Master's Thesis

Infinite Games

Miheer Dewaskar

June 2016

Chennai Mathematical Institute

Abstract

This is a literature survey on results for Parity, Mean payoff and Discounted payoff games. All of these are two player, perfect information, infinite duration games. We introduce these games, demonstrate algorithms to solve each of them, and the connections between them. These games have positional optimal strategies and their decision problems are in $\mathsf{UP} \cap \mathsf{coUP}$. However, whether any of them have a polynomial time algorithm or not is still an open problem.

Acknowledgements

I would like to thank my supervisor, Prof. B Srivathsan, for introducing me to infinite Automata and Games, and for all the helpful discussions that we have had.

I also wish to thank Prof. Madhavan Mukund, Prof. S P Suresh, Prof. Mandayam Srivas and Prof. Narayan Kumar for the exciting courses in Logic, Automata Theory and Verification that I took under them.

Finally, I would like to thank Prof. T Parthasarathy who kindled my interest in Game Theory and encouraged me.

Contents

1 Introduction

Two player (adversarial, perfect information) games are useful in Logic. For example the truth value of a formula $\phi \equiv \exists x \, \forall y \, \exists z \, (y \cdot z = x)$ on a structure $(\mathbb{N}, \cdot)_=$ can be described by the following two player game. The players are called $P_\exists, P_\forall$ respectively. First $P_\exists$ chooses an $x \in \mathbb{Z}$, then $P_\forall$ chooses an $y \in \mathbb{Z}$ and finally $P_\exists$ chooses an $z \in \mathbb{Z}$. $P_\exists$ wins the resulting play when $y \cdot z = x$, otherwise $P_\forall$ wins. Observe that ϕ is true exactly when $P_\exists$ has a winning strategy in this game. This shows how Games are related to quantifier alternation. Similarly, the acceptance conditions for Alternating automata can also be described a game.

This analysis is very useful in the infinite case. Büchi [18] used automata on infinite words to give a decision procedure for S1S – the MSO (Monadic second order) theory of natural numbers $(\mathbb{N}, S)_=$ with successor function $S(x) := x + 1$ and equality. MSO means that apart from first order quantification $\exists x$, terms like $x \in A$ and quantification $\exists A$ are also allowed over set variables A. Later, Rabin [17] used automata on infinite trees to give a similar decision procedure for S2S – the MSO theory on the infinite binary tree $(\{0,1\}^*, S_0, S_1)_=$ where $S_0(x) = x0$ and $S_1(x) = x1$ are left and right successors. The decidability of S2S is a powerful result. Many interesting logics can be decided by interpreting into S2S (see [17],[2, Chap 7]). Rabin's original proof is complicated. The combinatorics of the hardest part – the complementation lemma for Tree automata can be neatly simplified using infinite (duration) games [22, Chap 8]. The acceptance condition for the Tree automata is interpreted as a Parity game (introduced in Chapter 3); this allows to use the results on Parity games for the emptiness and complementation problem of the Tree automata.

Another important application of Infinite games is the Synthesis problem for Reactive systems. Reactive systems are computer systems that continuously interact with the environment, e.g., Lift Controller and an Operating System. An execution of such a system can be captured by an infinite string $(\mathcal{I} \times \mathcal{O})^\omega$ of input/output pairs over the input alphabet $\mathcal{I}$ and output alphabet $\mathcal{O}$. The desired behaviour of such a system is given by a specification $Spec \subseteq (\mathcal{I} \times \mathcal{O})^\omega$. Given $Spec$, the Synthesis problem is to find an $f : \mathcal{I}^* \mapsto \mathcal{O}$ (if it exists) so that for any input sequence $i_0 i_1 \ldots \in \mathcal{I}^\omega$ the input/output sequence $(i_0, f(i_0))(i_1, f(i_0 i_1))(i_2, f(i_0 i_1 i_2)) \ldots \in Spec$. Intuitively we want an algorithm that when presented with a (finite description of) $Spec$ finds an implementation f satisfying $Spec$ or concludes that there is none.

When $Spec$ is expressed in S1S (or LTL), [3] showed that Synthesis problem is decidable. The idea is to express the Synthesis problem as an infinite game between the Environment and the System (designer). The turns of Environment and System alternate during the game. Initially the Environment picks $i_1 \in \mathcal{I}$, then the System picks $o_1 \in \mathcal{O}$. Then the Environment picks $i_2 \in \mathcal{I}$ and the System picks $o_2 \in \mathcal{O}$, and the game thus continues. If $(i_1, o_1)(i_2, o_2) \ldots \in \mathit{Spec}$ then the play is winning for System else the Environment wins. Observe that implementations (f) satisfying *Spec* are exactly the winning strategies for the System in this game. Hence this reduces the Synthesis problem to finding the winning strategy in an infinite game. Since *Spec* is expressed in S1S, *Spec* is an ω-regular set. For ω-regular winning conditions on finite graphs (as is the case here) there are algorithms to compute the winning strategies.

Hence algorithms to decide the winner and find the winning strategies in (finitely presented) infinite games are of importance. We will keep this mind while discussing infinite games. Chapter 2 will provide a general formalism for infinite games on graphs. In Chapter 3 we discuss Parity games. Parity games may not be the most intuitive games to start with (see [22, Chap 2] for Reachability and Büchi games), but they are powerful enough to express ω-regular winning conditions (via deterministic parity automata) and yet have positional winning strategies. The problem of deciding the winner of Parity games is in $\mathsf{NP} \cap \mathsf{coNP}$ (hence unlikely to be NP-complete), but it is not known whether (or not) a polynomial time algorithm exists. An efficient algorithm for deciding the winner and winning strategies for Parity games will help improve the runtime for the algorithms mentioned above, and for many others which rely on infinite games.

Chapter 4 and Chapter 5 discuss Mean and Discounted payoff games. Although these games are no longer Win-Lose games, they are related to Parity games. Algorithms for such games can be useful for more general Controller-Synthesis problems with Quantitative objectives.

2 Graph games

We will look at a class of two player games on graphs. The two players will be called Player 0 and Player 1 (abbreviated as P_0, P_1). Although we will only look at specific games later, all of them fit into a general framework which will be discussed now.

2.1 Formalism

The game is played on a directed graph $G = (V_0, V_1, E)$ whose vertices are $V = V_0 \sqcup V_1$ and edges are $E \subseteq V \times V$. The vertices V are partitioned as (V_0, V_1) – those belonging to P0 and P1 respectively. We will always assume that G has no dead ends ($\forall v \in V\ \exists w\, (v, w) \in E$). For $n \geq 0$, let $\mathcal{P}^n$ denote the set of paths in G of length n (in particular $\mathcal{P}^0 = V$, $\mathcal{P}^1 = E$). $\mathcal{P}^*$ will denote the set of all finite paths in G, and by $\mathcal{P}^\omega$ the set of all infinite paths. $N(u) = \{v \mid (u, v) \in E\}$ is the set of outgoing neighbours of u.

The game starting from $v_0 \in V$ is played as follows

- A token is initially placed in v_0
- At any stage if the token is in a vertex $v \in V_i$, then P_i has to move the token to a vertex $w \in N(v)$

A concrete realization of this play will be an infinite path $\pi \in \mathcal{P}^\omega$

$$\pi = v_0 v_1 v_2 \dots$$

Where

- v_0 was the vertex where the play started
- If $v_i \in V_j$, then at $i + 1^{\text{th}}$ stage P_j decided to move the token to $v_{i+1} \in N(v_i)$

After this infinite path π is played, P_0 pays $f(\pi)$ units of money to P_1. Where $f : \mathcal{P}^\omega \to \mathbb{R}$ is the payoff function. When $f(\pi) < 0$ this is interpreted as P_1 paying $|f(\pi)|$ units to P_0. P_0's objective is to minimize $f(\pi)$ while P_1's objective is to maximize it.

This game will be denoted by $\mathcal{G} = (G, f)$. And in particular, the game starting at v_0 by $(\mathcal{G}, v_0)$.

2.1.1 Strategy

A strategy for P_i is a function $\sigma : \mathcal{P}^* \cap V^*V_i \mapsto V$, which assigns to each possible finite path αu ending in $u \in V_i$ a neighbour $\sigma(\alpha u) \in N(u)$.

Intuitively this is a recipe for P_i to make his moves in the game. If at the n^{th} stage it is P_i's turn, and the token has been moved to $v_0v_1 \ldots v_{n-1}$ so far (with $v_{n-1} \in V_i$), the strategy recommends playing $v_n = \sigma(v_0v_1 \ldots v_{n-1})$

An infinite path $\pi = v_0v_1 \ldots$ is said to conform with a strategy σ of P_i if whenever $v_j \in V_i$, $v_{j+1} = \sigma(v_0v_1 \ldots v_j)$.

Consider strategies (σ, τ) for P_0 and P_1 respectively. Starting at v_0 there is a unique path $\pi^{v_0}_{\sigma\tau} = v_0v_1 \ldots$ which conforms with both σ and τ, given by

$$v_{i+1} = \begin{cases} \sigma(v_0v_1 \ldots v_i) & \text{if } v_i \in V_0 \\ \tau(v_0v_1 \ldots v_i) & \text{if } v_i \in V_1 \end{cases}$$

Let $f^v(\sigma, \tau)$ denote $f(\pi^v_{\sigma\tau})$.

Finite memory strategies

From a computational perspective it is important to have a finite implementation of a strategy (which is generally an infinite object). Finite memory strategies are a subclass for which this is possible. A finite memory strategy for P_i is a tuple (M, δ, g, m_0), where M is a finite set (also called as "the memory"), and $m_0 \in M$.

$$\delta : M \times V \to M$$

is the update function for the memory and

$$g : M \times V_i \to V$$

is the action function (with $\forall u\ g(., u) \in N(u)$).

To implement this strategy start with m_0. At the n^{th} move if $v_{n-1} \in V_i$ then play $v_n = g(m_{n-1}, v_{n-1})$ (otherwise the opponent decides v_n) and set $m_n = \delta(m_{n-1}, v_n)$.

Positional strategy

An important special case of finite memory strategies is when M is singleton. These are called positional or memoryless strategies. In this case

$$g : V_i \to V$$

and the resulting strategy σ is just

$$\sigma(v_0 v_1 \dots v_n) = g(v_n)$$

which only depends on the current state the token is in.

Let ρ be a positional strategy for P_i in G, then G_ρ will denote the graph obtained from G by restricting the outgoing edges of every $u \in V_i$ to the unique edge given by ρ. Moreover if $\bar{\rho}$ is a strategy for P_{1-i} then $G_{\rho\bar{\rho}}$ is defined as $(G_\rho)_{\bar{\rho}}$.

An infinite path π is called ultimately periodic if for some $k \geq 0$

$$\pi = v_0 v_1 \dots v_{k-1}(v_k v_{k+1} \dots v_r)^\omega \tag{2.1}$$

Denote $\text{Prefix}(\pi) = v_0 v_1 \dots v_{k-1}$ and $\text{Cycle}(\pi) = v_k v_{k+1} \dots v_r$.

Let (σ, τ) be positional strategies for P_0 and P_1. Starting any $v_0 \in V$, the path $\pi^{v_0}_{\sigma\tau}$ is ultimately periodic with all v_i's distinct in (2.1). $\text{Cycle}(\pi^{v_0}_{\sigma\tau})$ is the unique cycle reachable from v_0 in $G_{\sigma\tau}$, and $\text{Prefix}(\pi^{v_0}_{\sigma\tau})$ is the path leading to that cycle.

Denote by $\mathcal{S}_i$ the set of all strategies of P_i. Often σ, τ (and their variants) will be used to denote the strategies for P_0 and P_1 respectively.

2.2 Optimal play

How should P_i play so as to best fulfil his objective of minimizing/maximizing the payoff? Generally this depends on the opponents play - so it is not exactly a standard optimization problem. Instead (whenever it exists) the following concept for optimal play is used.

2.2.1 Minimax equilibrium

For the game $(\mathcal{G}, v)$, if there is an $\eta_v \in \mathbb{R}$ and strategies σ^*, τ^* for player 0 and 1 respectively so that

$$\begin{aligned} f^v(\sigma^*, \tau) &\leq \eta_v \qquad \text{for all } \tau \in \mathcal{S}_1 \\ f^v(\sigma, \tau^*) &\geq \eta_v \qquad \text{for all } \sigma \in \mathcal{S}_0 \end{aligned} \tag{2.2}$$

then the game $(\mathcal{G}, v)$ has a minimax equilibrium and (σ^*, τ^*) are called optimal strategies for P_0 and P_1 respectively. η_v is called the value of the game.

The first inequality says that if P_0 follows σ^*, then payoff would always (against any play of P_1) be below η_v, while the second inequality says that if P_1 follows τ^* then the payoff always be above η_v. In particular

$$f^v(\sigma^*, \tau^*) = \eta_v$$

There may be many optimal strategy pairs (σ^*, τ^*) but value η_v must be unique. This is because (2.2) are equivalent to

$$\eta_v = \min_{\sigma \in \mathcal{S}_0} \sup_{\tau \in \mathcal{S}_1} f^v(\sigma, \tau) = \max_{\tau \in \mathcal{S}_1} \inf_{\sigma \in \mathcal{S}_0} f^v(\sigma, \tau) \tag{2.3}$$

(The distinction between max vs sup (or min vs inf) is that the optimal value must be attained in the former, but may not be attained in the latter)

The RHS of (2.3) is the optimal payoff in the game where P_1 declares his strategy to P_0 before the game begins (which is always disadvantageous to P_1). Similarly LHS of (2.3) is the optimal payoff when P_0 declares his strategy to P_1 in advance (which is always advantageous to P_1). The equality means that both can reveal their optimal strategies (σ^*, τ^*) to the opponent and none will have an incentive to change their strategy.

Not every game has a minimax equilibrium, however a weaker version of (2.3) (also called as ϵ-equilibrium) holds for a large class of payoffs (see [19])

$$\inf_{\sigma \in \mathcal{S}_0} \sup_{\tau \in \mathcal{S}_1} f^v(\sigma, \tau) = \sup_{\tau \in \mathcal{S}_1} \inf_{\sigma \in \mathcal{S}_0} f^v(\sigma, \tau)$$

For the games to be discussed, (2.3) will explicitly be shown. The minimax will exist from each $v \in V$. Since (σ^*, τ^*) are history dependent they can be chosen independent of v. Hence we can bundle up the values and associate to $\mathcal{G}$ a value vector $\eta \in \mathbb{R}^V$ given by $\eta = (\eta_v)_{v \in V}$. Let $f(\sigma, \tau) = (f^v(\sigma, \tau))_{v \in V}$, then we have (inequalities are coordinate-wise) –

$$\begin{aligned} f(\sigma^*, \tau) &\leq \eta \qquad \forall \tau \in \mathcal{S}_1 \\ f(\sigma, \tau^*) &\geq \eta \qquad \forall \sigma \in \mathcal{S}_0 \end{aligned} \tag{2.4}$$

2.3 Finite Games

We will now start with a simple yet important class of payoffs, and show the existence of minimax and optimal strategies for it.

A payoff function f is finitely determined if there is a $N \in \mathbb{N}$ so that f only depends on the outcomes in the first N rounds. More precisely

$$\forall \alpha, \beta \in \mathcal{P}^\omega \; \alpha|_N = \beta|_N \implies f(\alpha) = f(\beta)$$

where $(v_0 v_1 \ldots)|_m = v_0 v_1 \ldots v_m$.

If f is finitely determined then call $\mathcal{G} = (G, f)$ a finite duration game. We can assume that the game stops after N rounds have been played and the payoff is given by

$$f : \mathcal{P}^N \to \mathbb{R}$$

This is defined as the payoff for an arbitrary infinite extension, which is well defined since G has no dead ends and f depends only on first N rounds.

The following theorem is well known - it is sometimes called the backward induction technique. The backward induction, however, is not very explicit in the proof presented here.

Theorem 2.1. *Let $\mathcal{G} = (G, f)$ be a finite game. There is a $\eta \in \mathbb{R}^V$ and strategies (σ^*, τ^*) which satisfy* (2.4)

Proof. The proof is by induction on N – the duration of the game $\mathcal{G}$.

Base Case $N = 1$: Hence the game ends after the first move. Let

$$\eta_u = \begin{cases} \min\limits_{v \in N(u)} f(uv) & \text{if } u \in V_0 \\ \max\limits_{v \in N(u)} f(uv) & \text{if } u \in V_1 \end{cases}$$

And

$$\sigma^*(u) = \operatorname*{argmin}_{v \in N(u)} f(uv) \quad \text{if } u \in V_0$$
$$\tau^*(u) = \operatorname*{argmax}_{v \in N(u)} f(uv) \quad \text{if } u \in V_1$$

Then (2.2) follows for any v by considering the cases $v \in V_0$ and $v \in V_1$ separately. Hence (2.4) follows with $\eta = (\eta_v)_{v \in V}$

Inductive Case $N = k + 1$ with $k \geq 1$:

At the $k + 1^{\text{th}}$ round what will be the decision of the players? Suppose $v_0 v_1 \dots v_k$ has been played. Now if $v_k \in V_1$, since this is the last round, P_1 will chose an u which maximizes $f(v_0 v_1 \dots v_k u)$. Similarly if $v_k \in V_0$, then P_0 will choose an u which minimizes $f(v_0 v_1 \dots v_k u)$.

Motivated by this, define the k step game $\mathcal{G}' = (G, f')$ where

$$f'(v_0 v_1 \dots v_k) = \begin{cases} \min\limits_{u \in N(v_k)} f(v_0 v_1 \dots v_k u) & \text{if } v_k \in V_0 \\ \max\limits_{u \in N(v_k)} f(v_0 v_1 \dots v_k u) & \text{if } v_k \in V_1 \end{cases}$$

By induction hypothesis $\mathcal{G}'$ has a value vector η and optimal strategies (σ', τ'). Consider the aforementioned strategies for the $k + 1^{\text{th}}$ round -

$$\sigma_{k+1}(v_0 v_1 \dots v_k) = \operatorname*{argmin}_{u \in N(v_k)} f(v_0 v_1 \dots v_k u) \quad \text{if } v_k \in V_0$$
$$\tau_{k+1}(v_0 v_1 \dots v_k) = \operatorname*{argmax}_{u \in N(v_k)} f(v_0 v_1 \dots v_k u) \quad \text{if } v_k \in V_1$$

Let $\sigma^* = [\sigma', \sigma_{k+1}]$ be the strategy for P_0 which plays σ' for the first k rounds, and σ_{k+1} for the $k+1^{\text{th}}$ round. Analogously define $\tau^* = [\tau', \tau_{k+1}]$ for P_1.

The following shows that η is the value vector for $\mathcal{G}$ and (σ^*, τ^*) are the optimal strategies.

To prove the first inequality in (2.4), take any path $v_0v_1 \ldots v_{k+1}$ that conforms with σ^*. Then $v_0v_1 \ldots v_k$ must conform with σ'. Hence

$$f'(v_0v_1 \ldots v_k) \leq \eta_{v_0}$$

If $v_k \in V_0$ then $v_{k+1} = \sigma_{k+1}(v_0v_1 \ldots v_k)$, hence

$$f(v_0v_1 \ldots v_{k+1}) = f'(v_0v_1 \ldots v_k)$$

Otherwise $v_k \in V_1$ and from the definition of f'

$$f(v_0v_1 \ldots v_{k+1}) \leq f'(v_0v_1 \ldots v_k)$$

In any case $f(v_0v_1 \ldots v_{k+1}) \leq f'(v_0v_1 \ldots v_k) \leq \eta_{v_0}$. This shows

$$f(\sigma^*, \tau) \leq \eta \quad \forall \tau \in \mathcal{S}_1$$

The proof of the second inequality follows verbatim by reversing the inequalities, replacing σ by τ and interchanging V_0 and V_1. Hence this proves (2.4) for $\mathcal{G}$.

□

2.4 Win Lose games

When f is a $\{0, 1\}$ valued function, the game has a win/loss interpretation. P_1 wins the play (and P_0 loses) whenever the payoff is 1, otherwise P_1 loses the play (and P_0 wins). This game will be denoted by $\mathcal{G} = (G, Win)$ where

$$Win = \{\alpha \mid f(\alpha) = 1\} \subseteq \mathcal{P}^\omega$$

is the set for winning plays for P_1.

2.4.1 Determinacy

A strategy σ^* for P_0 is said to be a winning strategy from $v \in V$ if

$$\pi^v_{\sigma^*\tau} \notin Win \quad \forall \tau \in \mathcal{S}_1$$

Similarly a strategy τ^* for P_1 is a winning strategy from v if

$$\pi^v_{\sigma\tau^*} \in Win \quad \forall \sigma \in \mathcal{S}_0$$

Definition 2.5. Call a win-lose game *determined* if from every $v \in V$ either P_0 or P_1 has a winning strategy.

Notice that both the players cannot simultaneously have a winning strategy starting from v. It is also possible that neither of them has one, but axiom of choice is required to show this. See [15] for a very general theorem on determinacy by Martin.

Suppose the game $\mathcal{G} = (G, f)$ (where f is $\{0,1\}$ valued) has a minimax (2.2) starting from v, then $\eta_v \in \{0,1\}$ (because of (2.3)). If $\eta = 1$ then τ^* is a winning strategy for P_1, otherwise $\eta = 0$ and σ^* is a winning strategy P_0. Hence as a corollary of Theorem 2.1 we have –

Corollary 2.2. *If a winning condition Win is finitely determined ($\exists N \, \forall \alpha, \beta \in \mathcal{P}^\omega \; \alpha|_N = \beta|_N \implies \alpha \in Win \iff \beta \in Win$), then the game $\mathcal{G} = (G, Win)$ is determined.*

3 Parity

Parity games were introduced in [8] to solve the μ-calculus model checking problem. From the view of Complexity Theory, the problem of deciding the winner of a parity game is in $\mathsf{NP} \cap \mathsf{coNP}$ (so unlikely to be NP-complete), but the question of whether it has a polynomial time solution remains open.

Apart from Parity, the following chapters will also introduce Mean and Discounted Payoff Games. Although seemingly unrelated, each game can be reduced to the successive one. This is used in [12] to show that the decision problems for all of them are in $\mathsf{UP} \cap \mathsf{coUP}$ (here UP is the class of decision problems that have a unique polynomial size certificate, hence $\mathsf{P} \subseteq \mathsf{UP} \subseteq \mathsf{NP}$). Whether any of these games (i.e. their respective decision problems) has a polynomial time solution or not is again an open question.

3.1 Definition

Let $G = (V_0, V_1, E)$ be a graph and let (for some $M \in \mathbb{N}$)

$$p : V \mapsto \{0, 1, 2 \dots M\}$$

be an assignment of integer priorities to each vertex.

Parity game $\mathcal{G}_p = (G, Win_p)$ is a Win-Lose game (section 2.4) with

$$Win_p = \left\{ (v_0 v_1 \dots) \in \mathcal{P}^\omega \mid \left(\limsup_{i \geq 0} p(v_i) \right) \text{ is Odd} \right\} \tag{3.1}$$

Let the largest priority seen infinitely often, on a path π, be denoted by max-inf-priority of π. If max-inf-priority of π is odd then P_1 wins π; if the max-inf-priority is even then P_0 wins. Since V is finite, there's a vertex v which is visited infinitely often by π, and has the same priority as the max-inf-priority.

Notice that in contrast with finite games (section 2.3), the winning condition (3.1) only depends on long run behaviour. This property is also called prefix independence – for $w \in \mathcal{P}^*$ and $\alpha \in \mathcal{P}^\omega$ if $w \cdot \alpha \in \mathcal{P}^\omega$ then $\alpha \in Win_p \iff w \cdot \alpha \in Win_p$. However, due to path constraints in the graph G, the prefix might still be important as it can enable or disable certain long term behaviours.

3.2 Positional Determinacy

From every vertex $v \in V$ the parity game $(\mathcal{G}_p, v)$ is determined (section 2.5) – one can invoke the general determinacy theorem by Martin [15] to show this. However a stronger result was proved by Emerson [7] –

Theorem 3.1. *For any parity game $\mathcal{G}_p = (G, Win_p)$, there are positional strategies (σ^*, τ^*) and a partition of $V = W_0 \sqcup W_1$, so that σ^* is a winning strategy for P_0 from each $v \in W_0$ and τ^* is a winning strategy for P_1 from each $v \in W_1$.*

Hence, not only is the game $(\mathcal{G}_p, v)$ determined, the winner has a positional winning strategy independent of v (in the player's winning region).

We are interested in the following decision problem.

Decision Problem (PAR)**.** Given a graph G, priorities p, and a vertex $v \in V$, determine whether P_0 wins $(\mathcal{G}_p, v)$ or not.

Given $v \in V$ and positional strategies (σ, τ), since $\pi^v_{\sigma\tau}$ is ultimately periodic (page 5) we have

$$\pi^v_{\sigma\tau} \in Win_p \iff \left(\max_{v \in \mathrm{Cycle}(\pi^v_{\sigma\tau})} p(v) \right) \text{ is odd}$$

Hence given (σ, τ), the winner of $\pi^v_{\sigma\tau}$ can easily be decided by analysing the cycle reachable in $G_{\sigma\tau}$. Then by Theorem 3.1, one could find the winner of $(\mathcal{G}_p, v)$ by enumerating over all possible positional strategies (which are finitely many) of both the players, and find one which beats all the strategies of the opponent.

However something better can be done. Given a positional strategy σ, it is directly possible to check if this strategy is winning for P_0 in $(\mathcal{G}_p, v)$ or not. Look at the graph G_σ. σ is winning in $(\mathcal{G}_p, v)$ if and only if no run in G_σ starting at v has an odd max-inf-priority – this is the emptiness problem for parity word automata and can be solved in in $O\left((|V| + |E|) \log M\right)$ (see [14]). Hence to check if P_0 wins $(\mathcal{G}_p, v)$ one could guess a strategy σ (a polynomial size certificate) and check in polynomial time whether it is winning for P_0 or not. Hence the decision problem of whether P_0 wins $(\mathcal{G}_p, v)$ is in NP. This problem is also in coNP as one could apply the same procedure for P_1 (i.e. for "no" instances, guess a strategy for P_1 and verify it). Hence we have the following theorem from [22]

Theorem 3.2. *The decision problem PAR is in* $\mathsf{NP} \cap \mathsf{coNP}$

There are a couple of proofs of Theorem 3.1, some of which are constructive. In [13] the proof due to Zielonka [23] and McNaughton [16] is used to obtain sub-exponential algorithm $O(n^{\sqrt{n}})$ for deciding the winning regions. However whether a polynomial time algorithm exists is not known.

In the following, we will provide a proof from [1] of determinacy of Parity Games. We will not cover the existence of positional optimal strategies. As seen already, positional determinacy is an important property; it can be proven along the lines of the following proof using some more work (see [1]).

3.3 Finite Game

Now we introduce a finite game $\mathcal{G}_p^f = (G, Win_p^f)$ closely related to the Parity Game $\mathcal{G}_p = (G, Win_p)$. It is played on the same graph G but it stops as soon as a vertex repeats (which will happen within $|V|$ rounds). Assume that the resulting path is

$$\pi = v_0 v_1 \dots v_r \dots v_k v_{k+1}$$

with $v_r = v_{k+1}$ and all v_i's distinct for $i \leq k$. Denote the unique cycle formed by

$$\text{Cycle}(\pi) = v_r v_{r+1} \dots v_k \tag{3.2}$$

If the largest priority in Cycle(π) is odd then P_1 wins the play, otherwise P_0 wins. That is

$$\pi \in Win_p^f \iff \left(\max_{v \in \text{Cycle}(\pi)} p(v) \right) \text{ is Odd}$$

Notice that this is the same condition as the ultimately periodic path $v_0 \dots v_{r-1}(v_r \dots v_k)^\omega$ winning in $\mathcal{G}_p$. Hence $\mathcal{G}_p^f$ is the game that would result from $\mathcal{G}_p$ if both the players decided to play positional strategies. Since $\mathcal{G}_p^f$ is a finite game, by Corollary 2.2 starting at any $v \in V$ one of the players has a winning strategy in $(\mathcal{G}_p^f, v)$. Hindsight from Theorem 3.1 tells us that the same player will also be the winner of $(\mathcal{G}_p, v)$. We will show this without using Theorem 3.1 and hence prove determinacy of $(\mathcal{G}_p, v)$.

Theorem 3.3. *Given a (possibly history dependent) winning strategy ρ for P_i in $(\mathcal{G}_p^f, v)$, there is a finite memory winning strategy $\tilde{\rho}$ for P_i in $(\mathcal{G}_p, v)$.*

The following stack based technique will be used to construct the $\tilde{\rho}$. This will be useful later too. Notice that ρ is only defined (or relevant) for simple paths starting in v.

3.3.1 Stack based extension

Let ρ be a strategy for P_i defined on simple paths. The following describes $\tilde{\rho}$ – an extension of ρ to arbitrary plays using finite memory. At each stage P_i maintains a simple path of G in its memory. If the game starts at v_0, let $m_0 = v_0$. At the $k+1^{\text{th}}$ stage, if it is P_i's turn, play $v_{k+1} = \rho(m_k)$ (otherwise the opponent decides v_{k+1}), and update the memory as follows.
Assume

$$m_k = u_0 u_1 \dots u_r \quad r \geq 0$$

then

$$m_{k+1} = \begin{cases} u_0 u_1 \dots u_s & \text{if } u_s = v_{k+1} \text{ for some } s \le r \\ u_0 u_1 \dots u_r v_{k+1} & \text{otherwise} \end{cases}$$

In the first case $u_{s+1} \dots u_r v_{k+1}$ forms a cycle in G, which had to be eliminated for m_{k+1} to remain a simple path.

For each n, m_n is a simple path from v_0 to v_n which conforms with ρ. If $C_1, C_2 \dots C_{r_n}$ are the cycles eliminated by the n^{th} round then $m_n, C_1, C_2 \dots C_r$ forms a partition of $v_0 v_1 \dots v_n$

Proof of Theorem 3.3. Obtain $\tilde{\rho}$ from ρ by the above stack based extension. We will show that this strategy $\tilde{\rho}$ is winning for P_i in $(\mathcal{G}_p, v)$.

Assume ρ is a strategy for P_0. The proof for P_1 is similar. Let $\pi = v v_1 v_2 \dots$ be an infinite path that conforms with $\tilde{\rho}$. We will show that max-inf-priority for π is even. This shows that $\tilde{\rho}$ is winning for P_0 in $(\mathcal{G}_p, v)$.

Since ρ is winning for P_0 in $(\mathcal{G}_p^f, v)$, for any path that starts at v and conforms with ρ till the first cycle formed, the largest priority of the cycle will be even. In the stack based implementation of $\tilde{\rho}$ on π, each m_i conforms with ρ, hence the max priority in each of the eliminated cycles C_r will be even. Given this, the max-inf-priority of π cannot be odd. Indeed, let u be a vertex which is visited infinitely often in π, and has same priority as max-inf-priority of π. Then u must be a part of infinitely many C_i's. But if $p(u)$ is odd, then there will be a vertex in each of these cycles which has priority strictly greater than $p(u)$ (since max priority in every C_i is even). Since C_i's correspond to disjoint positions of π, this shows that the max-inf-priority of π is strictly greater than $p(u)$. A contradiction.

□

3.4 Summary

In this chapter we have introduced parity games and an equivalent finite duration game (section 2.3). A stack based technique (subsection 3.3.1) was used to extend winning strategies in the finite game to the parity game. Observe that if the original strategy is positional this extension will be the same positional strategy. This is used in [1] to give a complete proof of positional determinacy of parity games (Theorem 3.1). Positional determinacy also shows that the decision problem for parity games is in $\mathsf{NP} \cap \mathsf{coNP}$ (Theorem 3.2).

4 Mean payoff

Mean payoff is another infinite duration game, closely related to parity games. It was studied independently in [6] and [11].

4.1 Definition

Let $G = (V_0, V_1, E)$ be a graph and

$$w : E \mapsto \mathbb{Z}$$

be edge weights. Assume $|w(e)| \leq W$ for every $e \in E$.

The Mean payoff game is an infinite duration game $\mathcal{G}_w = (G, f_w)$ where we want the payoff

$$f_w(v_0 v_1 v_2 \ldots) = \lim_n \frac{1}{n} \sum_{i=0}^{n-1} w(v_i, v_{i+1}) \tag{4.1}$$

to be the mean weight on the infinite path. Unfortunately the limit in (4.1) may not always exist. We can also look at $\overline{\mathcal{G}_w} = (G, \overline{f_w})$ and $\underline{\mathcal{G}_w} = (G, \underline{f_w})$, where $\overline{f_w}$, $\underline{f_w}$ are obtained by replacing lim sup, lim inf in place of lim in (4.1). We will see that the choice will not matter for optimal play.

When the play π is ultimately periodic the limit in (4.1) exists and is equal to mean(Cycle(π)). Where

$$\text{mean}(v_1 v_2 \ldots v_k) = \frac{1}{k} \left(\sum_{i=1}^{k-1} w(v_i, v_{i+1}) + w(v_k, v_1) \right) \tag{4.2}$$

is the mean weight on the cycle $v_1 \ldots v_k \in \mathcal{P}^*$ (with $(v_k, v_1) \in E$). As a consequence when both P_0 and P_1 play finite memory strategies (σ, τ)

$$f_w(\sigma, \tau) = \overline{f_w}(\sigma, \tau) = \underline{f_w}(\sigma, \tau)$$

Just like the parity winning condition both $\overline{f_w}, \underline{f_w}$ are prefix independent (a payoff f is prefix independent if $f(\alpha) = f(w \cdot \alpha)$ for any $\alpha, w \cdot \alpha \in \mathcal{P}^\omega$).

4.2 Optimal strategies

The following theorem is the central result in [6].

Theorem 4.1. *There are positional strategies (σ^*, τ^*) for P_0, P_1 respectively and a value vector $\eta \in \mathbb{R}^V$ so that*

$$\begin{aligned} \overline{f_w}(\sigma^*, \tau) &\leq \eta \qquad \forall \tau \in \mathcal{S}_1 \\ \underline{f_w}(\sigma, \tau^*) &\geq \eta \qquad \forall \sigma \in \mathcal{S}_0 \end{aligned} \tag{4.3}$$

In other words starting from any v, P_0 can ensure that the lim sup of the means remains below η_v, while P_1 can ensure that the lim inf of the means remains above η_v. Call any (σ^*, τ^*) which satisfy (4.3) optimal strategies for $\mathcal{G}_w$ and η its value vector. When (σ^*, τ^*) is played, the limit in (4.1) exists and $f_w(\sigma^*, \tau^*) = \eta$

Since $\overline{f_w} \geq \underline{f_w}$ this shows that η is the minimax value for both $\overline{\mathcal{G}_w}$ and $\underline{\mathcal{G}_w}$ (hence η is unique), and the strategies (σ^*, τ^*) are optimal in these games too. It will follow from the proof later that if (σ, τ) are finite memory optimal strategies for $\overline{\mathcal{G}_w}$ (or $\underline{\mathcal{G}_w}$) they will also be optimal for $\mathcal{G}_w$ (i.e. they will satisfy (4.3)).

The following is a simple consequence of Theorem 4.1 which will be useful while approximating the value later.

Corollary 4.2. *Let η be the value of $\mathcal{G}_w$. Then for any $v \in V$, $\eta_v = \frac{n}{m}$ for some $n, m \in \mathbb{Z}$ with $|n| \leq W \cdot |V|$ and $1 \leq m \leq |V|$.*

Proof. Since (σ^*, τ^*) are positional

$$\eta_v = f_w^v(\sigma^*, \tau^*) = \text{mean}(\text{Cycle}(\pi^v_{\sigma^*\tau^*}))$$

is the average weight over some simple cycle in G. Hence it has the required form. □

We will prove a weaker version of Theorem 4.1 using the same technique used for Parity.

Theorem 4.3. *There are finite memory strategies (σ^*, τ^*) for P_0, P_1 and a value vector $\eta \in \mathbb{R}^V$ so that*

$$\begin{aligned} \overline{f_w}(\sigma^*, \tau) &\leq \eta \qquad \forall \tau \in \mathcal{S}_1 \\ \underline{f_w}(\sigma, \tau^*) &\geq \eta \qquad \forall \sigma \in \mathcal{S}_0 \end{aligned}$$

Just like in the case of parity, the complete proof of Theorem 4.1 can be obtained by doing some more work. In fact the presentation here, taken from [1], is the first part of a uniform proof for both Theorem 3.1 and Theorem 4.1.

4.3 Finite Game

Like in section 3.3, let us introduce a finite duration game $\mathcal{G}_w^f = (G, f'_w)$ related to $\mathcal{G}_w = (G, f_w)$. The game is played on G and stops the first time a vertex repeats. The payoff associated with such a path is

$$f'_w(\pi) = \text{mean}(\text{Cycle}(\pi))$$

Where Cycle(π) is the first cycle formed – see (3.2).

Notice that $\mathcal{G}_w^f$ is obtained from $\mathcal{G}_w$ when the players restrict to positional strategies. Since $\mathcal{G}_w^f$ is a finite game by Theorem 2.1 there is a payoff vector $\eta \in \mathbb{R}^V$ and optimal strategies (σ, τ) in $\mathcal{G}_w^f$. Note that for any path starting at v and conforming with σ till the first cycle formed, the mean value of the cycle will be $\leq \eta_v$. Similarly, for any path starting at v and conforming with τ till the first cycle formed, the mean value of the cycle will be $\geq \eta_v$. Now we will show that this η is also the value vector for $\mathcal{G}_w$.

Proof of Theorem 4.3. Let η be the value vector and (σ, τ) be the optimal strategies for $\mathcal{G}_w^f$. Use the stack based technique (subsection 3.3.1) to obtain σ^* from σ and τ^* from τ. We will now show that (4.3) is satisfied.

Let us show the first inequality. Take any path $\pi = v_0 v_1 \dots$ which conforms with σ^*. Then in the stack based implementation of σ^* on π, each m_i is a path starting at v_0 which confirms with σ. Hence each of the eliminated cycles C_i will have mean$(C_i) \leq \eta_{v_0}$. If by the n^{th} stage $m_n = u_0 \dots u_{s_n}$ and cycles $C_1, \dots, C_{r_n}$ have been eliminated, then

$$\sum_{i=0}^{n-1} w(v_i, v_{i+1}) = \sum_{i=0}^{s_n - 1} w(u_i, u_{i+1}) + \sum_{j=1}^{r_n} \text{sum}(C_j)$$

where

$$\text{sum}(v_1 v_2 \dots v_k) = \sum_{i=1}^{k-1} w(v_i, v_{i+1}) + w(v_k, v_1)$$

is the sum of the weights on the cycle. Hence if $|C|$ denotes the number of edges (or vertices) in C. We have

$$\begin{aligned}
\sum_{i=0}^{n-1} w(v_i, v_{i+1}) &= \sum_{i=0}^{s_n - 1} w(u_i, u_{i+1}) + \sum_{j=1}^{r_n} |C_j| \, \text{mean}(C_j) && (4.4) \\
&\leq s_n W + \eta_{v_0} \sum_{j=1}^{r_n} |C_j| && \text{since mean}(C_j) \leq \eta_{v_0} \\
&= s_n W + \eta_{v_0}(n - s_n) && \text{as } s_n + \textstyle\sum_{j=1}^{r_n} |C_j| = n \\
&= n\eta_{v_0} + s_n(W - \eta_{v_0})
\end{aligned}$$

But as m_n is a simple path, $|s_n| \leq |V|$; also $|\eta_{v_0}| \leq W$. Hence

$$\sum_{i=0}^{n-1} w(v_i, v_{i+1}) \leq n\eta_{v_0} + 2|V|W \tag{4.5}$$

Divide by n and let $n \to \infty$

$$\frac{1}{n}\sum_{i=0}^{n-1} w(v_i, v_{i+1}) \leq \eta_{v_0} + \frac{2|V|W}{n}$$

$$\limsup_n \left(\frac{1}{n}\sum_{i=0}^{n-1} w(v_i, v_{i+1})\right) \leq \limsup_n \left(\eta_{v_0} + \frac{2|V|W}{n}\right) = \eta_{v_0}$$

This shows

$$\overline{f_w}(v_0 v_1 \ldots) \leq \eta_{v_0}$$

Since $v_0 v_1 \ldots$ was any path that conformed with σ^* we have

$$\overline{f_w}(\sigma^*, \tau) \leq \eta \quad \forall \tau \in \mathcal{S}_1$$

To show the second inequality of (4.3) proceed similarly. Let $\pi = v_0 v_1 \ldots$ be a path the conforms with τ^*. In the stack based implementation of τ^* on π each of the eliminated cycles C_i will have mean$(C_i) \geq \eta_{v_0}$. Hence from (4.4)

$$\begin{aligned}\sum_{i=0}^{n-1} w(v_i, v_{i+1}) &\geq s_n(-W) + \eta_{v_0}(\sum_{j=1}^{r_n} |C_j|) \\ &\geq n\eta_{v_0} - s_n(\eta_{v_0} + W) \\ &\geq n\eta_{v_0} - 2|V|W \end{aligned} \tag{4.6}$$

Dividing by n and letting $n \to \infty$

$$\frac{1}{n}\sum_{i=0}^{n-1} w(v_i, v_{i+1}) \geq \eta_{v_0} - \frac{2|V|W}{n}$$

$$\liminf_n \left(\frac{1}{n}\sum_{i=0}^{n-1} w(v_i, v_{i+1})\right) \geq \liminf_n \left(\eta_{v_0} - \frac{2|V|W}{n}\right) = \eta_{v_0}$$

Hence for any $v_0 v_1 \ldots$ that conforms with τ^*

$$\underline{f_w}(v_0 v_1 \ldots) \geq \eta_{v_0}$$

This shows that

$$\underline{f_w}(\sigma, \tau^*) \geq \eta \qquad \forall \sigma \in \mathcal{S}_0$$

$\square$

In the above proof when (σ, τ) are positional strategies, (σ^*, τ^*) is just (σ, τ). The proof shows that when all the cycles reachable from v in G_σ have mean weight $\leq v$, then any infinite path π from v in G_σ will have $\overline{f_w}(\pi) \leq v$. Similarly if all the cycles reachable from v in G_τ have mean weight $\geq v$ then any infinite path π in G_τ from v will have $\underline{f_w}(\pi) \geq v$. This shows that if (σ, τ) are positional optimal strategies of $\overline{\mathcal{G}_w}$ (or $\underline{\mathcal{G}_w}$), they continue to be optimal strategies for $\mathcal{G}_w$. This argument can also be extended to the case when (σ, τ) are finite memory optimal strategies in $\overline{\mathcal{G}_w}$ (or $\underline{\mathcal{G}_w}$) by embedding the memory inside G.

4.4 Parity to Mean payoff

The corresponding finite games for Parity and Mean payoff are very similar. This helps establish a reduction from Parity to Mean payoff games.

Let G be a graph with priorities p. Consider the edge weights

$$w(u, v) = -(-|V|)^{p(u)} \tag{4.7}$$

w is defined so that for any simple cycle C in G, the max priority in C is odd then $\text{mean}(C) > 0$, and if the max priority is even then $\text{mean}(C) < 0$. Let v_r be the vertex with the largest priority $C = v_0 v_1 \ldots v_{k-1}$. Let $a = w(v_r, v_{r+1})$ (addition is modulo k). Notice that any other $w(v_i, v_{i+1})$ either equals a or has absolute value bounded by $\frac{a}{|V|}$. Since $|C| \leq |V|$, $\text{sum}(C)$ (and hence $\text{mean}(C)$) will have the same sign as $a = w(v_r, v_{r+1})$. When $p(v_r)$ is odd, $a > 0$; when it is even $a < 0$.

Now look at the finite games $(\mathcal{G}_p^f, v)$ and $(\mathcal{G}_w^f, v)$. P_1 wins $(\mathcal{G}_p^f, v)$ if and only if P_1 can ensure a payoff > 0 in $(\mathcal{G}_w^f, v)$. If η_v is the value of $(\mathcal{G}_w^f, v)$, this is the same condition as $\eta_v > 0$. Pass to their respective infinite games to obtain – P_1 has a winning strategy from $(\mathcal{G}_p, v)$ if and only if $\text{val}(\mathcal{G}_w, v) > 0$.

Hence if we define the following decision problem for the mean payoff game

Decision Problem (MP)**.** Given G, edge weights w and a $v \in V$, determine whether $val(\mathcal{G}_w, v) > 0$ or not.

Then this gives a reduction from PAR to MP. Notice that we can assume that

$$p : V \mapsto \{0, 1 \ldots 2|V|\}$$

Then the edge weights w (4.7) can be constructed in polynomial time. Hence this is a polynomial time reduction.

4.5 Summary

We have introduced Mean payoff games and the concept of optimal value and strategies. Similar to Parity games, an equivalent finite game was used to prove existence of optimal strategies (Theorem 4.3) – the proof uses the same stack based extension (subsection 3.3.1) as used for parity. The complete proof of positional determinacy (Theorem 4.1) can be obtained by doing some more work as presented in [1]. The similarity between the finite games for Parity and Mean payoff is used to give a polynomial time reduction from the decision problem for Parity (PAR) to the decision problem for mean payoff (MP). This brings a new set of techniques to tackle these problems. [24] describes a $O(|V|^3|E|W)$ time algorithm to find the value in the Mean payoff game, however this is only a pseudo polynomial time algorithm because of the linear dependence on W.

5 Discounted payoff

Discounted payoff game is yet another game related to Mean payoff (and hence Parity). Note that the payoffs for both Parity and Mean payoff were prefix independent; this is in stark contrast with the payoffs for finite games. Discounted payoff lies between these two - the payoff depends on the whole infinite path, but can be predicted to any desired accuracy by knowing a large enough prefix. This allows for the use of backward induction like technique (Theorem 2.1) to show (and compute) the minimax equilibrium.

Discounted payoff was introduced in [21] in a more general context of stochastic games. The following is a deterministic version of it.

5.1 Definition

Like in the Mean payoff, start with a graph G and edge weights

$$w : E \mapsto \mathbb{R}$$

Assume $|w(e)| \leq W$ for every $e \in E$.

The Discounted payoff game with parameter $0 < \lambda < 1$ is $\mathcal{G}_w^\lambda = (G, f_w^\lambda)$ with

$$f_w^\lambda(v_0 v_1 \ldots) = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i w(v_i, v_{i+1})$$

Since w is bounded and $|\lambda| < 1$, the series converges absolutely.

$\mathcal{G}_w^\lambda$ has an interpretation in terms of a stopping game – after each round (for instance after the n^{th} round) a coin (of bias λ) is tossed. With probability $(1 - \lambda)$ the game stops and the payoff is $w(v_{n-1}, v_n)$, otherwise with probability λ the game proceeds to the next round. Then $f_w^\lambda(v_0 v_1 \ldots)$ is the expected payoff for the path.

f_w^λ satisfies the following recursive equation

$$f_w^\lambda(v_0 v_1 \ldots) = (1 - \lambda) w(v_0, v_1) + \lambda f_w^\lambda(v_1 v_2 \ldots) \tag{5.1}$$

5.2 Optimal strategies

Fix w and λ for the rest of this section – their dependency might be suppressed at some places. The following presentation is adopted from the original proof by Shapely in [21] and can also be found in [24].

Theorem 5.1. *$\mathcal{G}_w^\lambda$ satisfies the minimax equilibrium (2.4). Moreover the value vector $\eta \in \mathbb{R}^V$ is the unique fixed point of*

$$F : \mathbb{R}^V \mapsto \mathbb{R}^V \tag{5.2}$$
$$(F(x))_u = \begin{cases} \min\limits_{v \in N(u)} (1-\lambda)w(u,v) + \lambda \eta_v & \textit{if } u \in V_0 \\ \max\limits_{v \in N(u)} (1-\lambda)w(u,v) + \lambda \eta_v & \textit{if } u \in V_1 \end{cases}$$

and

$$\begin{aligned} \sigma_\eta(u) &= \operatorname*{argmin}_{v \in N(u)} (1-\lambda)w(u,v) + \lambda x_v \quad \textit{if } u \in V_0 \\ \tau_\eta(u) &= \operatorname*{argmax}_{v \in N(u)} (1-\lambda)w(u,v) + \lambda x_v \quad \textit{if } u \in V_1 \end{aligned} \tag{5.3}$$

are positional optimal strategies.

Proof. For any $x \in \mathbb{R}^V$ and every $n \in \mathbb{N}$, consider the n step game $\mathcal{H}_x^n = (G, h_x^n)$ where

$$h_x^n : \mathcal{P}^n \mapsto \mathbb{R}$$
$$h_x^n(v_0 v_1 \dots v_n) = (1-\lambda)\left(\sum_{i=0}^{n-1} \lambda^i w(v_i, v_{i+1})\right) + \lambda^n x_{v_n}$$

For a fixed x, $\mathcal{H}_x^n$ approximates the game $\mathcal{G}_w^\lambda$ for large n. Moreover there is a simple recursion to find the value and optimal strategies for $\mathcal{H}_x^n$.

For $n = 1$ Consider the game $\mathcal{H}_x^1 = (G, h_x^1)$ with

$$h_x^1 : \mathcal{P}^1 \mapsto \mathbb{R}$$
$$h_x^1(v_0, v_1) = (1-\lambda)w(v_0, v_1) + \lambda x_{v_1}$$

Define

$$\begin{aligned} \sigma_x(u) &= \operatorname*{argmin}_{v \in N(u)} (1-\lambda)w(u,v) + \lambda x_v \quad \text{if } u \in V_0 \\ \tau_x(u) &= \operatorname*{argmax}_{v \in N(u)} (1-\lambda)w(u,v) + \lambda x_v \quad \text{if } u \in V_1 \end{aligned}$$

It is straightforward to check that $F(x)$ is the value vector and (σ_x, τ_x) are the optimal strategies for $\mathcal{H}_x^1$.

For $n \geq 2$ We have

$$h_x^n(v_0v_1\dots v_n) = (1-\lambda)w(v_0,v_1) + \lambda h_x^{n-1}(v_1v_2\dots v_n)$$

Let y is the value vector of $\mathcal{H}_x^{n-1}$ and (σ',τ') be optimal strategies. Then $F(y)$ will be the value vector of $\mathcal{H}_x^n$. Let $\sigma^* = [\sigma_y,\sigma']$ be the strategy for P_0 which plays σ_y in the first round and σ' after that. Similarly let $\tau^* = [\tau_y,\tau']$ be the corresponding strategy for P_1. Then (σ^*,τ^*) will be the optimal strategies for $\mathcal{H}_x^n$.
To show this, let $v_0v_1\dots v_n$ be a path conforming with σ^*. Since $v_1\dots v_n$ conforms with σ'

$$h_x^{n-1}(v_1v_2\dots v_n) \leq y$$

hence

$$\begin{aligned} h_x^n(v_0v_1\dots v_n) &= (1-\lambda)w(v_0,v_1) + \lambda h_x^{n-1}(v_1v_2\dots v_n) \\ &\leq (1-\lambda)w(v_0,v_1) + \lambda y \\ &\leq F(y)_{v_0} \end{aligned}$$

The last inequality follows as v_0v_1 conforms with σ_y. Similarly if $v_0v_1\dots v_n$ conforms with τ^*, $h_x^n(v_0v_1\dots v_n) \geq F(y)_{v_0}$. Hence $F(y)$ is the value vector for $\mathcal{H}_x^n$.

Hence by using induction this shows that $F^n(x)$ is the value vector for $\mathcal{H}_x^n$. Let $\sigma^* = [\sigma_{F^{n-1}(y)},\dots\sigma_{F(y)},\sigma_y]$ be the strategy that plays $\sigma_{F^{n-1}(y)}$ in the first round, $\sigma_{F^{n-2}(y)}$ in the second round and so on. Similarly let $\tau^* = [\tau_{F^{n-1}(y)},\dots\tau_y]$. Then (σ^*,τ^*) are optimal strategies for $\mathcal{H}_x^n$. Since $\mathcal{H}_x^n$ approximates $\mathcal{G}_w^\lambda$ as $n \to \infty$, we expect $F^n(x)$ to converge to the value vector for G_w^λ.

Now we will use this to show that $\mathcal{G}_w^\lambda$ has a minimax equilibrium. Consider the norm $\|.\|_\infty$ on $\mathbb{R}^V$ by

$$\|x\|_\infty = \max_{v\in V}|x_v|$$

Then (using $|\max_{i\in I} a_i - \max_{i\in I} b_i| \leq \max_{i\in I}|a_i - b_i|$ and similarly for min)

$$\|F(x) - F(y)\|_\infty \leq \lambda\,\|x-y\|_\infty$$

Hence F is a contraction mapping with coefficient $\lambda \in (0,1)$. It follows that (see [20, Chapter 9]) for any x, $\lim_n F^n(x) = \eta$ where $F(\eta) = \eta$ is the unique fixed point of F. Since $F^m(\eta) = \eta$ for any m, $\mathcal{H}_\eta^m$ has value η and $\sigma^* = \sigma_\eta, \tau^* = \tau_\eta$ are positional optimal strategies.

Let $v_0v_1\dots$ be a path in G which conforms with σ_η then

$$h_x^n(v_0v_1\dots v_n) = (1-\lambda)\left(\sum_{i=0}^{n-1}\lambda^i w(v_i,v_{i+1})\right) + \lambda^n\eta_{v_n} \leq \eta_{v_0}$$

for every n. Letting $n \to \infty$

$$f_w^\lambda(v_0 v_1 \ldots) \leq \eta_{v_0}$$

Similarly when $v_0 v_1 \ldots$ conforms with τ_η

$$f_w^\lambda(v_0 v_1 \ldots) \geq \eta_{v_0}$$

This shows that η is the value vector for $\mathcal{G}_w^\lambda$ and (σ_η, τ_η) are positional optimal strategies. □

Like in the Mean payoff case, positional determinacy gives us constraints on the optimal value.

Corollary 5.2. *Suppose w is integer valued. Let η be the payoff vector for the game $\mathcal{G}_w^\lambda$ and let $v \in V$. Then $\eta_v = p(\lambda)/q(\lambda)$ for polynomials p, q of degree $\leq |V| + 1$ with integer coefficients bounded by $4|W|$.*

Proof. Let η be the value vector for $\mathcal{G}_w^\lambda$. Since (σ_η, τ_η) are positional optimal

$$\pi^v_{\sigma_\eta \tau_\eta} = v_0 v_1 \ldots v_{r-1}(v_r v_{r+1} \ldots v_k)^\omega$$

for some $k \leq |V|$ and $f_w^\lambda(\pi^v_{\sigma_\eta \tau_\eta}) = \eta_v$. But

$$\begin{aligned}
f_w^\lambda(\pi^v_{\sigma_\eta \tau_\eta}) &= (1-\lambda) \sum_{i=0}^{\infty} \lambda^i w(v_i, v_{i+1}) \\
&= (1-\lambda) \left[\sum_{i=0}^{r-1} \lambda^i w(v_i, v_{i+1}) + \left(\sum_{i=r}^{k} \lambda^i w(v_i, v_{i+1}) \right) (1 + \lambda^{k-r+1} + \lambda^{2(k-r+1)} \ldots) \right] \\
&= (1-\lambda) \left[\sum_{i=0}^{r-1} \lambda^i w(v_i, v_{i+1}) + \left(\sum_{i=r}^{k} \lambda^i w(v_i, v_{i+1}) \right) \frac{1}{1 - \lambda^{k-r+1}} \right] \\
&= \frac{p(\lambda)}{q(\lambda)}
\end{aligned}$$

p, q are polynomial of degree $\leq |V| + 1$ (take $q(\lambda) = 1 - \lambda^{k-r+1}$) and all the coefficients are integers bounded modulus by $4W$. Note that coefficients not only depend on v, w but even on λ (via (σ_η, τ_η)). □

5.3 Mean payoff to Discounted payoff

For any bounded sequence $(a_i)_{i\in\mathbb{N}}$

$$\begin{aligned}(1-\lambda)\sum_{i=0}^{\infty} a_i\lambda^i &= (1-\lambda)^2\frac{1}{(1-\lambda)}\sum_{i=0}^{\infty} a_i\lambda^i \\ &= (1-\lambda)^2(1+\lambda+\lambda^2+\ldots)(\sum_{i=0}^{\infty} a_i\lambda^i) \\ &= (1-\lambda)^2\sum_{k=0}^{\infty}\left(\sum_{i=0}^{k} a_i\right)\lambda^k\end{aligned}$$

Hence we have

$$(1-\lambda)\sum_{i=0}^{\infty} a_i\lambda^i = (1-\lambda)^2\sum_{k=0}^{\infty}\left(\sum_{i=0}^{k} a_i\right)\lambda^k \tag{5.4}$$

And a special case (put $a_i = 1$ for each i)

$$(1-\lambda)^2\sum_{k=0}^{\infty}(k+1)\lambda^k = 1 \tag{5.5}$$

Using (5.4) and (5.5) one can prove the following.

Theorem 5.3. *Let* $u_\lambda = (1-\lambda)\sum_{i=0}^{\infty} a_i\lambda^i$ *and assume* $\lim_{n\to\infty}\frac{1}{n+1}\sum_{i=0}^{n} a_i = \alpha$. *Then* $\lim_{\lambda\to 1^-} u_\lambda = \alpha$.

Hence as $\lambda \to 1$ we expect the discounted game $\mathcal{G}_w^\lambda$ to approximate the mean payoff game $\mathcal{G}_w$. Now we will show this.

Theorem 5.4. *Consider a graph* (G, w) *with edge weights bounded (in modulus) by* $W > 0$. *Denote by* $\mathcal{G}_w$ *the Mean payoff game on this graph and by* $\mathcal{G}_w^\lambda$ *the discounted game with parameter* λ. *Let* η *be the value vector for* $\mathcal{G}_w$ *and* η_λ *be the value vector for* $\mathcal{G}_w^\lambda$. *Then*

$$\|\eta - \eta_\lambda\|_\infty \leq 2|V|W(1-\lambda)$$

Proof. From (5.4)

$$\begin{aligned}f_w^\lambda(v_0v_1\ldots) &= (1-\lambda)\sum_{i=0}^{\infty}\lambda^i w(v_i, v_{i+1}) \\ &= (1-\lambda)^2\sum_{k=0}^{\infty}\left(\sum_{i=0}^{k} w(v_i, v_{i+1})\right)\lambda^k \qquad (5.6)\end{aligned}$$

Let (σ^*, τ^*) be the stack based optimal strategies for $\mathcal{G}_w$ and let $(\sigma_\lambda, \tau_\lambda)$ be the optimal strategies for $\mathcal{G}_w^\lambda$.

Let $\pi = v_0 v_1 \ldots$ be an infinite path in G that conforms with σ^*. Then using the same argument as used in the proof of Theorem 4.3 we have from (4.5)

$$\sum_{i=0}^{k} w(v_i, v_{i+1}) \leq (k+1)\eta_{v_0} + 2|V|W$$

Combining this with (5.6)

$$\begin{aligned} f_w^\lambda(\pi) &\leq (1-\lambda)^2 \sum_{k=0}^{\infty} \left((k+1)\eta_{v_0} + 2|V|W\right)\lambda^k \\ &= \eta_{v_0}(1-\lambda)^2 \sum_{k=0}^{\infty}(k+1)\lambda^k + 2|V|W(1-\lambda)^2(\sum_{k=0}^{\infty}\lambda^k) \\ &= \eta_{v_0} + 2|V|W(1-\lambda) \qquad\qquad \text{(Using (5.5) and } \textstyle\sum_{i=0}^{\infty}\lambda^i = \frac{1}{1-\lambda}) \end{aligned}$$

Since π was any path that conformed with σ^*, consider $\pi^v_{\sigma^*\tau_\lambda}$. This shows

$$(\eta_\lambda)_v \leq f_w^\lambda(\pi^v_{\sigma^*\tau_\lambda}) \leq \eta_v + 2|V|W(1-\lambda) \tag{5.7}$$

for every $v \in V$.

Similarly if $\pi = v_0 v_1 \ldots$ is an infinite path in G that conforms with τ^*, from (4.6)

$$\sum_{i=0}^{k} w(v_i, v_{i+1}) \geq (k+1)\eta_{v_0} - 2|V|W$$

Similar to the above, this combined with (5.6) for $\pi^v_{\sigma_\lambda\tau^*}$ will give

$$(\eta_\lambda)_v \geq f_w^\lambda(\pi^v_{\sigma_\lambda\tau^*}) \geq \eta_v - 2|V|W(1-\lambda) \tag{5.8}$$

(5.7) and (5.8) together show that

$$\|\eta - \eta_\lambda\|_\infty \leq 2|V|W(1-\lambda)$$

□

Consider the following decision problem for Discounted payoff games

Decision Problem (DISC)**.** Given a graph G, vertex v, edge weights w, a discount λ and threshold a t. Determine whether $\mathrm{val}(\mathcal{G}_w^\lambda, v) \geq t$[1] or not.

[1] We can assume $t = 0$ by considering $w'(e) = w(e) - t$ instead of w

Corollary 5.5. *Given (G, v) with an integer valued w, MP can be reduced to an instance of DISC in polynomial time.*

Proof. Let W be the maximum modulus among all the edge-weights in w. Set $\lambda = 1 - \frac{1}{8|V|^2 W}$ and $t = \frac{3}{4|V|}$. For this choice of λ by Theorem 5.4, $|\operatorname{val}(\mathcal{G}_w, v) - \operatorname{val}(\mathcal{G}_w^\lambda, v)| \leq \frac{1}{4|V|}$. Combine this with Corollary 4.2 to obtain

$$\operatorname{val}(\mathcal{G}_w, v) > 0 \implies \operatorname{val}(\mathcal{G}_w, v) \geq \frac{1}{|V|} \implies \operatorname{val}(\mathcal{G}_w^\lambda, v) \geq \frac{3}{4|V|}$$

and

$$\operatorname{val}(\mathcal{G}_w^\lambda, v) \geq \frac{3}{4|V|} \implies \operatorname{val}(\mathcal{G}_w, v) \geq \frac{1}{2|V|} > 0$$

this shows that $\operatorname{val}(\mathcal{G}_w, v) > 0 \iff \operatorname{val}(\mathcal{G}_w^\lambda) \geq t$. Computing W, λ and t can be done time polynomial in the input size – hence this is a polynomial time reduction. □

5.4 Complexity

By Theorem 5.1, once the unique fixed point $\eta = F(\eta)$ is found, finding the optimal strategies (σ_η, τ_η) is easy. Hence to solve $\mathcal{G}_w^\lambda$ one has to find the fixed point of F.

By Corollary 5.2, the value vector has a polynomial representation in the input size. Hence one can guess a possible value η' (a polynomial size certificate) for the value vector and verify that $F(\eta') = \eta'$. Since F has a unique fixed point there will be a unique guess which works. Hence this shows that

Theorem 5.6. *DISC is in* $\mathsf{UP} \cap \mathsf{coUP}$.

> UP is the class of problems accepted by a polynomial time unambiguous Turing Machine. An unambiguous Turing machine is a nondeterministic Turing machine which has at most one accepting run. Hence $P \subseteq \mathsf{UP} \subseteq \mathsf{NP}$.

As mentioned already, for any x, $\lim_n F^n(x) = \eta$. Hence another way to compute η is to compute $F^n(0)$ for large n to get to the desired accuracy.

5.5 Summary

We have defined Discounted payoff games and shown that they have positional optimal strategies (Theorem 5.1). The value vector is the fixed point of the operator F and can be computed by taking repeated iterates. Next we show that the Discounted payoff value vector approximates the value vector for the corresponding Mean payoff game as $\lambda \to 1$ (Theorem 5.4). This is used to give a polynomial time reduction from the decision problem for Mean payoff MP to that for Discounted payoff DISC (Corollary 5.5). Finally we show that DISC is in $\mathsf{UP} \cap \mathsf{coUP}$ (Theorem 5.6).

6 Conclusion

We have introduced Parity, Mean payoff and Discounted payoff games. All of these games fit into the framework presented in Chapter 2 and have positional optimal strategies. In fact, [10] provides sufficient conditions (called fairly mixing property) on the payoff function f, under which the game (G, f) will have positional optimal strategies, and shows that Parity, Mean and Discounted payoffs satisfy those conditions.

Both Parity and Mean payoff games have prefix independent payoffs. Discounted payoff (with discount parameter λ) is prefix dependent, but the dependence on the prefix decreases as $\lambda \to 1$. On the other hand for a fixed $\lambda < 1$, the discounted payoff can be determined to any desired accuracy by knowing a large enough prefix. As a result, for a fixed λ, finding the value for the discounted game $\mathcal{G}_w^\lambda$ is easy, but as $\lambda \to 1$ the Discounted payoff approximates the Mean payoff (Theorem 5.4) and the problem becomes difficult.

The decision problem for Parity games PAR can be reduced in polynomial time to the decision problem for Mean payoff games MP (Section 4.4). This was made possible by the reduction between their corresponding finite games (which are very similar). Using the relation between Mean and Discounted payoff games MP can be reduced in polynomial time to the decision problem for Discounted payoff games DISC (Corollary 5.5)

Hence in this order – PAR, MP, DISC, each problem is harder than the previous. By Theorem 5.6, DISC (the hardest of them) is in $\mathsf{UP} \cap \mathsf{coUP}$ – hence each of them is in $\mathsf{UP} \cap \mathsf{coUP}$. Whether any of them have a polynomial time solution or not, is not known. There have been many attempts at better algorithms for PAR (see [22, Chap 7]), but recently [9] provided exponential lower bounds for some of the approaches which seemed promising.

From here one could look at Simple Stochastic Games (SSG) [5]. [24] provides a reduction from DISC to SSG. Hence SSG are harder than all the games presented here, however they too are in $\mathsf{UP} \cap \mathsf{coUP}$ (and have positional strategies). [4] also provides a direct reduction from PAR to SSG.

Bibliography

[1] Henrik Björklund, Sven Sandberg, and Sergei Vorobyov. "Memoryless determinacy of parity and mean payoff games: a simple proof". In: *Theoretical Computer Science* 310.1 (Jan. 2004), pp. 365–378. ISSN: 03043975. DOI: 10.1016/S0304-3975(03)00427-4. URL: http://linkinghub.elsevier.com/retrieve/pii/S0304397503004274.

[2] Egon Börger, Erich Grädel, and Yuri Gurevich. *The classical decision problem.* Springer Science & Business Media, 2001.

[3] J. Richard Büchi and Lawrence H. Landweber. "Solving sequential conditions by finite-state strategies". In: *Transactions of the American Mathematical Society* 138 (1969), pp. 295–311. ISSN: 0002-9947, 1088-6850. DOI: 10.1090/S0002-9947-1969-0280205-0. URL: http://www.ams.org/tran/1969-138-00/S0002-9947-1969-0280205-0/.

[4] Krishnendu Chatterjee and Nathanaël Fijalkow. "A reduction from parity games to simple stochastic games". In: *Electronic Proceedings in Theoretical Computer Science* 54 (June 4, 2011), pp. 74–86. ISSN: 2075-2180. DOI: 10.4204/EPTCS.54.6. URL: http://arxiv.org/abs/1106.1232v1.

[5] Anne Condon. "The complexity of stochastic games". In: *Information and Computation* 96.2 (Feb. 1992), pp. 203–224. ISSN: 08905401. DOI: 10.1016/0890-5401(92)90048-K. URL: http://linkinghub.elsevier.com/retrieve/pii/089054019290048K.

[6] Andrzej Ehrenfeucht and Jan Mycielski. "Positional strategies for mean payoff games". In: *International Journal of Game Theory* 8.2 (1979), pp. 109–113. URL: http://link.springer.com/article/10.1007/BF01768705.

[7] E. Allen Emerson. "Automata, tableaux, and temporal logics". In: *Logics of Programs*. Springer, 1985, pp. 79–88. URL: http://link.springer.com/content/pdf/10.1007/3-540-15648-8_7.pdf.

[8] E.A. Emerson and C.S. Jutla. "Tree automata, mu-calculus and determinacy". In: IEEE Comput. Soc. Press, 1991, pp. 368–377. ISBN: 978-0-8186-2445-2. DOI: 10.1109/SFCS.1991.185392. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=185392.

[9] Oliver Friedmann. "Exponential Lower Bounds for Solving Infinitary Payoff Games and Linear Programs". July 2011. URL: http://nbn-resolving.de/urn:nbn:de:bvb:19-132940.

[10] Hugo Gimbert and Wiesław Zielonka. "When Can You Play Positionally?" In: *Mathematical Foundations of Computer Science 2004*. Ed. by Jiří Fiala, Václav Koubek, and Jan Kratochvíl. Lecture Notes in Computer Science 3153. DOI: 10.1007/978-3-540-28629-5_53. Springer Berlin Heidelberg, Aug. 22, 2004, pp. 686–697. ISBN: 978-3-540-22823-3 978-3-540-28629-5. URL: http://link.springer.com/chapter/10.1007/978-3-540-28629-5_53.

[11] V.A. Gurvich, A.V. Karzanov, and L.G. Khachivan. "Cyclic games and an algorithm to find minimax cycle means in directed graphs". In: *USSR Computational Mathematics and Mathematical Physics* 28.5 (Jan. 1988), pp. 85–91. ISSN: 00415553. DOI: 10.1016/0041-5553(88)90012-2. URL: http://linkinghub.elsevier.com/retrieve/pii/0041555388900122.

[12] Marcin Jurdziński. "Deciding the winner in parity games is in UP∩ co-UP". In: *Information Processing Letters* 68.3 (1998), pp. 119–124. URL: http://www.sciencedirect.com/science/article/pii/S0020019098001501.

[13] Marcin Jurdzinski, Mike Paterson, and Uri Zwick. "A Deterministic Subexponential Algorithm for Solving Parity Games". In: *SIAM Journal on Computing* 38.4 (Jan. 2008), pp. 1519–1532. DOI: 10.1137/070686652. URL: http://epubs.siam.org/doi/abs/10.1137/070686652.

[14] Valerie King, Orna Kupferman, and Moshe Y. Vardi. "On the complexity of parity word automata". In: *Foundations of Software Science and Computation Structures*. Springer, 2001, pp. 276–286. URL: http://link.springer.com/chapter/10.1007/3-540-45315-6_18.

[15] Donald A. Martin. "Borel Determinacy". In: *Annals of Mathematics* 102.2 (1975), pp. 363–371. ISSN: 0003-486X. DOI: 10.2307/1971035. URL: http://www.jstor.org/stable/1971035.

[16] Robert McNaughton. "Infinite games played on finite graphs". In: *Annals of Pure and Applied Logic* 65.2 (1993), pp. 149–184. URL: http://www.sciencedirect.com/science/article/pii/016800729390036D.

[17] Michael O. Rabin. "Decidability of second-order theories and automata on infinite trees". In: *Transactions of the american Mathematical Society* 141 (1969), pp. 1–35. URL: http://www.jstor.org/stable/1995086.

[18] J. Richard Büchi. "Symposium on Decision Problems: On a Decision Method in Restricted Second Order Arithmetic". In: *Studies in Logic and the Foundations of Mathematics*. Vol. 44. Elsevier, 1966, pp. 1–11. ISBN: 978-0-8047-0096-2. URL: http://linkinghub.elsevier.com/retrieve/pii/S0049237X09705646.

[19] Stéphane Le Roux and Arno Pauly. "Equilibria in multi-player multi-outcome infinite sequential games". In: *arXiv preprint arXiv:1401.3325* (2014). URL: http://arxiv.org/abs/1401.3325.

[20] Walter Rudin. *Principles of mathematical analysis*. Vol. 3. McGraw-Hill New York, 1964.

[21] Lloyd S. Shapley. “Stochastic games”. In: *Proceedings of the National Academy of Sciences* 39.10 (1953), pp. 1095–1100. URL: http://www.pnas.org/content/39/10/1095.extract.

[22] Wolfgang Thomas, Thomas Wilke, et al. *Automata, logics, and infinite games: a guide to current research*. Vol. 2500. Springer Science & Business Media, 2002.

[23] Wieslaw Zielonka. “Infinite games on finitely coloured graphs with applications to automata on infinite trees”. In: *Theoretical Computer Science* 200.1 (June 1998), pp. 135–183. ISSN: 03043975. DOI: 10.1016/S0304-3975(98)00009-7. URL: http://linkinghub.elsevier.com/retrieve/pii/S0304397598000097.

[24] Uri Zwick and Mike Paterson. “The complexity of mean payoff games on graphs”. In: *Theoretical Computer Science* 158.1 (May 1996), pp. 343–359. ISSN: 03043975. DOI: 10.1016/0304-3975(95)00188-3. URL: http://linkinghub.elsevier.com/retrieve/pii/0304397595001883.