

Nearest fixed points and concurrent priority games

Bruno Karelovic and Wiesław Zielonka

LIAFA, Université Paris Diderot - Paris 7
bruno.karelovic@gmail.com
zielonka@liafa.univ-paris-diderot.fr

Abstract. As it is known the values of different states in parity games (deterministic parity games, or stochastic perfect information parity games or concurrent parity games) can be expressed by formulas of μ -calculus – a fixed point calculus alternating the greatest and the least fixed points of monotone mappings on complete lattices.

In this paper we examine concurrent priority games that generalize parity games and we relate the value of such games to a new form of fixed point calculus – the nearest fixed point calculus.

1 Introduction

As it is well known parity games are closely related to μ -calculus. This fact was first observed in the context of turn based deterministic games [1, 2], next for perfect information stochastic games [3, 4] and for concurrent stochastic games [5].

Intuitively, parity games capture a situation where we meet two types of properties, desirable ones and undesirable ones. Moreover properties are ordered by a priority relation. This leads to a classification of infinite runs of a system, a run is desirable iff the property associated with the maximal priority encountered infinitely often during the run is desirable.

We can try however a finer classification of properties by quantifying them by real numbers from a closed bounded interval $I = [p_1, p_2]$ of real numbers. To this end we associate with the most preferable properties the reward p_2 and with the most undesirable the reward p_1 . However in general we can have also a whole spectrum of intermediate properties with rewards between p_1 and p_2 . As in parity games the properties can be ordered by a priority relation, the priority over properties has nothing to do with the natural reward order on I , given two properties a and b with rewards $r(a), r(b) \in I$, it is possible to have $r(a) < r(b)$ (b gives a better payoff than a) with the priority of a greater than the priority of b , i.e. property a whenever happens then it “invalidates” property b . As in parity games, given an infinite run we take the property a of maximal priority encountered infinitely often during the run and define the reward of the run as the reward associated with this property.

We obtain in this way a class of games that we call priority games. Deterministic priority games can be reduced to parity games, in particular solving

a sequence of parity games we can find the values of all states in the priority games and optimal memoryless strategies for both players. Perfect information stochastic priority games also admit optimal memoryless strategies, however we do not know if they can be reduced to parity games. In this paper we examine concurrent stochastic priority games. As it turns out the values of such games can be obtained by a new kind of μ -calculus. The traditional μ -calculus alternates the greatest and the least fixed points, the μ -calculus in this paper defines for each $r \in I = [p_1, p_2]$ the nearest fixed point of a monotone function (nearest to r). The greatest and the least fixed points are just special cases of the nearest fixed points (they are nearest to p_1 and to p_2 respectively).

Even if priority games just extend parity games we think that our approach contributes also to a better comprehension of parity games. Indeed it is notoriously difficult to comprehend the μ -calculus formulas that give solutions to parity games. This follows from the fact that it is difficult to grasp the meaning of a μ -calculus formula alternating several greatest and least fixed points.

Our approach has advantage because it provides a natural interpretation in terms of games of a partially evaluated μ -calculus formula, where fixed points are applied only to some variables while other variables are left free. In our approach each variable in the formula corresponds to a state of the game and free variables correspond to absorbing states. Then adding a new fixed point over a free variable has the following interpretation in term of games – the state corresponding to this variable changes its nature from absorbing to non-absorbing. And the usual method for approximating a new fixed point giving the value of the state turns out to be nothing else but the natural algorithm for calculating the value the new non-absorbing state. At the end, when all fixed points are applied, then this corresponds to the final situation where all states are transformed from absorbing to non-absorbing.

2 Concurrent Stochastic Priority Games

A two-player arena $G = (\mathbf{S}, \mathbf{A}, \mathbf{B}, p)$ is composed of a finite set of states $\mathbf{S} = \{1, 2, \dots, n\} \subset \mathbb{N}$ (we assume without loss of generality that \mathbf{S} is a subset of positive integers) and finite sets \mathbf{A} and \mathbf{B} of actions of players Max and Min. For each state s , $\mathbf{A}(s) \subseteq \mathbf{A}$ and $\mathbf{B}(s) \subseteq \mathbf{B}$, are the set of actions that players Max and Min have at their disposal at s . We assume that \mathbf{A} and \mathbf{B} are disjoint and $(\mathbf{A}(s))_{s \in \mathbf{S}}$, $(\mathbf{B}(s))_{s \in \mathbf{S}}$ are partitions of \mathbf{A} and \mathbf{B} .

For $s, s' \in \mathbf{S}$, $a \in \mathbf{A}(s)$, $b \in \mathbf{B}(s)$, $p(s'|s, a, b)$ is the probability to move to s' if players Max and Min execute respectively actions a and b at s .

An infinite play is played by players Max and Min. At each stage, given the current state s , the players choose simultaneously and independently actions $a \in \mathbf{A}(s)$ and $b \in \mathbf{B}(s)$ and the game moves to a new state s' with probability $p(s'|s, a, b)$. The couple (a, b) is called a joint action.

A finite history is a sequence $h = (s_1, a_1, b_1, s_2, a_2, b_2, s_3, \dots, s_n)$ alternating states and joint actions and beginning and ending with a state. The length of h

is the number of joint actions in h , in particular a history of length 0 consists of just one state and no actions. The set of finite histories is denoted H .

A strategy of player Max is a mapping $\sigma : H \rightarrow \Delta(\mathbf{A})$, where $\Delta(\mathbf{A})$ is the set of probability distributions on \mathbf{A} . We require that $\text{supp}(\sigma(h)) \subseteq \mathbf{A}(s)$, where s is the last state of h and $\text{supp}(\sigma(h)) := \{a \in \mathbf{A} \mid \sigma(h)(a) > 0\}$ is the support of the measure $\sigma(h)$.

A strategy σ is *stationary* if $\sigma(h)$ depends only on the last state of h . Thus stationary strategies of player Max can be identified with mappings from \mathbf{S} to $\Delta(\mathbf{A})$ such that $\text{supp}(\sigma(s)) \subseteq \mathbf{A}(s)$ for each $s \in \mathbf{S}$.

A strategy σ is *pure* if $\text{supp}(\sigma(h))$ is a singleton for each h . Pure stationary strategies of player Max are identified with mappings $\sigma : \mathbf{S} \rightarrow \mathbf{A}$ such that $\sigma(s) \in \mathbf{A}(s)$.

Strategies for player Min are defined in a similar way.

We write $\Sigma(\mathbf{G})$ and $\mathcal{T}(\mathbf{G})$ to denote the sets of all strategies for player Max and Min respectively.

We omit \mathbf{G} and write Σ, \mathcal{T} if \mathbf{G} is clear from the context. We use σ and τ (with subscripts or superscripts) to denote strategies of players Max and Min respectively.

An infinite history or a play is an infinite sequence $h = (s_1, a_1, b_1, s_2, a_2, b_2, s_3, a_3, b_3, \dots)$ alternating states and joint actions. The set of infinite histories is denoted H^∞ . For a finite history h by h^+ we denote the cylinder generated by h consisting of all infinite histories with prefix h . We assume that H^∞ is endowed with the σ -algebra $\mathcal{B}(H^\infty)$ generated by the set of cylinders.

Strategies σ, τ of players Max and Min and the initial state s determine in the usual way a probability measure $\mathbb{P}_s^{\sigma, \tau}$ on $(H^\infty, \mathcal{B}(H^\infty))$.

A concurrent stochastic priority game is obtained by adding to \mathbf{G} a reward mapping

$$\rho : \mathbf{S} \rightarrow I$$

associating with each state s a reward $\rho(s)$ belonging to a closed interval $I \subset \mathbb{R}$.

The payoff $u_\rho(h)$ of an infinite history $h = (s_1, a_1, b_1, s_2, a_2, b_2, s_3, a_3, b_3, \dots)$ in the priority game is defined as

$$u_\rho(h) = \rho(\limsup_n s_n). \quad (1)$$

Thus the payoff is equal to the reward of the greatest (in the usual integer order) state visited infinitely often.

The aim of player Max (player Min) is to maximize (resp. minimize) the expected payoff

$$\mathbb{E}_s^{\sigma, \tau}[u_\rho] = \int_{H^\infty} u_\rho(h) \mathbb{P}_s^{\sigma, \tau}(dh).$$

Concurrent priority games contain two well known classes of games.

- (1) Concurrent parity games [6] correspond to concurrent priority games with the reward mapping having rewards in the two element set $\{0, 1\}$ rather than arbitrary rewards in the interval I .

- (2) The second subclass of concurrent priority games is the class of Everett's recursive games [7]. Everett games are concurrent priority games such that all non-absorbing states have reward 0 (a state s is absorbing if $p(s|s, a, b) = 1$ for all joint actions (a, b)).

Thus in Everett games players receive the payoff 0 if the play remains forever in non-absorbing states, otherwise, for plays ending in an absorbing state, the payoff is equal to the reward associated with this state. Note that Everett games contain as a subclass the class of reachability games which correspond to Everett games such that all absorbing states have reward 1.

From the determinacy of Blackwell games proved by Martin [8] it follows that concurrent priority games have values, i.e. for each state s , $\sup_{\sigma} \inf_{\tau} \mathbb{E}_s^{\sigma, \tau}[u_{\rho}] = \inf_{\tau} \sup_{\sigma} \mathbb{E}_s^{\sigma, \tau}[u_{\rho}]$. (Blackwell games do not have states but the result of Martin extends immediately to games with states as shown by Maitra and Sudderth [9].)

For two subclasses of concurrent priority games mentioned earlier we have more precise results. A proof of determinacy of concurrent stochastic parity games using fixed points was given by de Alfaro and Majumdar [5]. And for Everett's games Everett proved not only that such games have values but also that both players have ε -optimal stationary strategies [7].

Notation: For an arena \mathbf{G} by $\mathbf{G}(\rho)$ we will denote the priority game obtained by endowing \mathbf{G} with the reward mapping ρ . Another notation used frequently in the paper is $\mathbf{G}(x_1, \dots, x_n)$ which denotes the priority game with n states $\{1, \dots, n\}$ having rewards x_1, \dots, x_n respectively.

3 Interval Lattice and Nearest Fixed Point

Let us recall that a complete lattice is a partially ordered set (E, \leq) such that each subset X of E has the least upper bound $\bigvee X$ and the greatest lower bound $\bigwedge X$. A mapping $f : E \rightarrow F$ from a lattice E to a lattice F is monotone if for all $x, y \in E$, $x \leq y$ implies $f(x) \leq f(y)$. The set of such monotone mappings is denoted $\text{Mon}(E, F)$. The greatest and the least element of a complete lattice will be denoted respectively \top and \perp .

Theorem 1 (Tarski [10]). *For each complete lattice (E, \leq) and a monotone mapping $f : E \rightarrow E$ the set P of fixed points of f is non-empty and is a complete lattice, in particular P has the greatest and the least element.*

In this paper we are interested in the complete lattice $I \subseteq \mathbb{R}$ of real numbers from a closed interval I and in the product lattice I^n endowed with the componentwise order. In the sequel we will fix $I = [\perp, \top]$ and \perp, \top will always denote the minimal and maximal elements of I .

Let $f : x \mapsto f(x)$ be a monotone mapping from I to itself. For a monotone mapping $f : I \rightarrow I$ and $r \in I$ we define *the nearest fixed point* $\mu_r.x.f(x)$:

- (1) if $f(r) = r$ then $\mu_r.x.f(x) = r$,

- (2) if $f(r) < r$ then f maps the interval $[\perp, r]$ into itself and by Tarski's fixed point theorem there exists the greatest fixed-point of f in $[\perp, r]$ and $\mu_r x.f(x)$ denotes this fixed point (in other words, $\mu_r x.f(x)$ is the greatest fixed point of f in $[\perp, r]$),
- (3) if $f(r) > r$ then f maps the interval $[r, \top]$ into itself and by Tarski's fixed point theorem, there exists the least fixed-point of f in $[r, \top]$ and $\mu_r x.f(x)$ denotes this fixed point (in other words, $\mu_r x.f(x)$ is the least fixed point of f in $[r, \top]$).

A mapping $f : I \rightarrow I$ is *nonexpansive* if for all $x, y \in I$, $|f(x) - f(y)| \leq |x - y|$.

It is known that in general complete lattices a transfinite induction can be necessary in order to calculate the least and the greatest fixed points of monotone mappings. The following lemma shows that for monotone nonexpansive mappings from I to I the situation is much simpler:

Lemma 2. *Let $f : I \rightarrow I$ be monotone and nonexpansive.*

- (i) *Let $r \in I$ and let $(r_n)_{n \geq 0}$ be the sequence of real numbers such that $r_0 = r$ and $r_{n+1} = f(r_n)$. The sequence (r_n) is monotone and it converges (in the usual sense of convergence in \mathbb{R} with the euclidean metric) to $\mu_r x.f(x)$.*
- (ii) *The set of fixed points of f is a closed subinterval of I .*
- (iii) *Let e_1, e_2 be respectively the least and the greatest fixed points of f . Then $\mu_r x.f(x) = e_1$ if $r < e_1$, $\mu_r x.f(x) = r$ if $r \in [e_1, e_2]$ and $\mu_r x.f(x) = e_2$ if $r > e_2$.*

Note that (iii) shows that $\mu_r x.f(x)$ is indeed the fixed point which is closest (in the sense of the euclidean distance) to r .

Proof. (i) If $f(r) \leq r$ then by monotonicity $f^{n+1}(r) \leq f^n(r)$ for all $n \geq 0$, where \leq is either \leq or \geq . But a bounded monotone sequence of real numbers converges to some $r_\infty \in I$. Since f is nonexpansive $|f(r_\infty) - f^{n+1}(r)| \leq |r_\infty - f^n(r)|$. The left-hand side of this inequality converges to $|f(r_\infty) - r_\infty|$ while the right-hand side converges to 0.

(ii) Suppose that $x < y$ are two fixed points of f and $z \in [x, y]$. Then $x = f(x) \leq f(z) \leq f(y) = y$ and $|x - f(z)| = |f(x) - f(z)| \leq |x - z|$. Similarly $|f(z) - y| \leq |z - y|$. However both these inequalities can hold simultaneously only if $f(z) = z$. Thus if e_1 and e_2 are the least and the greatest fixed points of f then all elements of $[e_1, e_2]$ are fixed points of f .

(iii) Direct consequence of (ii).

3.1 Nested Nearest Fixed Points

For any set D we endow the set I^D of mappings from D to I with the order relation: for $f, g \in I^D$, $f \leq g$ if $f(x) \leq g(x)$ for all $x \in D$. Then (I^D, \leq) is also a complete lattice, for a set $F \subseteq I^D$, ΘF is a mapping f such that $f(x) = \Theta_{h \in F} h(x)$, $\Theta \in \{\bigvee, \bigwedge\}$. Note that the product lattice I^n can be seen as a mapping from $\{1, \dots, n\}$ into I , i.e. is covered by this definition.

The lattice I^I contains the lattice $\text{Mon}(I, I)$ of all monotone mappings from I to I . Note that for any set $F \subseteq \text{Mon}(I, I)$, ΘF calculated in the lattice I^I or in the lattice $\text{Mon}(I, I)$ gives the same result, $\Theta \in \{\wedge, \vee\}$.

Lemma 3. *Let $f, g \in \text{Mon}(I, I)$ and $r \in I$. If $f \leq g$ then for each $r \in I$, $\mu_r x.f(x) \leq \mu_r x.g(x)$.*

Proof. If $f(r) \leq r \leq g(r)$ then $\mu_r x.f(x) \leq f(r) \leq r \leq g(r) \leq \mu_r x.g(x)$.

If $r < f(r) \leq g(r)$ then f and g are monotone mappings from $[r, \top]$ to $[r, \top]$ and then $\mu_r x.f$ and $\mu_r x.g$ are the least fixed points of f and g considered as mappings from the lattice $[r, \top]$ to $[r, \top]$. However, if $f \leq g$, where f and g monotone, then the least (greatest) fixed point of f is \leq than the least (greatest) fixed point of g (Proposition 1.2.18 in [11]).

The case $f(r) \leq g(r) < r$ is symmetric to the previous one. \square

We endow \mathbb{R}^n with the norm $\|(x_1, \dots, x_n)\| = \max_i |x_i|$.

We say that a mapping $f \in \text{Mon}(I^n, I^m)$ is monotone nonexpansive if for all $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in I^n$, $\|f(\mathbf{x}) - f(\mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|$.

By $\text{Mon}_e(I^n, I^m)$ we denote the set of monotone nonexpansive mappings from I^n to I^m .

Given $f \in \text{Mon}(I^n, I)$ by

$$\mu_r z_i.f(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_n) \quad (2)$$

we denote the mapping from I^{n-1} to I which maps $(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n) \in I^{n-1}$ to the nearest fixed point of the mapping $z_i \mapsto f(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_n)$.

Lemma 4. *Let us fix $r \in I$.*

If $f \in \text{Mon}(I^n, I)$ then the mapping (2) belongs to $\text{Mon}(I^{n-1}, I)$.

If $f \in \text{Mon}_e(I^n, I)$ then the mapping (2) belongs to $\text{Mon}_e(I^{n-1}, I)$.

Proof. The first assertion follows immediately from Lemma 3.

For $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in I^{n-1}$ we define inductively a sequence of mappings:

$$g^0(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = f(x_1, \dots, x_{i-1}, r, x_{i+1}, \dots, x_n)$$

$$g^{k+1}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = f(x_1, \dots, x_{i-1}, g^k(x_1, \dots, x_i, x_{i+1}, \dots, x_n), x_{i+1}, \dots, x_n).$$

We shall prove that for all k and all $\mathbf{x} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n)$ in I^{n-1} ,

$$|g^k(\mathbf{x}) - g^k(\mathbf{y})| \leq \|\mathbf{x} - \mathbf{y}\|. \quad (3)$$

For $k = 0$ this follows directly from the fact that f is monotone nonexpansive. Suppose that (3) holds for k . Then

$$\begin{aligned} & |g^{k+1}(\mathbf{x}) - g^{k+1}(\mathbf{y})| = \\ & |f(x_1, \dots, x_{i-1}, g^k(\mathbf{x}), x_{i+1}, \dots, x_n) - f(y_1, \dots, y_{i-1}, g^k(\mathbf{y}), y_{i+1}, \dots, y_n)| \leq \\ & \max\{|x_1 - y_1|, \dots, |x_{i-1} - y_{i-1}|, |g^k(\mathbf{x}) - g^k(\mathbf{y})|, |x_{i+1} - y_{i+1}|, \dots, |x_n - y_n|\} \leq \\ & \max\{|x_1 - y_1|, \dots, |x_{i-1} - y_{i-1}|, \|\mathbf{x} - \mathbf{y}\|, |x_{i+1} - y_{i+1}|, \dots, |x_n - y_n|\} \leq \|\mathbf{x} - \mathbf{y}\|. \end{aligned}$$

Now it suffices to note that, by Lemma 2, $g^k(\mathbf{x})$ and $g^k(\mathbf{y})$ converge respectively to $\mu_r x_i \cdot f(x_1, \dots, x_i, \dots, x_n)$ and to $\mu_r y_i \cdot f(y_1, \dots, y_i, \dots, y_n)$ when $k \uparrow \infty$. \square

Let $\mathbf{f} \in \text{Mon}_e(I^n, I^n)$. Thus \mathbf{f} is a vector of mappings $\mathbf{f} = (f_1, \dots, f_n)$ where, for each i , $f_i \in \text{Mon}_e(I^n, I)$. Let $(r_1, \dots, r_n) \in I^n$.

For each k , $1 \leq k \leq n$, we define a monotone nonexpansive mapping $\mathbf{F}^{(k)} : I^{n-k} \rightarrow I^k$:

$$I^{n-k} \ni (x_{k+1}, \dots, x_n) \mapsto \mathbf{F}^{(k)}(x_{k+1}, \dots, x_n) \in I^k.$$

(for $k = n$, $\mathbf{F}^{(n)}$ will be just a constant from I^n not depending on any variable).

Since $\mathbf{F}^{(k)}$ is a mapping into I^k , it is composed of k mappings into I , $\mathbf{F}^{(k)} = (F_1^{(k)}, \dots, F_k^{(k)})$.

For $k = 1$, $\mathbf{F}^{(1)}$ is mapping into I and we will identify it with $F_1^{(1)}$.

We define $\mathbf{F}^{(k)}$ by induction. For $k = 1$,

$$\mathbf{F}^{(1)}(x_2, \dots, x_n) = F_1^{(1)}(x_2, \dots, x_n) = \mu_{x_1} r_1 \cdot f_1(x_1, x_2, \dots, x_n).$$

Suppose that $\mathbf{F}^{(k-1)}(x_k, \dots, x_n) = (F_1^{(k-1)}(x_k, \dots, x_n), \dots, F_{k-1}^{(k-1)}(x_k, \dots, x_n))$ is already defined.

Intuitively, given $\mathbf{F}^{(k-1)}$ as above to obtain $\mathbf{F}^{(k)}$ we should eliminate the variable x_k . To this end we use the k th component mapping f_k of \mathbf{f} .

First we define

$$\begin{aligned} F_k^{(k)}(x_{k+1}, \dots, x_n) = \\ \mu_{r_k} x_k \cdot f_k(F_1^{(k-1)}(x_k, \dots, x_n), \dots, F_{k-1}^{(k-1)}(x_k, \dots, x_n), x_k, x_{k+1}, \dots, x_n), \end{aligned} \quad (4)$$

and subsequently we put

$$F_i^{(k)}(x_{k+1}, \dots, x_n) = F_i^{(k-1)}(F_k^{(k)}(x_{k+1}, \dots, x_n), x_{k+1}, \dots, x_n), \quad \text{for } 1 \leq i < k. \quad (5)$$

By Lemma 4 and since the composition of monotone nonexpansive mappings is monotone nonexpansive we can see that all mappings $\mathbf{F}^{(k)}$ are monotone nonexpansive.

We shall write

$$\mu_{r_k} x_k \dots \mu_{r_1} x_1 \cdot \mathbf{f}(x_1, \dots, x_n)$$

to denote the mapping $\mathbf{F}^{(k)}$ defined above and we call it the k -th nested fixed point of \mathbf{f} . For $k = n$ we will speak about the nested fixed point without mentioning k .

4 Value of the Concurrent Priority Game as the Nested Nearest Fixed Point

4.1 Auxiliary One-Shot Game

In this section we define auxiliary matrix games.

Let $\mathbf{x} = (x_1, \dots, x_n) \in I^n$ and let $\mathbf{G}(\mathbf{x})$ be a priority game with n states.

A one shot game $\Gamma_k(x_1, \dots, x_n)$ is the game played on \mathbf{G} in the following way. The game starts at state k , players Max and Min choose independently and simultaneously actions $a \in \mathbf{A}(k)$ and $b \in \mathbf{B}(k)$. Suppose that the next state is m . Then player Max receives from player Min the payoff x_m and the game $\Gamma_k(x_1, \dots, x_n)$ ends.

Note that $\Gamma_k(x_1, \dots, x_n)$ can be seen a zero-sum matrix game where the payoff obtained by player Max from player Min when players play actions a, b respectively is equal to $\sum_{m \in \mathbf{S}} x_m \cdot p(m|k, a, b)$. The value of the game $\Gamma_k(x_1, \dots, x_n)$ will be denoted by

$$\Phi_k(x_1, \dots, x_n) := \text{val}(\Gamma_k(x_1, \dots, x_n)). \quad (6)$$

We will be interested in $\Phi_k(x_1, \dots, x_n)$ as a function of the reward vector $\mathbf{x} = (x_1, \dots, x_n)$.

Since all entries in the matrix game $\Gamma_k(x_1, \dots, x_n)$ belong to I , $\text{val}(\Gamma_k(x_1, \dots, x_n)) \in I$, i.e. Φ_k is a mapping from I^n into I .

The following well known properties of matrix games are essential (see for example [12]):

Theorem 5. *Let M_1 and M_2 be two matrix games of the same size. Then*

- *If $M_1 \leq M_2$ (where the inequality holds componentwise) then $\text{val}[M_1] \leq \text{val}[M_2]$.*
- *$|\text{val}[M_1] - \text{val}[M_2]| \leq \|M_1 - M_2\|$, where $\|M\| = \max_{i,j} |M(i, j)|$.*

From Theorem 5 it follows that

Proposition 6. *The mapping Φ_k defined in (6) is monotone and nonexpansive.*

4.2 Priority Games with One Non-absorbing State

In this section we will study concurrent priority games with one non-absorbing state. Let us recall that a state i is absorbing if for all $(a, b) \in \mathbf{A}(i) \times \mathbf{B}(i)$, $p(i|i, a, b) = 1$.

We shall write $\mathbf{G}_k(x_1, \dots, x_n)$ to denote a priority game \mathbf{G} with n states having rewards x_1, \dots, x_n and such that all states, except state k , are absorbing. We shall call such a game absorbing. A game starting in an absorbing state i , $i \neq k$, is trivial, the game remains forever in i and the payoff is equal to the reward x_i associated with state i . For plays starting in the non-absorbing state k either at some moment we hit an absorbing state i and the payoff obtained for

such plays is x_i (and it is irrelevant what players play once an absorbing state is attained) or we remain forever in k and the payoff for such a play is x_k . Such games are equivalent to Everett games with one non-absorbing state.

Thus only the value and players' strategies in the non-absorbing state k are of interest in $\mathbf{G}_k(x_1, \dots, x_n)$.

In the sequel we will use the following notation. For $\mathbf{x} = (x_1, \dots, x_n) \in I^n$ and $e \in I$ we write (\mathbf{x}_{-k}, e) to denote the element $(x_1, \dots, x_{k-1}, e, x_{k+1}, \dots, x_n)$ of I^n .

Moreover if σ, τ are strategies of players Max and Min in the one shot game Γ (i.e. for each state $s \in \mathbf{S}$, $\sigma(s) \in \Delta(\mathbf{A}(s))$ and $\tau(s) \in \Delta(\mathbf{B}(s))$ are mappings from states to distributions over actions) then σ^∞ and τ^∞ will denote the stationary strategies in the priority game \mathbf{G} such that at each stage players select actions independently of the past history with probabilities given by σ and τ .

Lemma 7. *Let $\mathbf{G}_k(\mathbf{x})$ be an absorbing priority game and $r \in I$. Then*

$$\text{val}(\mathbf{G}_k(\mathbf{x}_{-k}, r)) = \mu_r x_k \cdot \Phi_k(x_1, \dots, x_k, \dots, x_n).$$

- (i) *If $\mu_r x_k \cdot \Phi_k(\mathbf{x}) \geq r$ then player Min has an optimal stationary strategy while, for each $\varepsilon > 0$, player Max has an ε -optimal stationary strategy.*
- (ii) *If $\mu_r x_k \cdot \Phi_k(\mathbf{x}) \leq r$ then player Max has an optimal stationary strategy in $\mathbf{G}_k(\mathbf{x}_{-k}, r)$ while, for each $\varepsilon > 0$, player Min has an ε -optimal stationary strategy.*

Note that from this lemma it follows that if $\mu_r x_k \cdot \Phi_k(\mathbf{x}) = r$ then both players have optimal stationary strategies in the absorbing game $\mathbf{G}_k(\mathbf{x}_{-k}, r)$.

4.3 General Priority Games

Let \mathbf{G} be a priority game.

Note that

$$(x_1, \dots, x_n) \mapsto \Phi(x_1, \dots, x_n) = (\Phi_1(x_1, \dots, x_n), \dots, \Phi_n(x_1, \dots, x_n))$$

where Φ_i defined in (6) are monotone and nonexpansive mappings from I^n to I .

By $\mathbf{G}_{\leq k}$ we will denote the priority game obtained from \mathbf{G} by transforming all states i , such that $i > k$ into absorbing states. On the other hand, all states $j \leq k$ have the same available actions and transition probabilities as they have in \mathbf{G} .

Of course, the value of each absorbing state j of the game

$\mathbf{G}_{\leq k}(r_1, \dots, r_k, x_{k+1}, \dots, x_n)$, $k < j \leq n$, is x_j thus only the values of non-absorbing states $1, \dots, k$ are of interest.

It turns out that these values are obtained as nested fixed points:

Theorem 8. *Let $(r_1, \dots, r_k) \in I^k$. Then the nested fixed point*

$$(F_1^{(k)}(x_{k+1}, \dots, x_n), \dots, F_k^{(k)}(x_{k+1}, \dots, x_n)) := \mu_{r_k} x_k \cdot \dots \cdot \mu_{r_1} x_1 \cdot \Phi(\mathbf{x})$$

is the vector of values of non-absorbing states $(1, \dots, k)$ of the game

$\mathbf{G}_{\leq k}(r_1, \dots, r_k, x_{k+1}, \dots, x_n)$.

As it is known for parity games, which form a special subclass of priority games, the winning regions (in the deterministic case [2]) or the values (for concurrent stochastic parity games [5]) can be described by an appropriate formulas of μ -calculus – a fixed point calculus over an appropriate complete lattice where we alternate the greatest and the least fixed points. From this point of view Theorem 8 looks just as an extension of known results to a wider framework of priority games. However there is one ingredient of Theorem 8 that seems to be new.

It is notoriously difficult to comprehend a μ -calculus formula alternating several greatest and least fixed point.

Theorem 8 provides a natural interpretation in the term of games of a formula where only some initial fixed points are applied.

Let

$$(v_1, \dots, v_k) = \mu_{r_k} x_k \cdot \mu_{r_{k-1}} x_{k-1} \cdot \dots \cdot \mu_{r_1} x_1 \cdot \Phi(x_1, \dots, x_k, r_{k+1}, \dots, r_n) \quad (7)$$

Then (v_1, \dots, v_k) are the values of states $1, \dots, k$ in the priority game

$$\mathbf{G}_{\leq k}(r_1, \dots, r_k, r_{k+1}, \dots, r_n) \quad (8)$$

which differs from the original priority game $\mathbf{G}(r_1, \dots, r_k, r_{k+1}, \dots, r_n)$ in that the states $k+1, k+2, \dots, n$ are absorbing in the game (8).

Now when we add another fixed point to (7) to obtain

$$(v'_1, \dots, v'_k, v'_{k+1}) = \mu_{r_{k+1}} x_{k+1} \cdot \mu_{r_k} x_k \cdot \mu_{r_{k-1}} x_{k-1} \cdot \dots \cdot \mu_{r_1} x_1 \cdot \Phi(x_1, \dots, x_k, x_{k+1}, r_{k+2}, \dots, r_n)$$

then this can be interpreted as an operation transforming state $k+1$ from a non-absorbing in the game (8) into a non-absorbing in the game

$$\mathbf{G}_{\leq k+1}(r_1, \dots, r_k, r_{k+1}, \dots, r_n).$$

5 Algorithmic Issues

One can wonder if the recursive formulas for the nested nearest point cannot be used to approximate the values of fixed points, i.e. the values of the states in the priority game. Unfortunately in general this seems to be difficult, if at levels $1, \dots, k$ we stop iterations before attaining the fixed points then the resulting errors can even change the direction of iterations at level $k+1$. Moreover it is difficult to see when we can stop iterations (there is no criterion to estimate the distance between the value obtained at some iteration and the limit fixed point).

However there is one case when the recursive formulas developed in this paper can be used to solve the priority game, this is the case of perfect information stochastic priority games where for each state only one of the two players chooses actions to play (the other player can be seen as having only one action at this state). First we have the following counterpart of Lemma 7:

Lemma 9. *Let $\mathbf{G}_k(\mathbf{x})$ be a priority game with the unique non-absorbing state k , $r \in I$. Let k be controlled by one player (either Max or Min) who chooses an action to play and the probability distribution over next states depends uniquely upon the action chosen by the controlling player. Then the value*

$$\text{val}(\mathbf{G}_k(\mathbf{x}_{-k}, r)) = \mu_r x_k \cdot \Phi_k(x_1, \dots, x_k, \dots, x_n).$$

of state k can be calculated in polynomial time and the controlling player has an optimal pure strategy.

Proof. For each action a of the player controlling k we have the following formula for the expected reward after playing a once:

$$\mathbb{E}_k^a[R] = \mathbb{E}_k^a[R|A] \cdot p(A|k, a) + r \cdot p(k|k, a)$$

where R is the expected reward after playing a once, A is the event that the next state is absorbing, $p(A|k, a) = \sum_{j \in A} p(j|k, a)$ is the probability that the next state is absorbing when a is executed, $r \cdot p(k|k, a)$ is the probability that we remain in k when the player plays a in k . From this formula we can calculate $\mathbb{E}_k^a[R|A]$ i.e. the expected reward under condition that the next state is absorbing.

The expected payoff obtained when we play the strategy a^∞ (play only a as long as the current state is k) is equal

$$\mathbb{E}_k^{a^\infty}[u] = \begin{cases} \mathbb{E}_k^a[R|A] & \text{if } p(A|k, a) > 0, \\ r & \text{otherwise.} \end{cases} \quad (9)$$

Thus if k is controlled by player Max (Min) then he should play action a that maximizes (minimizes) (9) as long as we are in k . \square

Thus Lemma 9 shows how to solve one state perfect information stochastic priority game. To solve a perfect information priority game with any number of non-absorbing states we use the induction. However instead of value iteration algorithm (which can be non-terminating) we use the strategy iteration (which always terminate as the number of pure strategies is finite). In fact this algorithm just tries to accelerate and optimize the procedure calculating the nested nearest fixed points.

Algorithm 1 on page 12 implements a recursive procedure `SolveGame`($k, (r_1, \dots, r_k, x_{k+1}, \dots, x_n)$) that calculates the vector (a_1, \dots, a_k) of actions played in non-absorbing states $1, \dots, k$ by optimal pure stationary strategies in the perfect information stochastic priority game $\mathbf{G}_{\leq k}(r_1, \dots, r_k, x_{k+1}, \dots, x_n)$.

References

1. Emerson, E., Jutla, C.: Tree automata, μ -calculus and determinacy. In: FOCS'91, IEEE Computer Society Press (1991) 368–377
2. Walukiewicz, I.: Monadic second-order logic on tree-like structures. Theoretical Computer Science **275** (2002) 311–346

3. McIver, A., Morgan, C.: Games, probability and the quantitative μ -calculus qmu. In: Proc. LPAR. Volume 2514 of LNAI., Springer (2002) 292–310 full version arxiv.org/abs/cs.LO/0309024.
4. McIver, A., Morgan, C.: A novel stochastic game via the quantitative mu-calculus. In Cerone, A., Wiklicky, H., eds.: Proc. of the Third Workshop on Quantitative Aspects of Programming Languages (QAPL 2005). Volume 153, Issue 2 of ENTCS., Elsevier (2005) 195–212
5. de Alfaro, L., Majumdar, R.: Quantitative solution to omega-regular games. Journal of Computer and System Sciences **68** (2004) 374–397
6. Chatterjee, K., de Alfaro, L., Henzinger, T.: Qualitative concurrent parity games. ACM Transactions on Computational Logic **12** (2011) 28:1–28:51
7. Everett, H.: Recursive games. In: Contributions to the Theory of Games, Vol. III, Princeton University Press (1957) 47–78
8. Martin, D.: The determinacy of Blackwell games. Journal of Symbolic Logic **63**(4) (1998) 1565–1581
9. Maitra, A., Sudderth, W.: Stochastic games with Borel payoffs. In Neyman, A., Sorin, S., eds.: Stochastic Games and Applications. Volume 570 of NATO Science Series C, Mathematical and Physical Sciences. Kluwer Academic Publishers (2004) 367–373
10. Tarski, A.: A lattice-theoretical fixpoint theorem and its applications. Pacific J. Math. **5** (1955) 285–309
11. Arnold, A., Niwiński, D.: Rudiments of μ -calculus. Volume 146 of Studies in Logic and the Foundations of Mathematics. Elsevier (2001)
12. Parthasarathy, T., Raghavan, T.: Some Topics in Two-Person Games. Elsevier (1971)

Algorithm 1: Calculate optimal pure stationary strategies in a perfect information stochastic priority game with k non-absorbing states. We assume that $\text{OneState}(k, c_1, \dots, c_n)$ is the procedure described in Lemma 9 returning the optimal action for the game with one non-absorbing state k and reward vector (c_1, \dots, c_n) and $\text{value}(k, (a_1, \dots, a_k))$ is the value of state k when players play actions (a_1, \dots, a_k) in states $1, \dots, k$ respectively.

```

1 SolveGame( $k, (r_1, \dots, r_k, x_{k+1}, \dots, x_n)$ ) Result: a vector  $(a_1, \dots, a_k)$  of
   actions,  $a_i$  action played in state  $i$  by the optimal strategy of the player
   controlling  $i$ .
2 if  $k = 1$  then
3   | return  $\text{OneState}(1, r_1, x_2, \dots, x_n)$ ;
4 end
5  $w \leftarrow r_k$ ;
6 while true do
7   |  $(a_1, \dots, a_{k-1}) \leftarrow \text{SolveGame}(k-1, r_1, \dots, r_{k-1}, w, x_{k+1}, \dots, x_n)$ ;
8   |  $a_k \leftarrow \text{OneState}(k, a_1, \dots, a_{k-1}, w, x_{k+1}, \dots, x_n)$ ;
9   |  $z \leftarrow \text{value}(k, (a_1, \dots, a_k))$ ;
10  | if  $z = w$  then
11  |   | return  $(a_1, \dots, a_k)$ ;
12  | end
13  |  $w \leftarrow z$ ;
14 end

```