(x)). \end{cases}$$

For any sequence of strategies $(\mu_1, \dots, \mu_h) \in M^h$ of Player 1 and $(\nu_1, \dots, \nu_h) \in N^h$ of Player 0, and for any starting state x, the integer $[T_{\mu_1,\nu_1} \dots T_{\mu_h,\nu_h} 0](x)$ is the total payoff obtained on the h-long trajectory where Player 1 (resp. Player 0) uses μ_t (resp. ν_t) for picking transitions at time t (though the player never pick transitions simultaneously, the exact time steps where each actually picks a transition depends on x, and all the decision that have been made in the past).

For any pair of strategies $\mu \in M$ and $\nu \in N$, consider also the operators T_μ, $\tilde{T}_\nu$ and T that also act on functions $v : X \to \mathbb{N}$:

$$\begin{cases} \forall x \in X_0, \ [T_\mu v](x) = g(x) + \min_{y;(x,y)\in E} v(y), \\ \forall x \in X_1, \ [T_\mu v](x) = g(x) + v(\mu(x)), \end{cases}$$
$$\begin{cases} \forall x \in X_0, \ [\tilde{T}_\nu v](x) = g(x) + v(\nu(x)), \\ \forall x \in X_1, \ [\tilde{T}_\nu v](x) = g(x) + \max_{y;(x,y)\in E} v(y), \end{cases}$$
$$\begin{cases} \forall x \in X_0, \ [T v](x) = g(x) + \min_{y;(x,y)\in E} v(y), \\ \forall x \in X_1, \ [T v](x) = g(x) + \max_{y;(x,y)\in E} v(y). \end{cases}$$

Observe that, for any v, we have the following relations expressed in a somewhat compact form:

$$T_\mu v = \min_{\nu \in N} T_{\mu,\nu} v, \qquad \tilde{T}_\nu v = \max_{\mu \in M} T_{\mu,\nu} v, \qquad T v = \max_{\mu \in M} T_\mu v = \min_{\nu \in N} \tilde{T}_\nu = \max_{\mu \in M} \min_{\nu \in N} T_{\mu,\nu} v = \min_{\nu \in N} \max_{\mu \in M} T_{\mu,\nu} v.$$

where the max and min are taken componentwise.

Given a sequence of strategies $(\mu_1, \dots, \mu_h) \in M^h$ of Player 1, and any starting state x, the integer $[T_{\mu_1} \dots T_{\mu_h} 0](x)$ is the *minimal* total cost that Player 0 can induce if Player 0 uses μ_t to pick a transition at time t. Symmetrically, given a sequence of strategies $(\nu_1, \dots, \nu_h) \in N^h$ of Player 0, and any starting state x, the integer $[\tilde{T}_{\nu_1} \dots \tilde{T}_{\nu_h} 0](x)$ is the *maximal* total cost that Player 1 can induce if Player 1 uses ν_t to pick a transition at time t. Eventually, for any starting state x, $[T^h 0](x)$ is the value of the h-horizon payoff game and any pair of sequence of strategies $(\mu_1, \dots, \mu_h) \in M^h$ of Player 1 and $(\nu_1, \dots, \nu_h) \in N^h$ of Player 0 that satisfy $T^h 0 = T_{\mu_1,\nu_1} \dots T_{\mu_h,\nu_h} 0$ are optimal policies for the h-horizon problem.

Starting with $v_0 = 0$, consider the sequence of values obtained by iteratively applying the operator T:

$$\forall k \geq 0, \qquad v_{k+1} = T v_k.$$

As we have just said, $v_k = [T^k 0]$ corresponds to the solution of the k-horizon game. Let v_∞ be limit of this sequence:

$$v_\infty = \lim_{k \to \infty} v_k.$$

Note that $v_\infty(x)$ can be infinite for some x. Zwick and Paterson [1996] have shown that the mean payoff is equal to the limit of the average k-horizon cost when k tends to infinity:

$$\overline{v} = \lim_{k \to \infty} \frac{v_k}{k}.$$

It can be seen that $\overline{v}(x) > 0$ if and only if $v_\infty(x) = \infty$.

Let us make a few observations.

Lemma 1. *The sequence $(v_k)_{k\in\mathbb{N}}$ is non-decreasing.*

Proof. This follows from the facts that 1) T is a monotone operator in the sense that

$$\forall v, v',\ v \le v' \Rightarrow Tv \le Tv',$$

and 2) that $v_0 \le Tv_0$ since costs are non-negative. □

Lemma 2. *For any state x, and any $k \ge n$,*

$$v_k(x) = 0 \iff v_\infty(x) = 0.$$

Proof. The "$\Leftarrow$" implication is a consequence of Lemma 1.

To prove the "$\Rightarrow$" implication, we shall prove that for any starting state x_0 and any infinite sequence $(\mu_1, \mu_2, \dots) \in M^{\mathbb{N}}$ of decisions of Player 1. We shall prove that Player 0 can ensure that the total cost is 0. Let $(\nu_1, \dots, \nu_n) \in N^n$ be a sequence of policies of Player 0 such that:

$$v_k = T^n v_{k-n} = \tilde{T}_{\nu_1,\dots,\nu_n} v_{k-n}.$$

By assumption, we have

$$0 = [T_{\nu_1,\dots,\nu_n} v_{k-n}](x)$$

stack... cycle. check. □

Lemma 3. *For any state x, and any $k \ge n^2$,*

$$v_k(x) \ge n \iff v_\infty(x) = \infty.$$

Proof. □

Lemma 4. *Any positional policy $\nu \in N$ that ensures that the mean payoff from some state x is 0 necessarily reach the set in at most n steps.*

Proof. ...!!!... □

Consider the following partition of X:

$$\begin{aligned} A &= \{\ x\ ;\ v_N(x) = 0\ \}, \\ B &= \{\ x\ ;\ 0 < v_N(x) < n\ \}, \\ C &= \{\ x\ ;\ v_N(x) \ge n\ \}. \end{aligned}$$

Lemma 5. *The following properties hold:*

1. *For any state $x \in A$, there exists a strategy for Player 1 such that for all strategies of Player 0, plays from x in $\mathcal{G}'$ will only visit nodes with cost 0.*

2. *For any state $x \in B$, there exists a strategy for Player 1 such that for all strategies of Player 0, plays from x in $\mathcal{G}'$ will reach A in at most n steps.*

3. *For any state $x \in C$, there exists a strategy for Player 0 such that for all strategies of Player 1, plays from x in $\mathcal{G}'$ will visit nodes with cost 1 infinitely often.*

2 The optimal priority of a parity game

Puri [1996] has introduced the following reduction of any parity game $\mathcal{G} = (X,\ E,\ p)$ to a mean-payoff game $\mathcal{G}' = (X,\ E,\ w)$ that involves the exact same graph (X, E) and the cost function:

$$\forall x, \quad g(x) = (-n)^{p(x)}.$$

Indeed, consider an optimal cycle $(x_1^*, \ldots, x_c^*)$ in this mean-payoff game (using the above-mentionned positional strategies). Let $p = \max_{1\leq i\leq c} p(x_i^*)$ be the maximal priority obtained in this cycle. If p is even then

$$0 < n^p - (n-1)n^{p-1} \leq \sum_{i=1}^{c} (-n)^{p(x_i^*)}.$$

If p is odd, we similarly have:

$$0 > -n^p + (n-1)n^{p-1} \geq \sum_{i=1}^{c} (-n)^{p(x_i^*)}.$$

As a consequence, from any starting node x_0, Player 0 (resp. Player 1) wins the parity game if the value $v(x_0)$ of the mean-payoff game is positive (resp. negative). Furthermore, an optimal pair of strategies for the mean-payoff game is also optimal for the parity game (note that the opposite is in general not true).

We shall consider a slight variation of Puri's reduction that consists in choosing the alternative cost function:

$$\forall x, \quad g(x) = (-K)^{p(x)}$$

with any K such that $K - (n-1) > n^2$ (for instance one may take $K = (n+1)^2$).

Consider an optimal cycle $(x_1^*, \ldots, x_c^*)$ in this mean-payoff game. Let $p = \max_{1\leq i\leq c} p(x_i^*)$ be the maximal priority obtained in this cycle. If p is even then

$$n^2 K^{p-1} < (K-(n-1))K^{p-1} = K^p - (n-1)K^{p-1} \leq \sum_{i=1}^{c} (-K)^{p(x_i^*)} \leq nK^p,$$

and the value $v(x_0)$ from any node x_0 that reaches this cycle is such that

$$nK^{p-1} \leq \frac{n^2 K^{p-1}}{c} < v(x_0) \leq \frac{n}{c} K^p \leq nK^p.$$

When p is odd, we have

$$-n^2 K^{p-1} > -(K-(n-1))K^{p-1} = -K^p + (n-1)K^{p-1} \geq \sum_{i=1}^{c} (-K)^{p(x_i^*)} \geq -nK^p,$$

and the value $v(x_0)$ from any node x_0 that reaches this cycle is such that

$$-nK^{p-1} \geq \frac{-n^2 K^{p-1}}{c} > v(x_0) \geq -\frac{n}{c} K^p \geq -nK^p.$$

From any starting node x_0, we shall say that the *optimal priority* $p_*(x_0)$ is the value p such that $nK^{p-1} < v(x_0) \leq nK^p$ or $-nK^{p-1} > v(x_0) \geq -nK^p$ (by our choice of K this value is indeed unique). If this priority is even (resp. odd), Player 0 (resp. Player 1) wins the parity game from x_0.

Through this slightly modified reduction, one makes the parity game more precise: from any starting node, each player that cannot win tries to make the priority with which the game is won by the other player as low as possible.

3 An algorithm for computing the optimal priority

We shall now describe a recursive algorithm that computes the *optimal priority* $p_* : X \to [1, d]$ of a game $\mathcal{G} = \{X,\ E,\ p)$ that has some similarity with the original algorithm proposed by Zielonka [1998] for computing the winning regions of a parity game.

Terminal condition: If the game $\mathcal{G}$ only contains only one priority q, then we know that for all nodes, the optimal parity is q.

Recursion When the game $\mathcal{G}$ has at least two different priorities, let q be the maximal priority. For concreteness, let us assume that q is even (the other case is similar). Let us consider the sub-problem whether Player 0 can win the game with priority q or whether Player 1 can force Player 0 to cycle in nodes with priorities (strictly) lower than q (in $\mathcal{G}$, Player 1 may win or lose, but if he loses, it will be with a priority lower than q). This sub-problem can be cast as a mean payoff game $\mathcal{G}' = (X,\ E,\ w)$ with cost function with values in $\{0, 1\}$:

$$\forall x,\ g(x) = \mathbb{1}_{p(x)=q}.$$

...

We recursively solve the parity game restricted to the set A, i.e. the game $\mathcal{G}\backslash(B \cup C)$, a game which only contains priorities (strictly) lower than q, i.e. obtain for each node $x \in A$ its optimal parity $p_*(x)$. From this, we can propagate ? this optimal priority from A to B by iterating (at most n times):

$$\forall x \in X_0 \cap B, \quad p_*(x) = \max_{y;(x,y)\in E} p_*(y),$$
$$\forall x \in X_1 \cap B, \quad p_*(x) = \min_{y;(x,y)\in E} p_*(y),$$

where the max and min operators above use the order relation $\preceq$ on priorities:

$$p \prec p' \Leftrightarrow (-2)^p < (-2)^{p'}.$$

As there is only one recursive call, and as the maximal priority necessarily decreases at each iteration, the above procedure takes at most d iterations, and

Theorem 1. *A parity game can be solved in polynomial time.*

References

Andrzej Ehrenfeucht and Jan Mycielski. Positional strategies for mean payoff games. *International Journal of Game Theory*, 8:109–113, 1979.

Anuj Puri. *Theory of Hybrid Systems and Discrete Event Systems*. PhD thesis, USA, 1996. UMI Order No. GAX96-21326.

Wiesław Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. *Theor. Comput. Sci.*, 200(1–2):135–183, June 1998. ISSN 0304-3975. doi: 10.1016/S0304-3975(98)00009-7. URL `https://doi.org/10.1016/S0304-3975(98)00009-7`.

Uri Zwick and Mike Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1):343–359, 1996. ISSN 0304-3975. doi: https://doi.org/10.1016/0304-3975(95)00188-3. URL `https://www.sciencedirect.com/science/article/pii/0304397595001883`.