Games where you can play optimally without any memory ⋆

Hugo Gimbert and Wiesław Zielonka

Université Paris 7 and CNRS, LIAFA, case 7014
2, place Jussieu
75251 Paris Cedex 05, France
{hugo,zielonka}@liafa.jussieu.fr

Abstract. Reactive systems are often modelled as two person antagonistic games where one player represents the system while his adversary represents the environment. Undoubtedly, the most popular games in this context are parity games and their cousins (Rabin, Streett and Muller games). Recently however also games with other types of payments, like discounted or mean-payoff [5,6], previously used only in economic context, entered into the area of system modelling and verification. The most outstanding property of parity, mean-payoff and discounted games is the existence of optimal positional (memoryless) strategies for both players. This observation raises two questions: (1) can we characterise the family of payoff mappings for which there always exist optimal positional strategies for both players and (2) are there other payoff mappings with practical or theoretical interest and admitting optimal positional strategies. This paper provides a complete answer to the first question by presenting a simple necessary and sufficient condition on payoff mapping guaranteeing the existence of optimal positional strategies. As a corollary to this result we show the following remarkable property of payoff mappings: if both players have optimal positional strategies when playing solitary one-player games then also they have optimal positional strategies for two-player games.

Introduction

We investigate deterministic games of infinite duration played on finite graphs. We suppose that there are only two players, called Max and Min, with exactly opposite interests. The games are played in the following way. Let G be a finite graph such that each vertex is controlled either by player Max or by player Min. Initially, a pebble is put on some vertex of G. At each step of the play, the player controlling the vertex with the pebble chooses an outgoing edge and moves the pebble along it to the next vertex. Players interact in this way an infinite number of times and a play of the game is simply an infinite path traversed by the pebble.

⋆ This research was supported by European Research Training Network: Games and Automata for Synthesis and Validation and ACI Sécurité Informatique 2003-22 VERSYDIS.

We assume that the edges of G are coloured by elements of a set C of colours. Thus a play yields an infinite sequence of visited colours and it is this sequence that is used to determine the amount of money paid by player Min to player Max; namely we assume that there is a payoff mapping that maps each infinite sequence of colours to the set $\mathbb{R} \cup \{\pm\infty\}$ of extended reals. The objective of player Max is to maximise the outcome of the game while player Min will seek to minimise it. Players plan their actions and such plans are called strategies. Thus a strategy indicates which move to choose in a given situation and this decision may depend on the whole history of previous moves.

For several well-known games: parity, mean-payoff, discounted games, both players can play optimally using particularly simple positional (or memoryless) strategies; their moves depend then only on the current vertex and all previous history is irrelevant [8,12,7,14]. (In fact, for all three payoffs cited above, if the state and action spaces are finite then even more general perfect information stochastic games have optimal deterministic positional strategies). In computer science, the most popular of these games is the parity game used in model-checking and μ-calculus while discounted and mean-payoff games were studied mainly in economics, see however [5,6].

Games with optimal positional strategies are of much interest in computer science since to implement such strategies no memory is needed, which saves computational resources and there is an ongoing quest for new positionally optimal games, especially on push-down graphs [2,1,11,9].

Recently, Colcombet and Niwiński [4] have shown that for infinite graphs if the payoff takes only values 0 and 1 and is prefix independent (the finite prefix of a play has not influence on the payoff value) then only parity games have positional optimal strategies.

While in our paper we consider only games over finite graphs, contrary to [4] we allow general real valued payoff and do not impose any supplementary restriction (like prefix independence). In the previous paper [10] we provided necessary conditions for a payoff mapping guaranteeing the existence of optimal positional strategies. These conditions were robust enough to hold for all popular positional payoffs as well as for several new ones. Nevertheless, there are some trivial positional payoff mappings that do not satisfy the criteria of [10]. In the present paper we improve on the result of [10] by giving a complete characterisation of positional payoff mappings, i.e. we provide conditions that are both sufficient and necessary.

As an application, we describe how to construct, by means of priorities, new positional payoff mappings. As a particular case, we obtain a positional payoff mapping for which both the parity and mean-payoff games are just special cases. This example may be of interest by itself combining qualitative criteria expressed by parity condition with quantitative measures expressed by mean-payoff. Note that recently another combination of parity and mean-payoff games was proposed in [3], however the payoff of [3] happens to be very different from ours, in particular it is not positional.

1 Games, Arenas, Preferences and Optimal Strategies

For any set C, we write C^*, C^+, C^ω to denote respectively the sets of finite, finite non-empty and infinite words over C. In general, for $X \subset C^*$, $X^* = \sum_{i=0}^{\infty} X^i$ is the usual Kleene iteration operation.

We begin by defining arenas where our players meet to confront each other. Let us fix a set C of *colours*. An *arena* coloured by C is a triple

$$G = (S_{\text{Max}}, S_{\text{Min}}, E),$$

where S_{Max} and S_{Min} are two disjoint sets of states and E is the set of coloured transitions. More specifically, if $S = S_{\text{Max}} \cup S_{\text{Min}}$ is the set of all states then $E \subset S \times C \times S$. For a transition $e = (s, c, t) \in E$, the states s, t and the colour c are respectively called the *source*, the *target* and the *colour* of e and we note $\text{source}(e) = s$, $\text{target}(e) = t$ and $\text{colour}(e) = c$. For a state $s \in S$, $sE = \{e \in E \mid \text{source}(e) = s\}$ is the set of transitions outgoing from s.

Throughout this paper, we always assume that arenas have finitely many states and transitions and that each state has at least one outgoing transition.

A *path* in G is a finite or infinite sequence of transitions $p = e_0e_1e_2\ldots$ such that, for all $i \geq 0$, $\text{target}(e_i) = \text{source}(e_{i+1})$. The source $\text{source}(p)$ of p is the source of the first transition e_0. If p is finite then $\text{target}(p)$ is the target of the last transition in p. It is convenient to assume that for each state s there exists an empty path λ_s with no transitions and such that $\text{source}(\lambda_s) = \text{target}(\lambda_s) = s$. The set of finite paths in G, including the empty paths, is denoted P_G^*.

Two players Max and Min play on the arena G in the following way: if the current game position is a state $s \in S_P$ controlled by player $P \in \{\text{Max}, \text{Min}\}$ then player P chooses an outgoing transition $e \in sE$ and the state $\text{target}(e)$ becomes the new game position. If the initial position is s then in this way the players traverse an infinite path $p = e_0e_1e_2\ldots$ in G such that $\text{source}(p) = s$. In the sequel, finite and infinite paths in G are called often (finite and infinite) *plays*.

Every play $p = e_0e_1e_2\ldots$ generates a sequence

$$\text{colour}(p) = \text{colour}(e_0)\,\text{colour}(e_1)\,\text{colour}(e_2)\ldots$$

of visited colours; we call $\text{colour}(p)$ the *colour* of p (i.e. a colour of a play is a sequence of colours rather than a colour).

Players express their preferences for the game outcomes by means of preference relations.

A *preference relation* over a set C of colours is a binary complete, reflexive and transitive relation over the set C^ω of infinite colour sequences (complete means here that for all $x, y \in C^\omega$ either $x \sqsubseteq y$ or $y \sqsubseteq x$). Thus $\sqsubseteq$ is in fact a complete preorder relation over infinite colour sequences.

Intuitively, if $x \sqsubseteq y$ then the player whose preference relation is $\sqsubseteq$ appreciates the sequence y at least as much as the sequence x. On the other hand, if $x \sqsubseteq y$ and $y \sqsubseteq x$ then the outcomes x and y have the same value for our player, we shall say that x and y are *equivalent for* $\sqsubseteq$. By $\sqsubseteq^{-1}$ by denote the inverse of $\sqsubseteq$, $x \sqsubseteq^{-1} y$ iff $y \sqsubseteq x$.

We shall write $x \sqsubset y$ to denote that $x \sqsubseteq y$ but not $y \sqsubseteq x$.

A two-person game is a triple $(G, \sqsubseteq_{\text{Max}}, \sqsubseteq_{\text{Min}})$, where G is a finite arena and $\sqsubseteq_{\text{Max}}, \sqsubseteq_{\text{Min}}$ are preference relations for players Max and Min. The obvious aim of each player is to obtain the most favourable for him infinite colour sequence.

We will investigate only antagonistic games where the preference relation for player Min is just the inverse of the preference relation of player Max. One of the preference relations being redundant in this case, *antagonistic games* (or simply *games* in the sequel) are just pairs $(G, \sqsubseteq)$, where $\sqsubseteq$ is the preference relation of player Max and G a finite arena.

Most often preference relations are introduced by means of *payoff* or utility mappings. Such a mapping $u : C^{\omega} \to \mathbb{R} \cup \{-\infty, +\infty\}$ maps infinite colour sequences to extended real numbers. If u is the payoff mapping of player Max for example and the game outcome is an infinite colour sequence $x \in C^{\omega}$ then player Max receives the payoff $u(x)$. A payoff mapping u induces a natural preference relation $\sqsubseteq_u$ compatible with u and defined by $x \sqsubseteq_u y$ iff $u(x) \leq u(y)$.

Although in game theory preference relations are slightly less employed than payoff mappings they are still standard, for example preference relations are largely used in the popular textbook of Osborne and Rubinstein [13]. We have chosen here to base our exposition on preference relations rather than on payoffs for several reasons: first of all the proofs are more comprehensive when written in the language of preference relations, secondly, one really does not need precise payoff values unless the so-called ϵ-optimal strategies are considered which is not the case in this paper, finally, for some preference relations it would be artificial, cumbersome and counterintuitive to define a corresponding payoff mapping (while, as noted before, the converse is always true, a payoff defines immediately a preference relation).

Intuitively, a strategy of a player is a method he uses to choose his moves during the play. Thus for each finite play p that arrives at a state controlled by player P, $\text{target}(p) \in S_P$, the strategy indicates a transition with the source in the state $\text{target}(p)$ to be taken by player P after p. Therefore in general a *strategy for player* P is a mapping

$$\sigma_P : \{p \in P_G^* \mid \text{target}(p) \in S_P\} \to E,$$

such that $\sigma_P(p) \in sE$ if $s = \text{target}(p)$.

A finite or infinite play $p = e_0 e_1 e_2 \ldots$ is said to be *consistent* with the strategy σ_P if whenever $\text{target}(e_i) \in S_P$ then $e_{i+1} = \sigma_P(e_0 ... e_i)$ and moreover $e_0 = \sigma_P(\lambda_s)$ if $s = \text{source}(p) \in \sigma_P$.

A *positional* (or memoryless) strategy for player P is a mapping $\sigma_P : S_P \to E$ such that for all $s \in S_P$, $\sigma_P(s) \in sE$. Using such a strategy σ_P, after a finite play p with $\text{target}(p) \in V_P$ player P chooses the transition $\sigma_P(\text{target}(p))$, i.e. the chosen transition depends only on the current game position. Our interest in positional strategies is motivated by the fact that they are especially easy to implement, no memory of the past history is needed.

In the sequel σ and τ, possibly with subscripts or superscripts, will always denote strategies for players Max and Min respectively.

Given a state t and strategies σ and τ for players Max and Min, there exists a unique play in G, denoted by $p_G(t, \sigma, \tau)$, with source t consistent with both σ and τ.

Strategies $\sigma^\#$ and $\tau^\#$ are called *optimal* if for all states $s \in S$ and all strategies σ and τ of both players

$$\mathrm{colour}(p_G(s, \sigma, \tau^\#)) \sqsubseteq \mathrm{colour}(p_G(s, \sigma^\#, \tau^\#)) \sqsubseteq \mathrm{colour}(p_G(s, \sigma^\#, \tau)) \ . \tag{1}$$

Inequalities above mean that players Max and Min have no incentive to deviate unilaterally from their optimal strategies.

It is easy to see that if $(\sigma_1^\#, \tau_1^\#)$ and $(\sigma_2^\#, \tau_2^\#)$ are pairs of optimal strategies then $(\sigma_1^\#, \tau_2^\#)$ and $(\sigma_2^\#, \tau_1^\#)$ are optimal and in fact $\mathrm{colour}(p_G(s, \sigma_1^\#, \tau_1^\#)))$ and $\mathrm{colour}(p_G(s, \sigma_2^\#, \tau_2^\#)))$ are equivalent for $\sqsubseteq$.

2 Preferences Relations with Optimal Positional Strategies.

The main aim of this section it to provide a complete characterisation of preference relations for which both players have optimal positional strategies for all games on finite arenas.

Let $\mathrm{Rec}(C)$ be the family of recognizable subsets of C^* (C can be infinite and then $L \in \mathrm{Rec}(C)$ means that there exists a finite subset B of C such that L a recognizable subset of B^*). For any language of finite words $L \subset C^*$, $\mathrm{Pref}(L)$ will stand for the set of all prefixes of the words in L. We define an operator $[\cdot]$ that associates with each language $L \subset C^*$ of finite words a set $[L] \subset C^\omega$ of infinite words:

$$[L] = \{x \in C^\omega \mid \text{every finite prefix of } x \text{ is in } \mathrm{Pref}(L)\} \ .$$

We extend the preference relation $\sqsubseteq$ to subsets of C^ω: for $X, Y \subset C^\omega$,

$$X \sqsubseteq Y \quad \text{iff} \quad \forall x \in X, \exists y \in Y, x \sqsubseteq y \ .$$

Obviously, for $x \in C^\omega$ and $Y \subset C^\omega$, $x \sqsubseteq Y$ and $Y \sqsubseteq x$ stand for $\{x\} \sqsubseteq Y$ and $Y \sqsubseteq \{x\}$ respectively. We write also

$$X \sqsubset Y \quad \text{iff} \quad \exists y \in Y, \forall x \in X, x \sqsubset y \ .$$

Definition 1. *A preference relation* $\sqsubseteq$ *is said to be* monotone *if for all recognizable sets* $M, N \in \mathrm{Rec}(C)$,

$$\exists x \in C^*, [xM] \sqsubset [xN] \implies \quad \forall y \in C^*, [yM] \sqsubseteq [yN] \ .$$

A preference relation $\sqsubseteq$ *is said to be* selective *if for each finite word* $x \in C^*$ *and all recognizable languages* $M, N, K \in \mathrm{Rec}(C)$,

$$[x(M \cup N)^* K] \sqsubseteq [xM^*] \cup [xN^*] \cup [xK] \ .$$

Now we are ready to state the main result of this paper.

Theorem 2. *Given a preference relation $\sqsubseteq$, both players have optimal positional strategies for all games $(G, \sqsubseteq)$ over finite arenas G if and only if the relations $\sqsubseteq$ and its inverse $\sqsubseteq^{-1}$ are monotone and selective.*

Before proceeding to the proof of Theorem 2 it can be useful to convey some intuitions behind the definitions of monotone and selective properties.

Roughly speaking, a preference relation of Max is monotone if at each moment during the play the optimal choice of player Max between two possible futures does not depend on the preceding finite play. For example, consider the payoff function u defined on the set $C = \mathbb{R}$ of colours by the formula

$$u(x_1 x_2 \ldots) = \sup_{n \in \mathbb{N}} \frac{1}{n} \sum_{k=1}^{n} x_k, \tag{2}$$

where $x_1 x_2 \ldots$ is an infinite sequence of real numbers. Consider the finite sequences $x = 0000$ and $y = 1111$ and the infinite sequences $v = 2000 \ldots = 20^\omega$ and $w = 1111 \ldots = 1^\omega$. Then $u(xv) < u(xw)$ while $u(yw) < u(yv)$, hence the preference relation $\sqsubseteq_u$ associated with u is not monotone. This means that player Max has no optimal positional strategy in the one-player arena depicted on the left of Fig 1, if Max plays optimally the transition to take at state z depends on whether he arrives from s or from t. It is worth to note that the payoff (2) is selective.

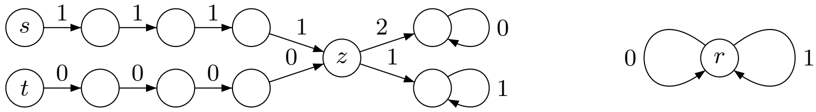

Fig. 1. When playing on the left arena using the non-monotone payoff (2), or playing on the right arena using the non-selective payoff "wins 1 if the colours 0 and 1 appear infinitely often and 0 otherwise" player Max has no optimal positional strategies.

The selective property expresses the fact that player Max cannot improve his payoff by switching between different behaviors. Typical non selective payoff is provided by the Muller condition. Let u be the payoff function for $C = \{0, 1\}$ defined by $u(x_0 x_1 \ldots) = 1$ if the colours 0 and 1 occur infinitely often, otherwise the payoff is 0. This payoff mapping is monotone (as are all payoffs that do not depend on finite prefixes) but is not selective. It is clear that when Max plays with this payoff on the one-player arena depicted on the right of Fig 1 then he should alternate infinitely often between the two transitions to maximize his payoff.

We begin the proof of Theorem 2 by noting the following trivial property of the operator $[\cdot]$:

Lemma 3. *For all* $L, M \subset C^*$, $[L \cup M] = [L] \cup [M]$.

A finite (non-deterministic) automaton over C is a tuple $\mathcal{A} = (Q, i, F, \Delta)$, where Q is a finite set of states, $i \in Q$ the initial state, $F \subset Q$ the set of final states and $\Delta \subset Q \times C \times Q$ is the transition relation. A path in $\mathcal{A}$ is a path in the one-player arena $(Q, \emptyset, \Delta)$ that we can construct from $\mathcal{A}$ and the notions of source, target and colour of a path are defined as for arenas. So, in this terminology, the language recognized by $\mathcal{A}$ is simply the set
$\{\text{colour}(p) \mid p \text{ is a finite path in } \mathcal{A} \text{ such that } \text{source}(p) = i \text{ and } \text{target}(p) \in F\}$.
The automaton $\mathcal{A}$ is said to be *co-accessible* if from any state there is a (possibly empty) path to a final state.

Lemma 4. *Let* $\mathcal{A} = (Q, i, F, \Delta)$ *be a co-accessible finite automaton recognizing a language* $L \subset C^*$. *Then*

$$[L] = \{\text{colour}(p) \mid p \textit{ is an infinite path in } \mathcal{A} \textit{ with } \text{source}(p) = i\}.$$

Proof. Let $p = e_0 e_1 e_2 \ldots$ be an infinite path in $\mathcal{A}$, where $\forall j, e_j \in \Delta$ and $\text{source}(e_0) = i$. Since $\mathcal{A}$ is co-accessible, for every n there is a path from the state $\text{target}(e_n)$ to a final state. Therefore the finite word $\text{colour}(e_0 \ldots e_n)$ is a prefix of some word recognized by $\mathcal{A}$. Hence $\text{colour}(p) \in [L]$.

Conversely, let $x = c_0 c_1 c_2 \cdots \in [L]$. Let T be the directed tree defined as follows. The vertices of T are finite paths q in $\mathcal{A}$ such that $\text{colour}(q)$ is a prefix of x and $\text{source}(q) = i$. There is an edge from a vertex q of T to a vertex q' iff there is a transition $e \in \Delta$ such that $q' = qe$. The root of T is the empty path λ_i with the source and target i. Clearly, T is infinite since x is infinite and the degree of vertices of T is bounded by the cardinality of Δ. Hence, by the Koenig Lemma, there exists an infinite path in T starting from the root λ_i. This infinite path corresponds to an infinite path in $\mathcal{A}$ coloured by x. □

It turns out that already for one-player games controlled by player Max to guarantee that Max has an optimal positional strategy it is necessary for his preference relation $\sqsubseteq$ to be monotone and selective:

Lemma 5. *Suppose that player* Max *has optimal positional strategies for all games* $(G, \sqsubseteq)$ *over finite one-player arenas* $G = (S_{\text{Max}}, \emptyset, E)$, *where he controls all states. Then* $\sqsubseteq$ *is monotone and selective.*

Proof. We want to use finite automata as one-player arenas with all states controlled by player Max. Technically however, this raises a problem since we require that arenas have always at least one outgoing transition for each state s and this condition may fail for automata. For this reason we introduce the following notion.

For any finite automaton $\mathcal{A} = (Q, i, F, \Delta)$, a state $s \in Q$ is said to be *essential* if there exists an infinite path in $\mathcal{A}$ with source s. A transition is essential if its target is essential. Note that for any essential state s there is at least one essential transition with source s and any infinite path in $\mathcal{A}$ traverses uniquely essential states and transitions. Moreover, by Lemma 4, if $\mathcal{A}$ is co-accessible and recognizes L then $[L] \neq \emptyset$ iff the initial state is essential. By arena($\mathcal{A}$) we shall denote the arena $(Q', \emptyset, \Delta')$, where Q' and Δ' are respectively the sets of essential states and essential transitions of $\mathcal{A}$.

Suppose that $\sqsubseteq$ satisfies the hypothesis of our lemma. We show first that $\sqsubseteq$ is monotone. Let $x, y \in C^*$ and $M, N \in \mathrm{Rec}(C)$ and

$$[xM] \sqsubset [xN] \ . \tag{3}$$

We shall prove that this implies

$$[yM] \sqsubseteq [yN] \ . \tag{4}$$

Let $\mathcal{A}_x$ and $\mathcal{A}_y$ be the usual deterministic co-accessible automata recognizing the one-word languages $\{x\}$ and $\{y\}$. Let $\mathcal{A}_M, \mathcal{A}_N$ be finite co-accessible automata recognizing respectively M, N. Without loss of generality we can assume that neither $\mathcal{A}_M$ nor $\mathcal{A}_N$ has a transition with the initial state as the target.

If $[M]$ is empty then (4) holds trivially. Thus we can assume that $[M]$ and $[N]$ are non-empty and the initial states of $\mathcal{A}_M$ and $\mathcal{A}_N$ are essential.

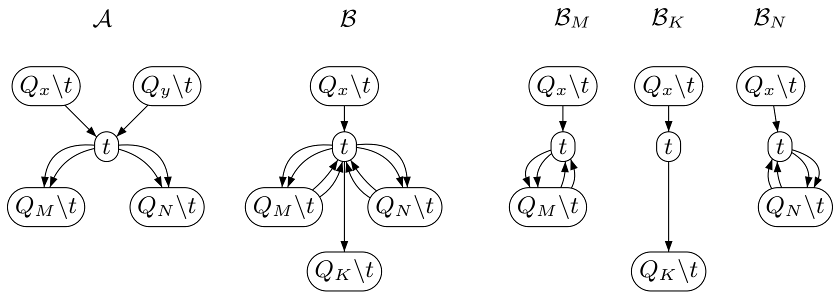

Fig. 2. Automaton $\mathcal{A}$ used to prove that $\sqsubseteq$ is monotone is obtained by "gluing" together the final states of $\mathcal{A}_x$ and $\mathcal{A}_y$ with initial states of $\mathcal{A}_M$ and $\mathcal{A}_N$. Q_x, Q_y, Q_M, Q_N are the states of the corresponding automata. Automaton $\mathcal{B}$ used to prove that $\sqsubseteq$ is selective is obtained by "gluing" together the final state of $\mathcal{A}_x$, the initial and the final states of $\mathcal{A}_M$ and $\mathcal{A}_N$ and the initial state of $\mathcal{A}_K$.

From automata $\mathcal{A}_x, \mathcal{A}_y, \mathcal{A}_M, \mathcal{A}_N$ we obtain a new automaton $\mathcal{A}$ by identifying the following four states: the final state of $\mathcal{A}_x$, the final state of $\mathcal{A}_y$, the initial state of $\mathcal{A}_M$ and the initial state of $\mathcal{A}_N$. We note t the state obtained in this way. The transitions of $\mathcal{A}_x$ and $\mathcal{A}_y$ with target in the final state have target t in $\mathcal{A}$ while the transitions of $\mathcal{A}_M, \mathcal{A}_N$ with the source in the initial state have source t in $\mathcal{A}$. All the other states and transitions remain unchanged in $\mathcal{A}$, see Fig 2. The final states of $\mathcal{A}_M$ and $\mathcal{A}_N$ are final in $\mathcal{A}$ while the initial state of $\mathcal{A}_x$ is initial in $\mathcal{A}$. Note that since $\mathcal{A}_M$ and $\mathcal{A}_N$ are co-accessible $\mathcal{A}$ is also co-accessible. Moreover, $\mathcal{A}$ recognizes the language $x(M \cup N)$ (since we have assumed that no transition of $\mathcal{A}_M$ and $\mathcal{A}_N$ returns to the initial state).

Let $\sigma^\#$ be an optimal positional strategy of player Max in the game $(\text{arena}(\mathcal{A}), u)$. Then, by Lemma 4 applied to $\mathcal{A}$, the set of plays in arena($\mathcal{A}$) starting from the initial state of $\mathcal{A}$ is $[x(M \cup N)]$, which is equal to $[xM] \cup [xN]$ by Lemma 3. Let p be the unique infinite play in arena($\mathcal{A}$) with source in the initial state of $\mathcal{A}$ and consistent with the strategy $\sigma^\#$. Then by optimality of $\sigma^\#$, $[xM] \cup [xN] \sqsubseteq \text{colour}(p)$ implying, by (3), $\text{colour}(p) \notin [xM]$.

Therefore, play p reaching the state t takes a transition leading to the states of $\mathcal{A}_N$ (Fig. 2) and stays forever in $\mathcal{A}_N$ in the sequel. In other words, we can conclude that $\sigma^\#(t)$ is a transition of $\mathcal{A}_N$.

Now let us examine the unique infinite play q in arena($\mathcal{A}$) consistent with $\sigma^\#$ and starting at the initial state of $\mathcal{A}_y$. Since q is consistent with $\sigma^\#$ and $\sigma^\#(t)$ is a transition of $\mathcal{A}_N$, play q traverses first the states of automaton $\mathcal{A}_y$ and next the states of $\mathcal{A}_N$.

Since from all the states traversed by q we can reach in $\mathcal{A}$ the final states of $\mathcal{A}_N$, we have

$$\text{colour}(q) \in [yN] \ . \tag{5}$$

On the other hand, for the same reasons as for $\mathcal{A}$ but now with the initial state of A_y, the optimality of $\sigma^\#$ yields $[yM] \cup [yN] \sqsubseteq \text{colour}(q)$. This and (5) imply immediately (4).

It remains to prove that $\sqsubseteq$ is selective. Let $x \in C^*$, $M, N, K \in \text{Rec}(C)$. Without loss of generality we can assume that M and N do not contain the empty word and choose the automata $\mathcal{A}_M$ and $\mathcal{A}_N$ recognizing M and N to be co-accessible, with one initial and one final state and with no transition returning to the initial state and no transition leaving the final state. Let A_K be a co-accessible automaton recognizing K with no transition returning to its initial state. We glue together the final states of automata $\mathcal{A}_x, \mathcal{A}_M, \mathcal{A}_N$ and the initial states of $\mathcal{A}_M, \mathcal{A}_N, \mathcal{A}_K$. The resulting state is called t. Taking the initial state from $\mathcal{A}_x$ and the final states from $\mathcal{A}_K$ we obtain an automaton $\mathcal{B}$.

Let $\sigma^\#$ be an optimal positional strategy of player Max in the game $(\text{arena}(\mathcal{B}), \sqsubseteq)$. Let p be the infinite play consistent with $\sigma^\#$ and with the initial state of $\mathcal{B}$ as the source. Automaton $\mathcal{B}$ is co-accessible and recognizes the language $x(M \cup N)^*K$, therefore, by Lemma 4 and optimality of $\sigma^\#$,

$$[x(M \cup N)^*K] \sqsubseteq \text{colour}(p) \ . \tag{6}$$

Since $\sigma^\#$ is positional, each time p traverses the state t, $\sigma^\#$ chooses the same outgoing transition. This means that p is an infinite path in one of the three co-accessible automata $\mathcal{B}_M, \mathcal{B}_N, B_K$ depicted on Fig. 2. By Lemma 4, $\text{colour}(p) \sqsubseteq [xM^*] \cup [xN^*] \cup [xK]$. This and (6) imply that u is selective. □

With each arena G with a state set S and a transition set E we associate the *index* n_G of G defined as $n_G = |E| - |S|$. Note that since in arenas each state has at least one outgoing transition the index is always non-negative. The proof of Theorem 2 will be carried on by induction on the value of n_G and the decisive inductive step is provided by the following lemma.

Lemma 6. *Let G be an arena and $\sqsubseteq$ a monotone and selective preference relation. Suppose that players* Max *and* Min *have optimal positional strategies in all games $(H, \sqsubseteq)$ over the arenas H such that $n_H < n_G$. Then* Max *has an optimal positional strategy in the game $(G, \sqsubseteq)$.*

Proof. Let $G = (S_{\text{Max}}, S_{\text{Min}}, E)$ and let $\sqsubseteq$ be monotone and selective. If for every $t \in S_{\text{Max}}$ there is only one transition with the source t then Max has never any choice and he has therefore a unique strategy which is positional and optimal.

Suppose now that there exists a state $t \in S_{\text{Max}}$ such that $|tE| > 1$. Fix a partition of tE into two disjoint non-empty sets A_0, A_1. We define two new arenas $G_i = (S_{\text{Max}}, S_{\text{Min}}, E_i)$, $i = 0, 1$, where $E_i = E \setminus A_{1-i}$. In other words, G_i is obtained by removing from G the transitions with the source t not belonging to A_i. Since $n_{G_i} < n_G$ we can apply the hypothesis of our lemma to the games $\mathbf{G}_i = (G_i, \sqsubseteq)$ to conclude that in both games $\mathbf{G}_i$ players Max and Min have optimal positional strategies $\sigma_i^{\#}$, $\tau_i^{\#}$ respectively. Let us note $\mathbf{G} = (G, \sqsubseteq)$ the initial game over G.

Let $M_i \subset C^*$ be the set of finite colour sequences colour(p) of all finite plays p in G_i that are consistent with strategy $\tau_i^{\#}$ and have source and target t.

To see that $M_i \in \text{Rec}(C)$, we can build a finite automaton with the same state space as for the arena G_i, we keep also all transitions of G_i that have the source in the set S_{Max}, however for each state $s \in S_{\text{Min}}$ controlled by player Min we keep only one outgoing transition, namely the transition $\tau_i^{\#}(s) \in sE_i$ chosen by the strategy $\tau_i^{\#}$. Then M_i is the set of words recognized by such an automaton if we take t as the initial and the final state.

Now we define the sets $K_i \subset C^*$, $i = 0, 1$ consisting of colours colour(p) of all finite plays p in the arena G_i that have source t and are consistent with $\tau_i^{\#}$ (but can end in any state of G_i). Again it should be obvious that $K_i \in \text{Rec}(C)$.

The monotonicity of $\sqsubseteq$ implies that either $\forall x \in C^*, [xK_0] \sqsubseteq [xK_1]$ or $\forall x \in C^*, [xK_1] \sqsubseteq [xK_0]$ Since the former condition is symmetric to the latter, without loss of generality, we can assume that

$$\forall x \in C^*, [xK_1] \sqsubseteq [xK_0] \ . \tag{7}$$

Let us set

$$\sigma^{\#} = \sigma_0^{\#} \ . \tag{8}$$

We shall show that, if (7) holds then strategy $\sigma^{\#}$ is not only optimal for player Max in the game $\mathbf{G}_0$ but it is also optimal for him in $\mathbf{G}$. It is clear that $\sigma^{\#}$ is a well-defined positional strategy for Max in the game $\mathbf{G}$. To finish the proof of Lemma 6 we should construct a strategy $\tau^{\#}$ for player Min such that $(\sigma^{\#}, \tau^{\#})$ is a couple of optimal strategies. However, contrary to $\sigma^{\#}$, to implement the strategy $\tau^{\#}$ player Min will need some finite memory.

We define first a mapping $h : P^*_G \to \{0,1\}$ that assigns to each finite play $p \in P^*_G$ in G a one bit value $h(p)$:

$$h(p) = \begin{cases} 0 & \text{if either } p \text{ does not contain any transition with the source } t \text{ or} \\ & \text{the last transition of } p \text{ with the source } t \text{ belongs to } A_0, \\ 1 & \text{if the last transition of } p \text{ with the source } t \text{ belongs to } A_1. \end{cases}$$

Then the strategy $\tau^\#$ of Min in $\mathbf{G}$ is defined by

$$\tau^\#(p) = \begin{cases} \tau^\#_0(\mathrm{target}(p)) & \text{if } h(p) = 0, \\ \tau^\#_1(\mathrm{target}(p)) & \text{if } h(p) = 1, \end{cases}$$

for finite plays p with $\mathrm{target}(p) \in S_{\mathrm{Min}}$. In other words, playing in $\mathbf{G}$ player Min applies either his optimal strategy $\tau^\#_0$ from the game $\mathbf{G}_0$ or his optimal strategy $\tau^\#_1$ from the game $\mathbf{G}_1$ depending on the value $h(p)$. Initially, before the first visit to t, player Min uses the strategy $\tau^\#_0$. After the first visit to t the choice between $\tau^\#_0$ and $\tau^\#_1$ depends on the transition chosen by his adversary Max at the last visit to t, if the chosen transition was in A_0 then player Min uses the strategy $\tau^\#_0$, otherwise, if Max took a transition of A_1 then player Min plays according to $\tau^\#_1$. The intuition behind the definition of $\tau^\#$ is the following: If at the last visit to t player Max has chosen a outgoing transition from A_0 then this means that the play from this moment onward is like a play in $\mathbf{G}_0$ and therefore player Min tries to respond using his optimal strategy from $\mathbf{G}_0$. Symmetrically, if at the last visit to t player Max has chosen an outgoing transition from A_1 then from this moment onward the play is like a play in $\mathbf{G}_1$ and player Min tries to counter with his optimal strategy from $\mathbf{G}_1$.

It should be clear that the strategy $\tau^\#$ needs in fact just two valued memory $\{0,1\}$ for player Min to remember if during the last visit to t a transition of A_0 or a transition of A_1 was chosen by his adversary. This memory is initialised to 0 and updated only when the state t is visited.

We shall prove that $(\sigma^\#, \tau^\#)$ is a couple of optimal strategies in $\mathbf{G}$, i.e. (1) holds for any strategies σ, τ of players Max and Min and any initial state s.

In the sequel we shall write frequently $p \sqsubseteq q$ for infinite plays p and q as an abbreviation of $\mathrm{colour}(p) \sqsubseteq \mathrm{colour}(q)$.

Let τ be any strategy for player Min in the game $\mathbf{G}$ and let τ_0 be its restriction to the set $P^*_{G_0}$ of finite plays in the arena G_0. Clearly τ_0 is a valid strategy of Min over the arena G_0. Then for any state s of G

$$\begin{aligned} p_G(s, \sigma^\#, \tau^\#) &= p_{G_0}(s, \sigma^\#_0, \tau^\#_0) && \text{by definition of } \sigma^\# \text{ and } \tau^\#, \\ &\sqsubseteq p_{G_0}(s, \sigma^\#_0, \tau_0) && \text{by optimality of } (\sigma^\#_0, \tau^\#_0) \text{ in } \mathbf{G}_0, \\ &= p_G(s, \sigma^\#, \tau) && \text{by definition of } \sigma^\# \text{ and } \tau_0, \end{aligned}$$

which concludes the proof of the right hand side inequality in (1).

Now let σ be any strategy for player Max in $\mathbf{G}$ and s any state of G. There are two cases to examine depending on whether the play $p_G(s, \sigma, \tau^\#)$ traverses t or not.

Case 1: $p_G(s, \sigma, \tau^\#)$ does not traverse the state t.

In this case, according to the definition of $\tau^\#$, player Min uses in fact all the time during this play the strategy $\tau_0^\#$, never switching to $\tau_1^\#$.

Let us take any strategy σ_0 for player Max which is defined exactly as σ for all finite plays with the target different from t while for plays with target t the strategy σ_0 chooses always a transition of A_0. The last condition implies that σ_0 is also a valid strategy over the arena G_0. Moreover, since $p_G(s, \sigma, \tau^\#)$ never traverses t the strategies σ and σ_0 choose the same transitions for all finite prefixes of $p_G(s, \sigma, \tau^\#)$ with the target the state controlled by player Max, therefore $p_G(s, \sigma, \tau^\#) = p_{G_0}(s, \sigma_0, \tau_0^\#)$. However, $p_{G_0}(s, \sigma_0, \tau_0^\#) \sqsubseteq p_{G_0}(s, \sigma_0^\#, \tau_0^\#) = p_G(s, \sigma^\#, \tau^\#)$, where the first inequality follows from optimality of $\sigma_0^\#, \tau_0^\#$ in $\mathbf{G}_0$ while the last equality is just the consequence of (8) and definition $\tau^\#$. Therefore, $p_G(s, \sigma, \tau^\#) \sqsubseteq p_G(s, \sigma^\#, \tau^\#)$, i.e. the left-hand side of (1) holds in this case.

Case 2: $p_G(s, \sigma, \tau^\#)$ traverses the state t.

Let p' be the shortest finite play such that p' is a prefix of $p_G(s, \sigma, \tau^\#)$ and $\text{target}(p') = t$. Note that by the definition of $\tau^\#$ it follows that p' is in fact consistent with $\tau_0^\#$. Let $\text{colour}(p') = x$.

Then by definition of x, M_0, M_1, K_0 and K_1, any prefix of $\text{colour}(p_G(s, \sigma, \tau^\#))$ longer than x belongs to the set $x(M_0 \cup M_1)^*(K_0 \cup K_1)$, hence

$$\begin{aligned} \text{colour}(p_G(s, \sigma, \tau^\#)) &\in [x(M_0 \cup M_1)^*(K_0 \cup K_1)] \\ &\sqsubseteq [x(M_0)^*] \cup [x(M_1)^*] \cup [x(K_0 \cup K_1)] && \text{since } \sqsubseteq \text{ is selective,} \\ &\sqsubseteq [x(M_0)^*] \cup [x(M_1)^*] \cup [xK_0] \cup [xK_1] && \text{by Lemma 3,} \\ &\sqsubseteq [xK_0] \cup [xK_1] && \text{since } (M_i)^* \subset K_i, \\ &\sqsubseteq [xK_0] && \text{by (7).} \end{aligned} \tag{9}$$

Let us define a new transition set $\delta \subset E$, where E is the the set of transitions of the arena G: for any state r of G the set of transitions with source r under δ is defined by:

$$r\delta = \begin{cases} A_0 & \text{if } r = t, \\ rE & \text{if } r \in S_{\text{Max}} \setminus \{t\}, \\ \tau_0^\#(r) & \text{if } r \in S_{\text{Min}}. \end{cases} \tag{10}$$

Let Q be the set of states of G that are accessible from t under δ. Take a finite automaton $\mathcal{D}$ with the initial state t, the set of states Q all of which are final and the transition relation δ restricted to Q.

Automaton $\mathcal{D}$ is co-accessible, recognizes the language K_0 and therefore, by Lemma 4, $[K_0]$ is precisely the set of colour sequences $\text{colour}(q)$ of infinite plays q with source t that are consistent with $\tau_0^\#$.

Let U be the set of all colour sequences $\text{colour}(q')$ of infinite plays q' in G_0 with source s that are consistent with $\tau_0^\#$. Then $x[K_0] \subset U$ implying that

$$x[K_0] \sqsubseteq U \sqsubseteq \text{colour}(p_{G_0}(s, \sigma_0^\#, \tau_0^\#)), \tag{11}$$

where the last inequality follows from optimality of $\sigma_0^\#$ in the game $\mathbf{G}_0$. But, by definition of $\sigma^\#$ and $\tau^\#$, we get $p_{G_0}(s, \sigma_0^\#, \tau_0^\#) = p_G(s, \sigma^\#, \tau^\#)$, which together with (9) and (11) yield $p_G(s,\sigma,\tau^{\#}) \sqsubseteq p_G(s,\sigma^{\#},\tau^{\#})$ terminating the proof of the left hand-side of (1) in this case.

□

Proof. of Theorem 2. Note that, due to symmetry, we can permute players Max and Min and replace the preference relation $\sqsubseteq$ by $\sqsubseteq^{-1}$ in Lemmas 5 and 6. Since, clearly, players have optimal positional strategies for the preference relation $\sqsubseteq$ iff they have optimal positional strategies with the preference $\sqsubseteq^{-1}$ under the permutation (in fact these are the same strategies), Lemma 5 shows that to be monotone and selective for $\sqsubseteq$ and $\sqsubseteq^{-1}$ is necessary for the existence of optimal positional strategies.

Now Lemma 6 allows us to apply a trivial induction over the arena index to conclude immediately that these conditions are also sufficient. □

The following corollary turns out to be much more useful in practice than Theorem 2 itself.

Corollary 7. *Suppose that $\sqsubseteq$ is such that for each finite arena $G = (S_{\mathrm{Max}}, S_{\mathrm{Min}}, E)$ controlled by one player, i.e. such that either $S_{\mathrm{Max}} = \emptyset$ or $S_{\mathrm{Min}} = \emptyset$, the player controlling all states of G has an optimal positional strategy in the game $(G, \sqsubseteq)$. Then for all finite two-player arenas G both players have optimal positional strategies in the games $(G, \sqsubseteq)$.*

Proof. By Lemma 5 if both players have optimal positional strategies on one-player games then $\sqsubseteq$ and $\sqsubseteq^{-1}$ are monotone and selective and then, by Theorem 2, they have optimal positional strategies on all two-person games on finite arenas. □

3 An example: Priority Mean-payoff Games.

The interest in Corollary 7 stems from the fact that often it is quite trivial to verify if a given preference relation is positional for one-player games. To illustrate this point let us consider mean-payoff games [7]. Here colours are real numbers and for an infinite sequence $r_1 r_2 \dots$ of elements of $\mathbb{R}$ the payoff is calculated by $\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} r_i$. Suppose that G is an arena controlled by player Max. Take in G a simple cycle (in the sense of graph theory) with the maximal mean value. It is easy to see that any other infinite play in G cannot supply a payoff greater than the mean-payoff over this cycle. Thus the optimal positional strategy for player Max is to go as quickly as possible to this maximum payoff cycle and next go round this cycle forever. Clearly, player Min has also optimal positional strategies for all arenas where he controls all states and Corollary 7 allows us to conclude that in mean-payoff games both players have optimal positional strategies.

As a more sophisticated example illustrating Corollary 7 we introduce here *priority mean-payoff games.* Let $C = \{0,\dots,k\} \times \mathbb{R}$ be the set of colours, where for each couple $(m,r) \in \{0,\dots,k\} \times \mathbb{R}$ the non-negative integer m is called the *priority* and r is a real-valued reward. The payoff for an infinite sequence $x = (m_1, r_1), (m_2, r_2), \ldots$ of colours is calculated in the following way: let $k = \limsup_{i\to\infty} m_i$ be the maximal priority appearing infinitely often in x and let $i_1 < i_2 < \ldots$ be the infinite sequence of all positions in x with the priority k, i.e. $k = m_{i_1} = m_{i_2} = \ldots$. Then the priority mean-payoff is calculated as the mean payoff of the corresponding subsequence $r_{i_1} r_{i_2} \ldots$ of real rewards, $\limsup_{t\to\infty} \frac{1}{t} \sum_{n=1}^{t} r_{i_n}$. In other words priorities are used here to select an appropriate subsequence of real rewards for which the mean-payoff mapping is applied subsequently.

This payoff, rather contrived at first sight, is in fact a common natural generalization of mean-payoff and parity payoffs. On the one hand, we recover simple mean-payoff games if there is only one priority. On the other hand, if we allow only a subset of colours consisting of couples (m, r) such that r is 1 if m is odd and r is 0 for m even then the rewards associated with the maximal priority are constant and we just obtain the parity game coded in an unusual manner.

Instead of proving immediately that the priority mean-payoff mapping admits optimal positional strategies let us generalize it slightly before.

Let $u_0, \ldots, u_k$ be payoff mappings on the set C of colours. We define a payoff mapping u on the set $\mathcal{B} = \{0, \ldots, k\} \times C$ of colours which we shall call *priority product* of $u_0, \ldots, u_k$. In the sequel we call the elements of $\{0, \ldots, k\}$ priorities. Let $x = (p_1, c_1), (p_2, c_2), \ldots \in \mathcal{B}^\omega$ be an infinite colour sequence of elements of $\mathcal{B}$. Define priority(x) to be the highest priority appearing infinitely often in x: priority$(x) = \limsup_{i\to\infty} p_i$.

Let $(j_m)_{m=0}^{\infty}$ be the sequence of positions in x with priority priority(x), priority$(x) = p_{j_1} = p_{j_2} = p_{j_3} = \ldots$. Then the priority product gives us the payoff

$$u(x) = u_m(c_{j_1} c_{j_2} c_{j_3} \ldots), \quad \text{where } m = \text{priority}(x) \ .$$

A payoff mapping u is said to be *prefix-independent* if $\forall x \in C^*$, $\forall y \in C^\omega$, $u(xy) = u(y)$.

Lemma 8. *If u_i, $i = 0, \ldots, k$, are prefix-independent and admit all optimal positional strategies for both players for all games on finite arenas then their priority product u admits optimal positional strategies for both players on all finite arenas.*

Note first that the priority product of several mean-payoff mappings is just the priority mean-payoff mapping. Thus Lemma 8 implies that on finite arenas priority mean-payoff mapping admits optimal positional strategies for both players.

Proof. We prove that, under the conditions of Lemma 8, player Max has optimal positional strategies on one-player arenas. Let G be such an arena. For each simple cycle in G we can calculate the value of the payoff u for the play that turns round the cycle forever. Let a be the maximal payoff calculated in this way and c the cycle giving this value. We prove that for any infinite play p on G, $u(\text{colour}(p)) \leq a$, which means that an optimal strategy for player Max is to go to as quickly as possible to the cycle c and turn round c forever. This strategy is positional. Thus let p be any infinite path in G and let m be the maximal priority appearing infinitely often in p. This implies that in G there exists at least one simple cycle with the maximal priority m. Let b be the maximum payoff of u over all simple cycles with the maximal priority m, this quantity is well-defined since we noted that such cycles exist. It is not difficult to observe that $u(\text{colour}(p)) \leq b$, but $b \leq a$ just by the definition of a.

The proof for arenas controlled by player Min is symmetrical and Corollary 7 applies. □

References

1. A. Bouquet, O. Serre, and I. Walukiewicz. Pushdown games with unboundedness and regular conditions. In *FSTTCS*, volume 2380 of *LNCS*, pages 88–99, 2003.
2. T. Cachat, J. Duparc, and W. Thomas. Solving pushdown games with a Σ_3 winning condition. In *CSL*, volume 2471 of *LNCS*, pages 322–336, 2002.
3. K. Chatterjee, T.A. Henzinger, and M. Jurdziński. Mean-payoff parity games. In *LICS*, 2005. to appear.
4. T. Colcombet and D. Niwiński. On the positional determinacy of edge-labeled games. submitted.
5. L. de Alfaro, M. Faella, T. A. Henzinger, Rupak Majumdar, and Mariëlle Stoelinga. Model checking discounted temporal properties. In *TACAS 2004*, volume 2988 of *LNCS*, pages 77–92. Springer, 2004.
6. L. de Alfaro, T. A. Henzinger, and Rupak Majumdar. Discounting the future in systems theory. In *ICALP 2003*, volume 2719 of *LNCS*, pages 1022–1037. Springer, 2003.
7. A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. *Intern. J. of Game Theory*, 8:109–113, 1979.
8. E.A. Emerson and C. Jutla. Tree automata, μ-calculus and determinacy. In *FOCS'91*, pages 368–377. IEEE Computer Society Press, 1991.
9. H. Gimbert. Parity and exploration games on infinite graphs. In *Computer Science Logic 2004*, volume 3210 of *LNCS*, pages 56–70. Springer, 2004.
10. H. Gimbert and W. Zielonka. When can you play positionally? In *Mathematical Foundations of Computer Science 2004*, volume 3153 of *LNCS*, pages 686–697. Springer, 2004.
11. E. Grädel. Positional determinacy of infinite games. In *STACS*, volume 2996 of *LNCS*, pages 4–18. Springer, 2004.
12. A.W. Mostowski. Games with forbidden positions. Technical Report 78, Uniwersytet Gdański, Instytut Matematyki, 1991.
13. M. J. Osborne and A. Rubinstein. *A Course in Game Theory.* The MIT Press, 2002.
14. L. S. Shapley. Stochastic games. *Proceedings Nat. Acad. of Science USA*, 39:1095–1100, 1953.